
RTA-OS3.1

User Guide

Copyright

The data in this document may not be altered or amended without special notification from ETAS GmbH. ETAS GmbH undertakes no further obligation in relation to this document. The software described in it can only be used if the customer is in possession of a general license agreement or single license. Using and copying is only allowed in concurrence with the specifications stipulated in the contract. Under no circumstances may any part of this document be copied, reproduced, transmitted, stored in a retrieval system or translated into another language without the express written permission of ETAS GmbH.

©Copyright 2008-2010 ETAS GmbH, Stuttgart.

The names and designations used in this document are trademarks or brands belonging to the respective owners.

Document: 10483-UG-1.0.0

Contents

1	Welcome to RTA-OS3.1!	12
1.1	Related Documents	12
1.2	About You	12
1.3	Document Conventions	13
1.4	References	13
2	Introduction	15
2.1	Features of the RTA-OS3.x Kernel	17
2.1.1	OSEK	17
2.1.2	AUTOSAR	20
2.1.3	Unique RTA-OS3.x Features	22
2.2	Summary	23
3	Development Process	24
3.1	Configuration	24
3.1.1	OS Configuration	25
3.1.2	RTA-TRACE Configuration	28
3.1.3	Build	29
3.1.4	Project Files	29
3.1.5	Error Checking	31
3.1.6	Generating Reports	34
3.2	Assistants	36
3.3	Library Generation	36
3.3.1	Preparing the Tool Chain	36
3.3.2	Understanding AUTOSAR Dependencies	38
3.3.3	Running rtaosgen	41
3.3.4	Building the library	41
3.3.5	Generated Files	42
3.4	Integration	43
3.4.1	Accessing the OS in your Source Code	43
3.4.2	Implementing Tasks and ISRs	44
3.4.3	Starting the OS	44
3.4.4	Interacting with the RTA-OS3.x	45
3.4.5	Compiling and Linking	45
3.5	Memory Images and Linker Files	45
3.5.1	Sections	45
3.5.2	The Linker Control File	47
3.6	Summary	48

4	Tasks	49
4.1	Scheduling	49
4.2	Basic and Extended Tasks	51
4.2.1	Task States	52
4.2.2	Task Priorities	54
4.2.3	Queued Task Activation	55
4.3	Conformance Classes	56
4.4	Maximizing Performance and Minimizing Memory	57
4.5	Task Configuration	57
4.5.1	Scheduling Policy	59
4.5.2	Queued Activation	59
4.5.3	Auto-starting Tasks	60
4.6	Stack Management	61
4.6.1	Working with Extended Tasks	62
4.6.2	Mandatory Stack Information	64
4.6.3	Specifying Task Stack Allocation	66
4.6.4	Optimizing the Extended Task context save	67
4.6.5	Handling Stack Overrun	70
4.7	Implementing Tasks	72
4.8	Activating Tasks	73
4.8.1	Direct Activation	74
4.8.2	Indirect Activation	74
4.9	Controlling Task Execution Ordering	75
4.9.1	Direct Activation Chains	75
4.9.2	Using Priority Levels	76
4.10	Co-operative Scheduling in RTA-OS3.x	77
4.10.1	Optimizing out the Schedule() API	78
4.11	Terminating Tasks	78
4.11.1	Optimizing Termination in RTA-OS3.x	79
4.12	The Idle Mechanism	80
4.13	Pre and Post Task Hooks	81
4.14	Saving Hardware Registers across Preemption	83
4.15	Summary	87

5	Interrupts	88
5.1	Single-Level and Multi-Level Platforms	88
5.2	Interrupt Service Routines	88
5.3	Category 1 and Category 2 Interrupts	89
5.3.1	Category 1 Interrupts	89
5.3.2	Category 2 Interrupts	89
5.4	Interrupt Priorities	89
5.4.1	User Level	92
5.4.2	OS Level	92
5.5	Interrupt Configuration	93
5.5.1	Vector Table Generation	94
5.6	Implementing Interrupt Handlers	95
5.6.1	Category 1 Interrupt Handlers	95
5.6.2	Category 2 Interrupt Handlers	96
5.6.3	Dismissing Interrupts	96
5.6.4	Writing Efficient Interrupt Handlers	97
5.7	Enabling and Disabling Interrupts	98
5.8	Saving Register Sets	99
5.9	The Default Interrupt	99
5.10	Summary	101
6	Resources	103
6.1	Resource Configuration	105
6.2	Resources on Interrupt Level	105
6.3	Using Resources	106
6.3.1	Nesting Resource Calls	108
6.4	Linked Resources	108
6.5	Internal Resources	110
6.6	Using Resources to Minimize Stack Usage	112
6.6.1	Internal Resources	113
6.6.2	Standard Resources	114
6.7	The Scheduler as a Resource	114
6.8	Choosing a Preemption Control Mechanism	116
6.9	Avoiding Race Conditions	117
6.10	Summary	118

7	Events	119
7.1	Configuring Events	119
7.1.1	Defining Waiting Tasks	120
7.2	Waiting on Events	120
7.2.1	Single Events	122
7.2.2	Multiple Events	122
7.2.3	Deadlock with Extended Tasks	123
7.3	Setting Events	124
7.3.1	Setting Events with an Alarm	125
7.3.2	Setting Events with a Schedule Table Expiry Point	125
7.4	Clearing Events	125
7.5	Simulating Extended Tasks with Basic Tasks	126
7.6	Summary	127
8	Counters	128
8.1	Configuring Counters	128
8.2	Counter Drivers	129
8.2.1	Software Counter Drivers	130
8.2.2	Hardware Counter Drivers	134
8.3	Accessing Counter Attributes at Runtime	137
8.3.1	Special Counter Names	138
8.4	Reading Counter Values	138
8.5	Tick to Time Conversions	139
8.6	Summary	141
9	Alarms	142
9.1	Configuring Alarms	142
9.1.1	Activating a Task	143
9.1.2	Setting an Event	144
9.1.3	Alarm Callbacks	144
9.1.4	Incrementing a Counter	145
9.2	Setting Alarms	147
9.2.1	Absolute Alarms	147
9.2.2	Relative Alarms	150
9.3	Auto-starting Alarms	151
9.4	Canceling Alarms	152
9.5	Working out when an Alarm will occur	153
9.6	Non-cyclic (aperiodic) Alarms	154
9.7	Summary	154

10	Schedule Tables	156
10.1	Configuring a Schedule Table	158
10.1.1	Synchronization	160
10.2	Configuring Expiry Points	160
10.3	The Schedule Table Assistant	162
10.4	Starting Schedule Tables	163
10.4.1	Absolute Start	163
10.4.2	Relative Start	165
10.4.3	Synchronized Start	166
10.5	Expiry Point Processing	167
10.6	Stopping Schedule Tables	167
10.6.1	Restarting a stopped schedule table	167
10.7	Switching Schedule Tables	169
10.8	Choosing an Synchronization Strategy	170
10.9	Implicit Synchronization	171
10.10	Explicit Synchronization	173
10.10.1	Counter Constraints	173
10.10.2	Specifying Synchronization Bounds	174
10.10.3	Startup	177
10.10.4	Synchronizing the Schedule Table	179
10.10.5	Synchronization	180
10.10.6	Disabling Synchronization	183
10.11	Schedule Table Status	184
10.12	Summary	184
11	Writing Hardware Counter Drivers	186
11.1	The Hardware Counter Driver Model	186
11.1.1	Interrupt Service Routine	187
11.1.2	Callbacks	188
11.2	Using Output Compare Hardware	189
11.2.1	Callbacks	190
11.2.2	Interrupt Handlers	194
11.2.3	Handling a Hardware modulus not equal to TickType	198
11.3	Free Running Counter and Interval Timer	203
11.3.1	Callbacks	204
11.3.2	ISR	205
11.4	Using Match on Zero Down Counters	206
11.4.1	Callbacks	207
11.4.2	Interrupt Handler	208
11.5	Software Counters Driven by an Interval Timer	209
11.6	Summary	209

12	Startup and Shutdown	211
12.1	From System Reset to StartOS()	211
12.1.1	Power-on or Reset	211
12.1.2	C Language Start-up Code	212
12.1.3	Running main()	213
12.2	Starting RTA-OS3.x	216
12.2.1	Startup Hook	217
12.2.2	Application Modes	217
12.3	Shutting Down RTA-OS3.x	222
12.3.1	Shutdown Hook	222
12.4	Restarting RTA-OS3.x	223
12.5	Summary	224
13	Error Handling	225
13.1	Centralized Error Handling - the ErrorHandler()	226
13.1.1	Configuring Advanced Error Logging	227
13.1.2	Working out which Task is Running	229
13.1.3	Working out which ISR is Running	230
13.1.4	Generating a Skeleton ErrorHandler()	230
13.2	In-line Error Handling	231
13.3	Conditional Inclusion of Error Checking Code	231
13.4	Summary	232
14	Measuring and Monitoring Stack Usage	233
14.1	Stack Monitoring	233
14.1.1	Setting Defaults	234
14.1.2	Configuring Stack Allocation per Task/ISR	234
14.2	Using the Os_Cbk_StackOverrunHook()	236
14.3	Measuring Stack Usage	238
14.3.1	Marking the Worst Case for Function Calls	239
14.4	Summary	241
15	Measuring and Monitoring Execution Time	242
15.1	Enabling Time Monitoring	242
15.1.1	Providing a Stopwatch	243
15.1.2	Scaling the Stopwatch	244
15.2	Automatic Measurement of Task and ISR Execution Times	244
15.3	Manual Time Measurement	246
15.4	Imprecise Computation	247
15.5	Monitoring Execution Times against Budgets	248
15.6	Summary	250

16	Using an ORTI-Compatible Debugger	251
16.1	Development Process	251
16.2	Intrusiveness	252
16.3	Validity	253
16.4	Interactions	253
16.5	Summary	254
17	RTA-TRACE2.1 Integration	255
17.1	Basic Configuration	256
17.2	Controlling RTA-TRACE2.1	257
17.2.1	Controlling with Objects are Traced	258
17.3	User-Defined Trace Objects	260
17.3.1	Tracepoints	261
17.3.2	Task Tracepoints	262
17.3.3	Intervals	262
17.3.4	Controlling which User-Defined Objects are Traced	264
17.3.5	Format Strings	266
17.4	ECU Links	269
17.4.1	Debugger Links	270
17.4.2	Serial Links	271
17.5	Summary	276
18	Protected Systems	277
18.1	Customized Protection Schemes	278
18.2	Handling Protection Errors	279
18.2.1	Fallbacks	282
18.3	Forced Termination	282
18.3.1	Tasks and ISRs	282
18.3.2	OS-Applications	283
18.3.3	Disabling Interrupt Sources	283
18.4	Generating a Skeleton ProtectionHook()	284
18.5	Summary	284

19	Timing Protection	285
19.1	What is a timing fault?	285
19.2	Meeting Deadlines at Runtime	286
19.2.1	Why deadline monitoring is not sufficient	286
19.2.2	Meeting Deadlines and Identifying Violators . .	287
19.3	Execution Budgets and Time Frames	289
19.3.1	Execution Budgets	289
19.3.2	Execution Budget Detection Modes	292
19.3.3	Time Frames	292
19.4	Configuring Timing Protection	294
19.5	Lock Budgets	295
19.6	Providing a Time Reference	296
19.6.1	Passive Detection	296
19.6.2	Active Detection	297
19.6.3	Rounding Errors	297
19.7	Function-Level Timing Protection	298
19.8	Summary	298
20	Service Protection	299
20.1	OS-Applications	300
20.2	Configuring OS-Applications	301
20.2.1	Ownership of OS objects	301
20.2.2	Access to OS objects	302
20.3	Accessing Applications	303
20.3.1	Hidden Accesses	305
20.3.2	Safety Risks	307
20.4	Identifying the Running OS-Application	307
20.5	Access and Ownership	308
20.6	Terminating OS-Applications	309
20.6.1	The Restart Task	310
20.6.2	Interaction with Accessing Applications	311
20.7	OS-Application Hooks	311
20.7.1	Startup Hook	312
20.7.2	Shutdown Hook	313
20.7.3	Error Hook	313
20.8	Trusted Functions	314
20.8.1	Calling Trusted Functions	314
20.8.2	Implementing Trusted Functions	315
20.9	Summary	316

21	Memory Protection	317
21.1	Hardware Limitations	318
21.2	AUTOSAR Limitations	319
21.3	Defining and Using Memory Regions	320
	21.3.1 Placing Task and OS-Application Data into the same section	320
	21.3.2 Placing Task and OS-Application Data into unique sections	322
21.4	Locating	324
21.5	Interacting with an MPU or MMU	325
	21.5.1 Default Implementations	326
21.6	Interaction with Stack Monitoring	326
21.7	Checking Memory Access	326
	21.7.1 Setting Access Permissions	327
	21.7.2 Checking Access Permissions	329
21.8	Summary	329
22	Contacting ETAS	331
22.1	Technical Support	331
22.2	General Enquiries	331
	22.2.1 ETAS Global Headquarters	331
	22.2.2 ETAS Local Sales & Support Offices	331

1 **Welcome to RTA-OS3.1!**

This user guide tells you how to use RTA-OS3.x to build AUTOSAR OS-based applications and is structured as follows:

Chapter 2 introduces you to RTA-OS3.x, covering what tools are provided, which standards are supported by the kernel and gives a brief overview of kernel features.

Chapter 3 takes you through the stages of development with RTA-OS3.x, including how to use the tools supplied to configure and build a kernel library and how to integrate it with your application.

Chapters 4-12 explain in detail how to configure the OS for each major class of OS object and how to use the kernel APIs that manipulate those objects at runtime.

Chapters 13-17 explain what to do when things go wrong. They cover how to detect erroneous use of the kernel API during development, how to check for stack overruns and timing faults and how to integrate with external debugging and profiling tools to get additional insight into how the OS is behaving at runtime.

Chapters 18-21 describe how to build partitioned systems where multiple applications are integrated with a common OS. They describe how OS objects are mapped to partitions and how to protect those partitions against erroneous access, timing and memory faults at runtime.

1.1 Related Documents

A complete technical reference to RTA-OS3.x can be found in the *Reference Guide* in the same directory as this user guide. Specific technical details about the implementation of RTA-OS3.x for your choice of compiler and target hardware (what we call a *port*) is contained in the *Target/Compiler Port Guide*. There is one *Target/Compiler Port Guide* for each installed RTA-OS3.x port that can be found in the root of the port's installation directory.

1.2 About You

You are a trained embedded systems developer who wants to build real-time applications using a preemptive operating system. You should have knowledge of the C programming language, including the compilation, assembling and linking of C code for embedded applications with your chosen toolchain. Elementary knowledge about your target microcontroller, such as the start address, memory layout, location of peripherals and so on, is essential.

You should also be familiar with common use of the Microsoft Windows 2000, Windows XP or Windows Vista operating systems, including installing software, selecting menu items, clicking buttons, navigating files and folders.

1.3 Document Conventions

The following conventions are used in this guide:

Choose File > Open .	Menu options are printed in bold, blue characters.
Click OK .	Button labels are printed in bold characters
Press <Enter>.	Key commands are enclosed in angle brackets.
The “Open file” dialog box appears	The names of program windows, dialog boxes, fields, etc. are enclosed in double quotes.
Activate(Task1)	Program code, header file names, C type names, C functions and API call names all appear in a monospaced typeface.
See Section 1.3 .	Hyperlinks through the document are shown in blue letters .



Functionality that is provided in RTA-OS but may not be portable to another AUTOSAR OS implementation is marked with the ETAS logo.



Caution! Notes like this contain important instructions that you must follow carefully in order for things to work correctly.

1.4 References

OSEK is a European automotive industry standards effort to produce open systems interfaces for vehicle electronics. For details of the OSEK standards, please refer to:

<http://www.osek-vdx.org>

AUTOSAR (AUTomotive Open System ARchitecture) is an open and standardized automotive software architecture, jointly developed by automobile manufacturers, suppliers and tool developers. For details of the AUTOSAR standards, please refer to:

<http://www.autosar.org>

2 Introduction

RTA-OS3.x is a statically configurable, preemptive, real-time operating system (RTOS) for use in high-performance, resource-constrained applications. RTA-OS3.x is a full implementation of the open-standard AUTOSAR OS R3.x (Scalability Class 4) and includes functionality that is fully compliant to Version 2.2.3 of the OSEK/VDX OS Standard.

The RTA-OS3.x kernel has been designed to be:

high performance - the kernel is very small and very fast. The memory footprint of the kernel and its run-time performance are class leading, making RTA-OS3.x particularly suitable for systems manufactured in large quantities, where it is necessary to meet very tight constraints on hardware costs and where any final product must function correctly.

RTA-OS3.x provides a number of unique optimizations that contribute to reductions in unit cost of systems. The kernel uses a single-stack architecture for all types of task. This provides significant RAM savings over a traditional stack-per-task model. Furthermore, careful application design can exploit the single-stack architecture to offer significant stack RAM savings.

The offline tools analyze your OS configuration and use this information to build the smallest and fastest kernel possible. Code that you are not going to use is excluded from the kernel to avoid wasting execution time and memory space.

real-time - conventional RTOS designs normally have unpredictable overheads, usually dependent upon the number of tasks and the state of the system at each point in time. This makes it difficult to guarantee real-time predictability - no matter how 'fast' the kernel is. In RTA-OS3.x the kernel is fast and all runtime overheads - such as switching to and from tasks, handling interrupts and waking up tasks - have low worst-case bounds and little or no variability of execution times. In many cases, context switching happens in constant execution time, meaning that RTA-OS3.x can be used for the development of hard real-time systems where responses must be made within specific timing deadlines. Meeting hard deadlines involves calculating the worst-case response time of each task and Interrupt Service Routine (ISR) and ensuring that everything runs on time, every time. RTA-OS3.x is a *real* RTOS because it meets the assumptions of fixed priority schedulability analysis.

portable - RTA-OS3.x is available for a wide variety of microcontroller/compiler combinations (or *port*). All ports share the same common RTA-OS3.x code, which comprises about 97% of the total kernel functional-

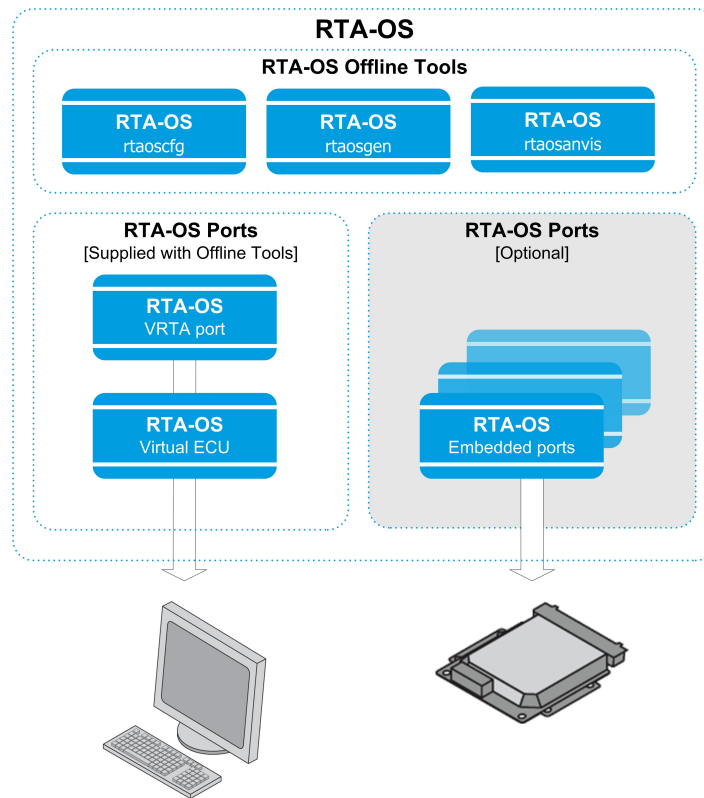


Figure 2.1: RTA-OS3.x Product Architecture

ity. The kernel is written in ANSI C that is MISRA-C 2004 compliant. A MISRA report for RTA-OS3.x can be generated by the offline tools.

As far as is possible, RTA-OS3.x does not impose on hardware - generally, there is no need to hand over control of hardware, such as the cache, watchdog timers and I/O ports. As a result of this, hardware can be used freely by your code, allowing legacy software to be integrated into the system.

The RTA-OS3.x product architecture is shown in Figure 2.1 and consists of:

- **rtaoscfg** a graphical configuration tool that reads and writes configurations in the AUTOSAR XML configuration language.
- **rtaosgen** a command-line tool for generating a RTA-OS3.x kernel library from your input configuration.
- Port plug-ins, one for each target/compiler combination for which you use RTA-OS3.x. You can install multiple ports at the same time and switch between them as desired. You can also install multiple versions of the

same port concurrently, allowing you to easily manage projects that use legacy compilers and/or microcontrollers.

- VRTA is a special port plug-in which provides the functionality of RTA-OS3.x on a standard Windows PC. This allows you to design and test application behavior without needing real target hardware¹. VRTA comes with a development kit that allows you to build Virtual ECUs that can simulate interrupts, I/O etc.

2.1 Features of the RTA-OS3.x Kernel

RTA-OS3.x builds on the proven technology of earlier ETAS operating systems which, to date, have been used in over 350 million ECUs worldwide. The kernel provides an implementation of the AUTOSAR OS R3.x open standard, a standard which subsumes features from the earlier OSEK OS standard². The kernel also provides a number of additional features which are unique to RTA-OS3.x. The following sections provide a short introduction to the standards and their features.

2.1.1 OSEK

OSEK is a European automotive industry standards effort to produce open systems interfaces for vehicle electronics. The full name of the project is OSEK/VDX. OSEK is an acronym formed from a phrase in German, which translates as *Open Systems and Corresponding Interfaces for Automotive Electronics*. VDX is based on a French standard (Vehicle Distributed eXecutive), which has now been merged with OSEK. OSEK/VDX is referred to as OSEK in this guide.

The goals of OSEK are to support portability and reusability of software components across a number of projects. This allows vendors to specialize in *Automotive Intellectual Property*, whereby a vendor can develop a purely software solution and run software in any OSEK-compliant ECU.

To reach this goal, however, detailed specifications of the interfaces to each non application-specific component are required. OSEK standards therefore include an Application Programming Interface (API) that abstracts away from the specific details of the underlying hardware and the configuration of the in-vehicle networks.

For further information see <http://www.osek-vdx.org>.

¹Note that VRTA is not a realtime environment, since it operates within the Windows operating system.

²For the sake of brevity, the term *AUTOSAR OS* is used throughout this document to refer to the combined AUTOSAR and OSEK OS standards.

OSEK OS

OSEK OS is the most mature and most widely used of the OSEK standards. OSEK OS has been adopted in all types of automotive ECUs, from powertrain, chassis and body to multi-media devices.

The most recent version of OSEK OS is 2.2.3, the third minor revision of the 2.2 standard originally introduced in September 2001. This version of OSEK OS is also part of the ISO17356 standard.

OSEK OS is entirely statically defined using an offline configuration language called OIL (OSEK Implementation Language). Since all objects are known at system generation time, implementations can be extremely small and efficient.

OSEK OS provides the following OS features:

Tasks are the main building block of OSEK OS systems. Unlike some other OS's, tasks in OSEK are not required to be *self-scheduling* (i.e. it is not necessary to place the body of the task inside an infinite loop³). There are four types of task in OSEK OS:

1. Basic tasks with unique priority and non-queued activation. These are the simplest form of task and ideally suited for hard real-time systems. Once a task is activated it must run and terminate before it can be activated again. This type of task cannot suspend itself mid-way through execution to wait for an event. In RTA-OS3.x these are called **BCC1** tasks because they correspond to OSEK OS's BCC1 conformance class (see Section 4.3 for more details about OSEK's Conformance Classes).
2. Basic tasks with shared priority and queued activation. These tasks can share priorities with other tasks in the system and do not need to terminate before being activated again. The OS queues pending task activations and runs the next activation when the current one has terminated. Like BCC1 tasks, this type of task cannot suspend itself mid-way through execution to wait for an event. In RTA-OS3.x these are called **BCC2** tasks because they correspond to OSEK OS's BCC2 conformance class.
3. Extended tasks with unique priority. An extended task is allowed to wait for events during execution (i.e. the task can *self suspend*). However, activations cannot be queued and the tasks must have unique priorities. In RTA-OS3.x these are called **ECC1** tasks because they correspond to OSEK OS's ECC1 conformance class.

³Though you can do this for the class of tasks called "extended tasks".

4. Extended tasks with shared priority. These are like ECC1 tasks but can share priorities with other tasks in the system. In this regard they are similar to BCC2 tasks. However, unlike BCC2 tasks, extended tasks cannot have queued activations. In RTA-OS3.x these tasks are called **ECC2** tasks.

A system can contain any combination of the above task types.

Scheduling Tasks can be scheduled either preemptively or non-preemptively and co-operative schedulers can be constructed easily.

Interrupts allow for the interaction of the OS with asynchronous external triggers. There are two types of interrupt in OSEK OS:

1. Category 1 interrupts are not handled by the OS;
2. Category 2 interrupts are handled by, and can interact with, the OS.

Resources are simple binary semaphores that allow you to provide mutual exclusion over critical sections shared between tasks and interrupts. Resources are managed by the OS using the *priority ceiling protocol* which guarantees freedom from deadlock and minimizes priority inversion at runtime⁴.

Counters and alarms are used to provide periodic (and aperiodic) scheduling of tasks. Counters, as the name suggests, count the occurrence of (domain specific) events and register values as 'ticks'. Alarms can be set to expire at run-time configurable count values, either at absolute count value or relative to the 'tick' value of the counter when the alarm is set.

Debugging Support is provided natively in the OS through the use of build levels. The OS provides two build levels:

1. **Standard** is 'lean and mean' and provides minimum error handling.
2. **Extended** is the 'debugging' build that provides extensive error detection facilities to check if you are using the OS correctly.

Debugging is also provided through the OSEK ORTI (OSEK Run-Time Interface) standard. This provides a common way for OS implementations, like RTA-OS3,0, to export symbol details to third-party debuggers

⁴Priority inversion is the situation where a low priority task is running in preference to a higher priority task. With priority ceiling protocol this situation can occur at most once each time a higher priority task is activated (and it is always at the start of execution) and is called the blocking time for the higher priority task. The blocking time is bounded by the longest time any single task shares data with the higher priority object - there is no cumulative blocking due to the interaction of lower priority tasks.

so that the debugger can display information about the internal state of the OS at runtime (e.g. which task is running, which tasks are ready to run etc.).

2.1.2 AUTOSAR

AUTOSAR (AUTomotive Open System ARchitecture) is an open and standardized automotive software architecture, jointly developed by automobile manufacturers, suppliers and tool developers worldwide.

AUTOSAR provides specifications for *Basic Software Modules* (BSW) such as operating systems, communication drivers, memory drivers and other micro-controller abstractions. The AUTOSAR standard also defines a component-based architectures model. This model defines a *Virtual Function Bus* (VFB) that defines an abstraction for communication between application *SoftWare Components* (SW-Cs). The VFB allows SW-Cs to be independent of the underlying hardware, making them portable between different ECUs and reusable across multiple automotive projects. The VFB abstraction is encapsulated by the AUTOSAR *Run-Time Environment* (RTE). The RTE provides the 'glue' between SW-Cs and the BSW.

For further information see <http://www.autosar.org>.

AUTOSAR OS

AUTOSAR OS is an extension to the OSEK OS specification. An AUTOSAR OS includes all the features of OSEK OS and adds some new functionality which is divided into four *Scalability Classes* as follows:

Scalability Class 1 includes OSEK OS plus:

Schedule Tables - Schedule tables provide an easier alternative to OSEK Alarms when programming repeating activities. Each schedule table can be managed as a single unit and you can switch between tables at runtime, allowing you to build 'modal' systems easily.

Software Counter Interface - The interaction between the OS and counters has been standardized (in OSEK this was vendor specific).

Stack Monitoring - Additional debugging support has been added to assist with stack-faults.

Scalability Class 2 includes Scalability Class 1 plus:

Schedule Table Synchronization - Schedule tables can be synchronized with a global time source (although this is trivially possible within Scalability Class 1).

Timing Protection - Protection is added to guard against tasks and interrupts executing for too long or too often. The protection scheme allows you to constrain at runtime those aspects of system timing that control whether your system meets its deadlines or not.

Scalability Class 3 includes Scalability Class 1 plus:

Memory Protection - Memory protection allows a system to be partitioned into **OS-Applications**. OS-Applications can be configured to be *trusted*, i.e. they run in what is typically called 'supervisor mode', or *non-trusted*, i.e. they run in what is typically called 'user mode'. Memory access constraints can be programmed for non-trusted OS-Applications and the OS manages the target microcontroller's memory management features at runtime to provide protection.

Service Protection - Access to the OS API can be allowed or denied for configured tasks/ISRs. For example you can forbid a task in one OS-Application from activating tasks in another OS-Application. API call protection also provides a mechanism for extending the API by adding *trusted functions* and granting or denying access to these functions as you would for the OS API.

Scalability Class 4 is a superset of Scalability Classes 2 and 3.

RTA-OS3.x 1.0.0 supports all AUTOSAR OS R3.x features from Scalability Classes 1-4.

As AUTOSAR OS is based on OSEK OS, it is backwards compatible to existing OSEK OS-based applications - i.e. applications written for OSEK OS will largely run on AUTOSAR OS without modification. However, the AUTOSAR OS standard also clarifies some of the ambiguities in the OSEK OS specification that arise when the behavior of OSEK OS is *undefined* or *vendor specific* because these represent a barrier to portability. Users who are migrating from an OSEK OS and rely upon a particular implementation of an OSEK OS feature should be aware that AUTOSAR OS defines the required OSEK OS behavior in the following cases:


OSEK OS	AUTOSAR OS
Behavior of relative alarms started at an offset of zero is undefined	Relative alarms cannot be started at a relative time of zero
The StartOS() API call may or may not return depending on the vendor implementation	StartOS() must not return
The behavior of ShutdownOS() is not defined if the ShutdownHook() returns.	ShutdownOS() disables all interrupts and enters an infinite loop.

The *Reference Guide* provides an API call compatibility listing between OSEK OS and AUTOSAR OS R3.x.

AUTOSAR OS replaces OSEK's OIL configuration format with an XML-based configuration language. AUTOSAR XML adopts the same configuration objects and concepts found in OIL, but uses a different syntax.

2.1.3 Unique RTA-OS3.x Features

RTA-OS3.x is much more than an AUTOSAR OS. The kernel is designed to support software engineers building and integrating real-time systems.

 *RTA-OS3.x-specific features are not guaranteed to be portable to other implementations of OSEK OS or AUTOSAR OS.*

The additional features include:

Time Monitoring to measure the execution time of tasks and Category 2 ISRs at runtime and optionally check times against pre-configured budgets.

Enhanced Stack Monitoring providing additional possibilities to help you debug stack problems

RTA-TRACE Integration providing automatic instrumentation of the OS kernel to support the ETAS RTA-TRACE real-time OS profiling and visualization tool so you can view exactly what the OS is doing in real-time.

User control of hardware so that there is no need to hand over control of hardware, such as peripheral timers, the cache and I/O ports etc. to the OS. All hardware interaction occurs through RTA-OS3.x's well-defined hardware interface.

Predictable run-time overheads such as switching to and from tasks, handling interrupts and waking up tasks, have low worst-case bounds and little variability within execution times.

Graphical offline configuration editor supporting AUTOSAR XML configuration of the OS.

Easy integration into your build process as RTA-OS3.x code generation requires just one command-line tool that can be driven from any build environment.

Highly scalable kernel architecture using offline tools that automatically optimize the kernel for your application.

2.2 Summary

- RTA-OS3.x is a preemptive RTOS for embedded systems
- The kernel provides the features specified in the AUTOSAR OS R3.x standard for all Scalability Classes, including support for the legacy OSEK OS.
- RTA-OS3.x provides additional features that make it easier to integrate AUTOSAR OS into your build process.

3 Development Process

This chapter provides a short overview of how to use RTA-OS3.x in your applications. The process involves the following steps:

1. Configure the features of the OS you want to use;
2. Generate a customized RTA-OS3.x kernel library;
3. Write application code that uses the OS;
4. Compile your application code and linking with the RTA-OS3.x library;
5. Run your application on your target.

The following sections cover each of these steps.

3.1 Configuration

RTA-OS3.x is statically configured, which means that every task and interrupt you need must be declared at configuration time, together with any critical sections, synchronization points, counters etc.

All configuration is held in XML files conforming to the AUTOSAR standard. The XML is not particularly easy to read, so the use of a tool is recommended - RTA-OS3.x includes **rtaoscfg**, a graphical configuration editor for configuring your RTA-OS3.x application. **rtaoscfg** accepts any AUTOSAR XML file as input and allows you to edit the OS-specific parts of a configuration. If the input file contains both OS and non-OS specific configuration then only the OS configuration will be modified.



The nature of XML parsers means that the ordering of configuration objects in the files may not be preserved when edited with different tools.

rtaoscfg has five main areas panes as shown in Figure 3.1.

1. Menu/Toolbar.
2. Project Navigator.
See an overview of the top-level objects in your configuration, switch between configuration workspaces and manage the files in the project.
3. Configuration Workspace.
This is where you do most of your configuration. RTA-OS3.x provides the following workspaces:
 - (a) OS Configuration;
 - (b) RTA-TRACE Configuration;
 - (c) Builder;

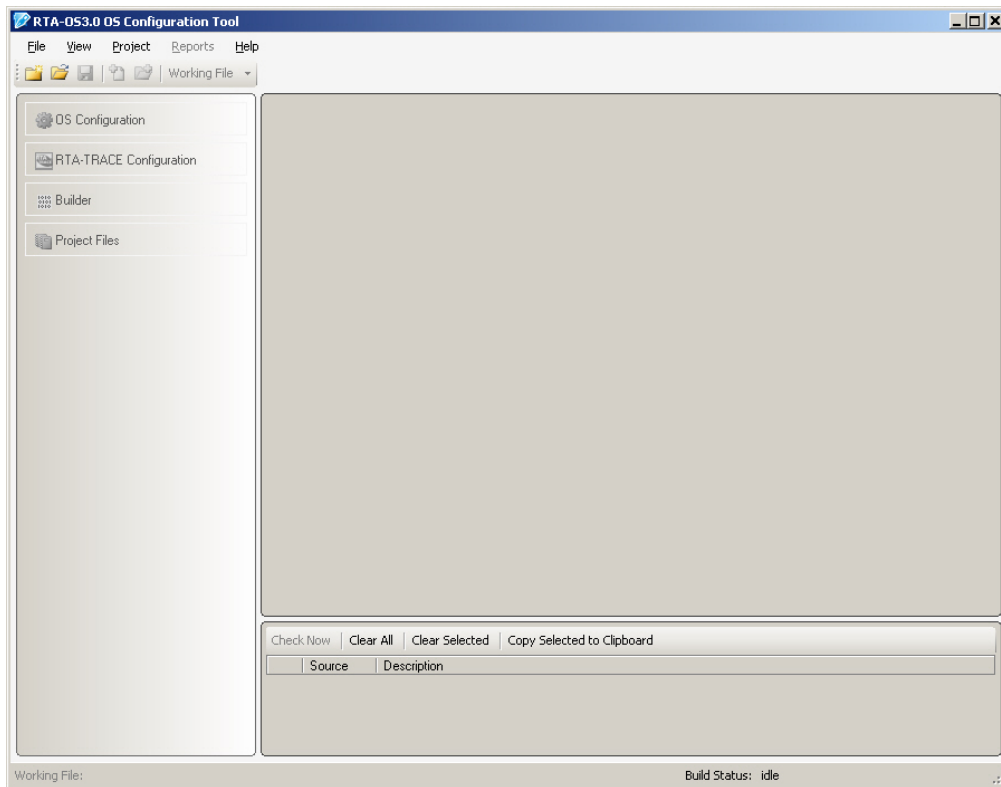


Figure 3.1: The **rtaoscfg** configuration tool

(d) Project Files.

4. Error Viewer.

This displays a list of errors in the currently loaded configuration.

5. Status Bar.

3.1.1 OS Configuration

The OS configuration navigator displays the logical structure of your OS configuration in the left hand window, grouped by OS object. You can expand each object to see the instances you have created. Clicking on an instance of an object displays the configuration panel in the right hand window. Individual items are configured in the workspace on the right hand side of the navigator.

To create a new configuration, select **File → New Project** from the menu or use the keyboard shortcut <Ctrl+N>. Each new configuration requires you to specify the *administrative* parts of an AUTOSAR XML configuration. This is required because parts of the OS configuration need to reference other parts (for example, tasks need to reference which resources they use) and these

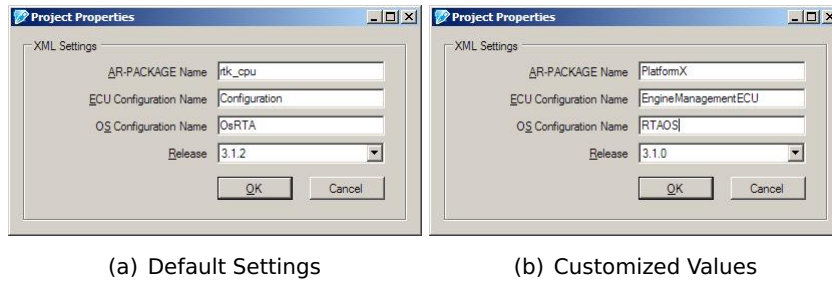


Figure 3.2: Defining project-wide settings

references are formed as an absolute path to an item in the AUTOSAR XML configuration. The items required are:

AR-PACKAGE Name defines the name of the AUTOSAR package. All AUTOSAR configuration items live in an AR-PACKAGE and a system may contain multiple packages. The OS configuration for a single ECU must live in a single package - it is not possible to split an OS configuration over multiple packages.

ECU Configuration Name defines the name of the ECU-CONFIGURATION of which this OS configuration will be a part. An ECU-CONFIGURATION contains all the configuration elements for all of the basic software for one ECU.

OS Configuration Name defines the name of the OS configuration MODULE-CONFIGURATION. This is the name that will be used to refer to the OS from the ECU-CONFIGURATION.

Schema defines which variant of AUTOSAR OS R3.x should be used. Since there are some minor differences between 3.0 and 3.1, this may need to be set to conform with other configuration files.

Figure 3.2 shows the default settings for a new project and example of how these might be customized for a particular project, in this case the ECU 'EngineManagementECU' for the vehicle platform 'PlatformX'.



*Apart from the Schema, the project settings can be modified later by selecting **Project → Properties**.*

Clicking on a 'root' object like "Tasks", "ISRs", "Resources" etc. in the project navigator takes you to the configuration workspace for that type of object. Clicking on individual objects takes you to the configuration workspace for the selected object. Figure 3.3 shows an OS configuration where TaskB has been selected.

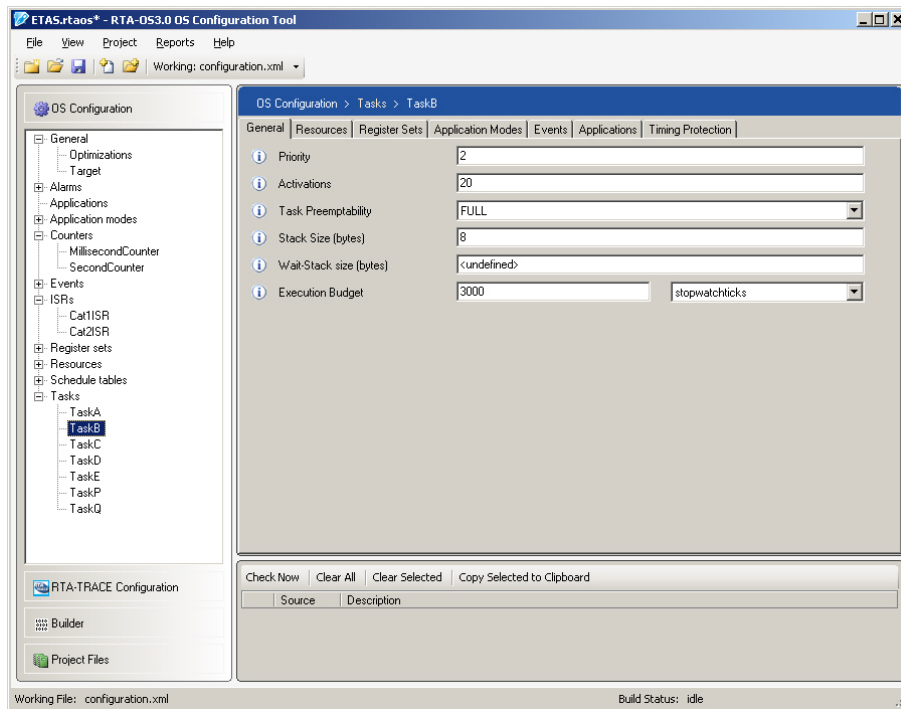


Figure 3.3: OS configuration in **rtaoscfg**

OS configuration is split into three main parts:

1. System-wide configuration for target-neutral general settings such as the level of debugging information you want to record, whether you monitor the stack at runtime, which hooks (callbacks) you are going to use etc.
2. Target-specific settings including:
 - the target device you are using. You can use **rtaoscfg** to configure any target device for which you have installed a licensed RTA-OS3.x port;
 - the variant of the target device if the port supports multiple variants of the target;
 - the version of the target to use if you have more than one version of a specific RTA-OS3.x port installed;
 - “Target Specific” aspects of configuration.
3. Object configuration for each of the OS objects you want to use. This includes tasks, interrupts, resources, events, alarms, counters and schedule tables.

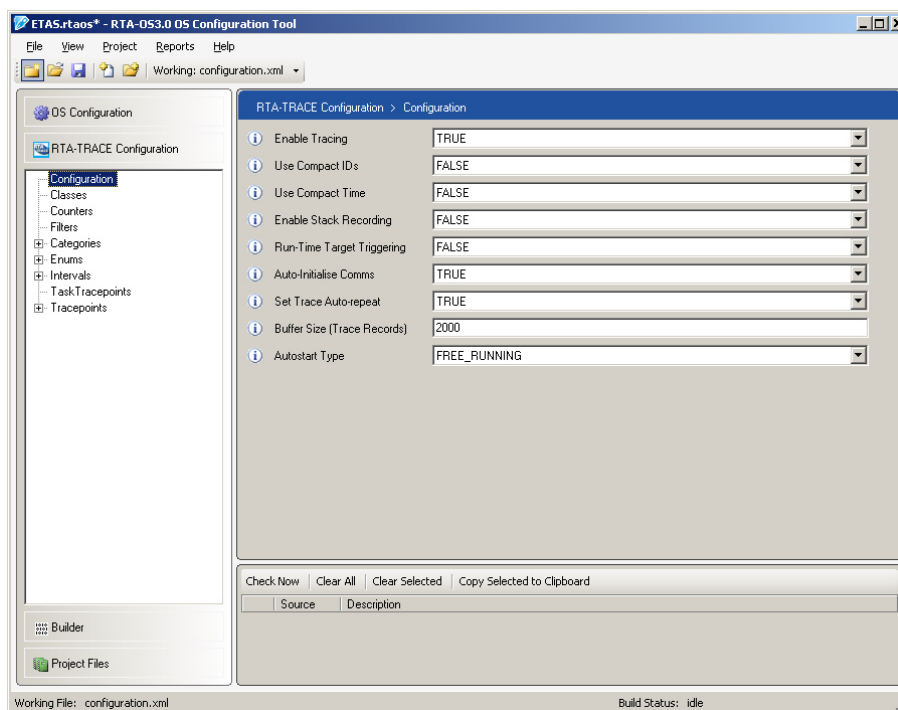


Figure 3.4: RTA-TRACE configuration in **rtaoscfg**

Specific aspects of configuration are covered in the later chapters of this user guide.

3.1.2 RTA-TRACE Configuration

The RTA-TRACE configuration view, shown in Figure 3.4, allows you to configure RTA-OS3.x to support ETAS' RTA-TRACE runtime profile and monitoring tool¹. Like the OS configuration view, an overview of the RTA-TRACE configuration is displayed in the project navigator and configuration windows are shown in the workspace. You can expand each object to see the instances you have created. Clicking on an instance of an object displays the configuration panel in the right hand window.

Configuring RTA-TRACE parameters tells RTA-OS3.x to include all the necessary OS instrumentation that allows RTA-TRACE to gather runtime data for your application. There is no harm configuring RTA-TRACE instrumentation if you do not have RTA-TRACE to view the trace data, but this will make your RTA-OS3.x configuration larger and slower than it would be without instrumentation.

¹RTA-TRACE is not supplied with RTA-OS3.x. For further information about RTA-TRACE contact your ETAS sales office.

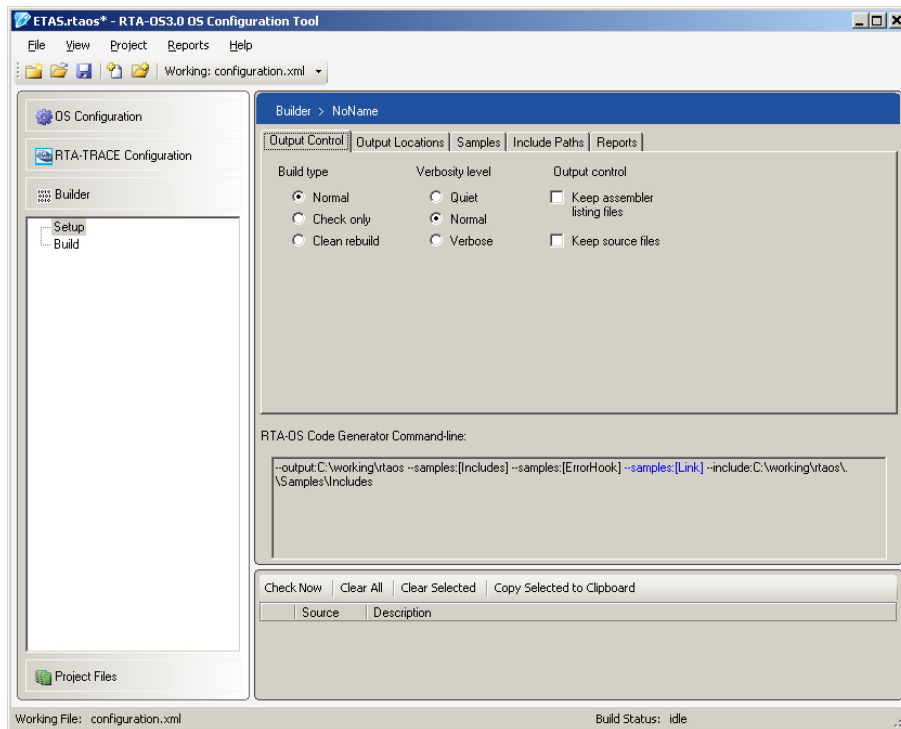


Figure 3.5: Running **rtaosgen** from **rtaoscfg**

3.1.3 Build

The RTA-OS3.x library is built using the command-line **rtaosgen** tool. If you prefer to configure and build within the same tool then you can do so from the build workspace shown Figure 3.5. The builder allows you to configure the build process and run **rtaosgen** from inside **rtaoscfg**. Further details on **rtaosgen** are in Section 3.3.

3.1.4 Project Files

By convention, **rtaoscfg** calls the working configuration a *project*. A project is simply one or more XML files that define your OS configuration. A default project contains a single AUTOSAR XML file.

The AUTOSAR XML language allows you to partition your OS configuration across multiple files according to the demands of your build and/or version control process. This is essential if you are working with other tools that generate fragments of OS configuration that need to be integrated into your main configuration. This is also useful if you want to maintain a ‘core’ configuration and then have multiple customizations of that core for different target hardware.

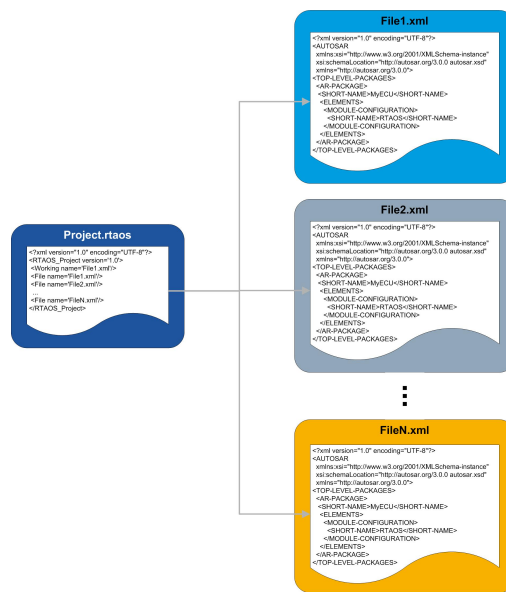


Figure 3.6: Relationship between Project File and AUTOSAR XML Files

The project viewer allows you to manage these complex projects. You can add/remove XML files to/from the project with ease. When you are working with a multi-file XML project, **rtaoscfg** internally merges all configuration data, allowing you to work with multiple configuration files simultaneously as if they are a single OS configuration. That way you can see the entire model and check for consistency even though configuration data is physically separated.

When the project is saved, **rtaosgen** remembers which parts of the configuration came from which XML file so that when you save your configuration each element of configuration data is written back to the correct file.

RTA-OS3.x also writes out a project file (with a .rtaos extension). A project file is a special RTA-OS3.x file that lists all the AUTOSAR XML files in your project. Figure 3.6 shows the basic concept.

Each XML file can be saved individually (i.e. without saving a project) from the context menu of the *Project Files* workspace.

ETAS Project files are specific to RTA-OS3.x

The Working File

When you create a new OS object (e.g. a task) or an attribute (e.g. a task's priority) in your configuration, **rtaoscfg** writes it to the *working file*. The project viewer shows you the current working file and allows you to switch the working file to be any one in your project.

If you create an object or attribute and realize that you added it to the wrong file, then you can simply move it to another file in your project by deleting it from the current working file, changing to a new working file and then re-creating the object or attribute.

If you work with many XML file fragments as part of your OS configuration then it would be tiresome to open each in turn when you want to make modifications. **rtaosgen** allows you to open the project file itself. This loads every XML file referenced by the project file automatically.

The current working XML file can be saved quickly (without saving the project) from the main menu (**File → Save Working File**).

3.1.5 Error Checking

When you add or remove configuration items **rtaoscfg** reports configuration errors in the Error Viewer.

Three types of errors are reported:

Information contain summary information about the configuration, such as how many objects of a type you have configured.

Warnings tell you that your configuration might not behave as expected because you might be missing something in your configuration.

Errors tell you that parts of your configuration are incorrect.

You can check your configuration for consistency by clicking on the **Check Now** button. Any errors that are found are reported in the error window. Figure 3.7 shows how errors are reported.

Unresolved References

An *unresolved reference* is an error reported by **rtaoscfg** when an object in the project references another object that is not declared in the same project. Examples of unresolved references are:

- a task or ISR referencing an undeclared resource;
- a task referencing an undeclared event;
- an alarm or schedule table referencing an undeclared task or event;
- an alarm or schedule table referencing an undeclared counter.

Check Now Clear All Clear Selected Copy Selected to Clipboard		
	Source	Description
	Check	Starting XML Checks
	Check	The parameter called [Priority] for [/AUTOSAR/0s/0sIsr] 'Cat2ISR' must have a value specified.
	Check	The parameter called [Address] for [/AUTOSAR/0s/0sIsr] 'Cat2ISR' must have a value specified.
	Check	The parameter called [Status] for [/AUTOSAR/0s/0sOS] '0sInfo' must have a value specified.
	Check	The parameter called [UseParameterAccess] for [/AUTOSAR/0s/0sOS] '0sInfo' must have a value specified.
	Check	The parameter called [UseResScheduler] for [/AUTOSAR/0s/0sOS] '0sInfo' must have a value specified.
	Check	The parameter called [Schedule] for [/AUTOSAR/0s/0sTask] 'TaskA' must have a value specified.
	Check	Skipping Pre-Build checks due to previous errors.

Figure 3.7: Viewing configuration errors

Check Now Clear All Clear Selected Copy Selected to Clipboard		
	Source	Description
	Check	Starting XML Checks
	Check	There is an unresolved reference to "/rtk_cpu/0sRTA/WakeUp".
	Check	There is an unresolved reference to "/rtk_cpu/0sRTA/StandardResource".
	Check	There is an unresolved reference to "/rtk_cpu/0sRTA/WakeUp".
	Check	There is an unresolved reference to "/rtk_cpu/0sRTA/WakeUp".
	Check	There is an unresolved reference to "/rtk_cpu/0sRTA/WakeUp".
	Check	Skipping Pre-Build checks due to previous errors.

Figure 3.8: Unresolved References

Typically references may be unresolved because you are editing a single XML file that is part of a larger configuration or because you are importing XML from a 3rd party source.

Unresolved references are reported as errors in the error window as normal, as shown in Figure 3.8.

However, when you try and edit the object that is makes the unresolved reference, **rtaosgen** will create the missing object automatically in an XML file called `MissingReferences.xml` and add that file to the project. Figure 3.9 shows the creation of the missing `StandardResource` referenced from `TaskA`.

You can see which missing references have been created by viewing the contents of `MissingReferences.xml` in the "Project Files" workspace. Figure 3.10 shows that `StandardResource` has been created.

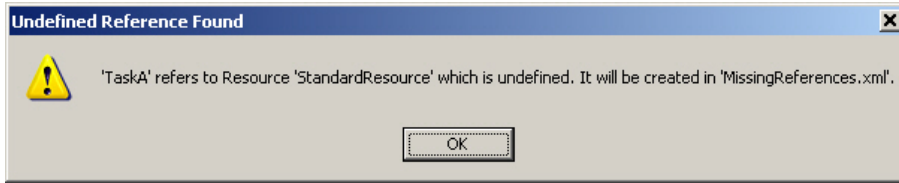


Figure 3.9: Creating an entry in MissingReferences.xml

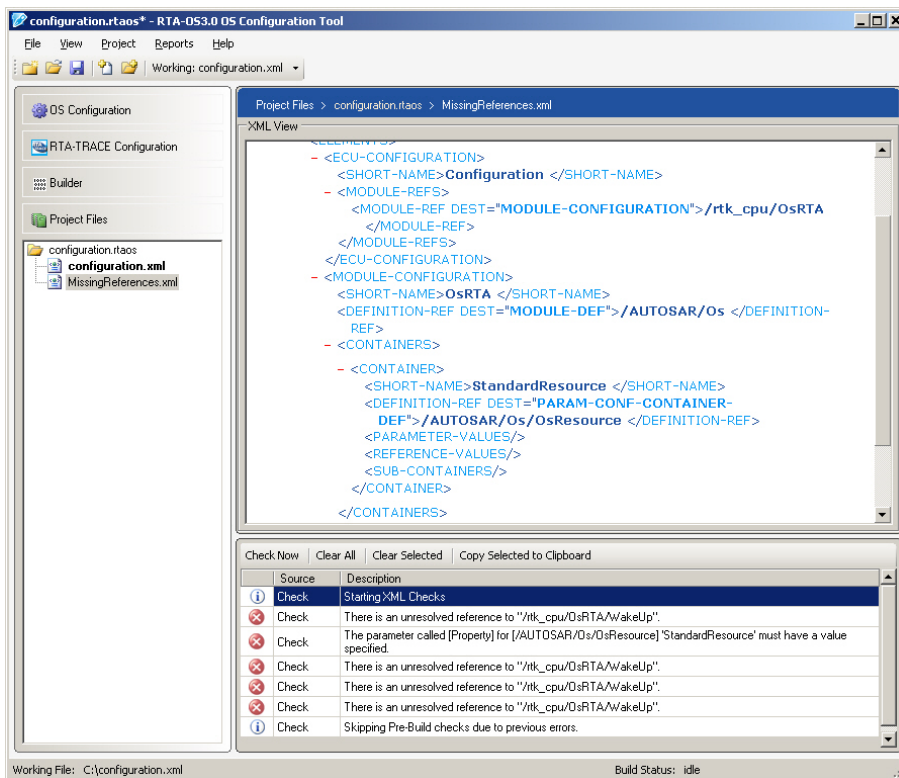


Figure 3.10: Viewing the content MissingReferences.xml



If the missing references are declared elsewhere and the referenced objects will be present at the point you build the kernel, you should remove `MissingReferences.xml` from the project immediately before saving the project.

3.1.6 Generating Reports

You can generate reports about your configuration from the menu/toolbar. Reports present summary information about your configuration that can be used for Quality Assurance audits, internal communication between departments, etc. The reports provided include:

Configuration Summary - an overview of the OS configuration.

OS API Reference - a customized reference guide for the configured OS. This may include documentation for target-specific features such as additional API calls and types that are not part of the standard *Reference Guide*.

Stack Usage - worst-case stack usage for the configured OS. This requires that stack allocations have been provided for each tasks and ISR.

MISRA Deviations - provides the MISRA-C 2004² deviations for the configured OS. This includes which deviations apply, why the deviation has been made and how many times it occurs.

All reports are generated in HTML and will be displayed in your default web-browser.

Generation of reports is actually done by **rtaosgen** so can be produced at the same time as the build process runs. As an example, you may wish to generate the MISRA Deviation report to provide evidence for your QA process. Build-time generation of reports also allows the format for the report to be selected. Reports are provided as plain text, XML and HTML.

Figure 3.11 shows an extract from the “Configuration Summary” report that lists every OS object and its associated configuration.

A full list of the reports available for your target can be obtained using the following command line:

```
rtaosgen --target:YourTarget --report:?
```

²Motor Industry Standards and Reliability Association

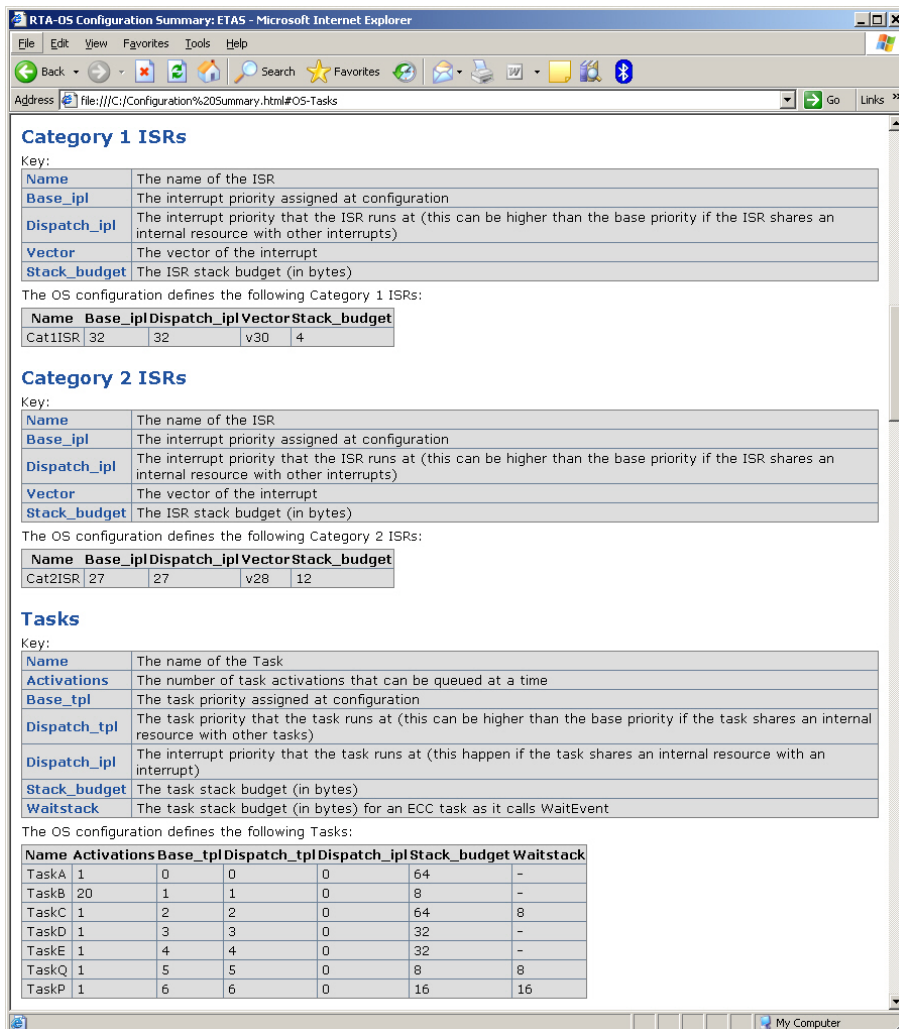


Figure 3.11: Automatically generated configuration summary report

3.2 Assistants

rtaoscfg supports *assistant* plug-ins to perform frequently-used activities. If any assistant plugins are detected at the start of **rtaoscfg**, an additional **Assistants** menu option will be visible in the main menu bar. Starting with V2.0.0, **rtaoscfg** ships with a single *Schedule Table Assistant* allowing periodic Schedule Tables to be easily configured. Further details of the Schedule Table Assistant can be found in Section 10.3.

3.3 Library Generation

Before you can use RTA-OS3.x in your application, you need to generate an RTA-OS3.x kernel library and associated header files. **rtaosgen** generates a customized RTA-OS3.x kernel library that is optimized for your OS configuration by:

- Analyzing your XML configuration and automatically optimizing the RTA-OS3.x kernel so that it contains only those features that you will use. This makes RTA-OS3.x as small and efficient as possible.
- Customizing the optimized kernel to your chosen target using information provided by the target plug-in.
- Building a kernel library using the same 3rd party tool chain that you are using for your application. This guarantees tool chain compatibility between RTA-OS3.x and your own code.

The tool integrates the core OS kernel, enables optional kernel features you have selected and customizes this with the information about the target supplied by the port plug-in. Figure 3.12 illustrates the basic process by showing how a kernel would be generated for the EDSAC target.

3.3.1 Preparing the Tool Chain

To generate a kernel library, **rtaosgen** needs access to the compiler tool chain for your target. **rtaosgen** knows how to run the compiler, assembler and librarian for your target and what options to use. You need only be concerned with two things:

1. Your compiler tool chain must be accessible on your PATH.
2. Your compiler tool chain must be compatible with RTA-OS3.x.

You can find out if your compiler is on your PATH by opening a Windows Command Prompt using **Start → Run** and running **cmd**.

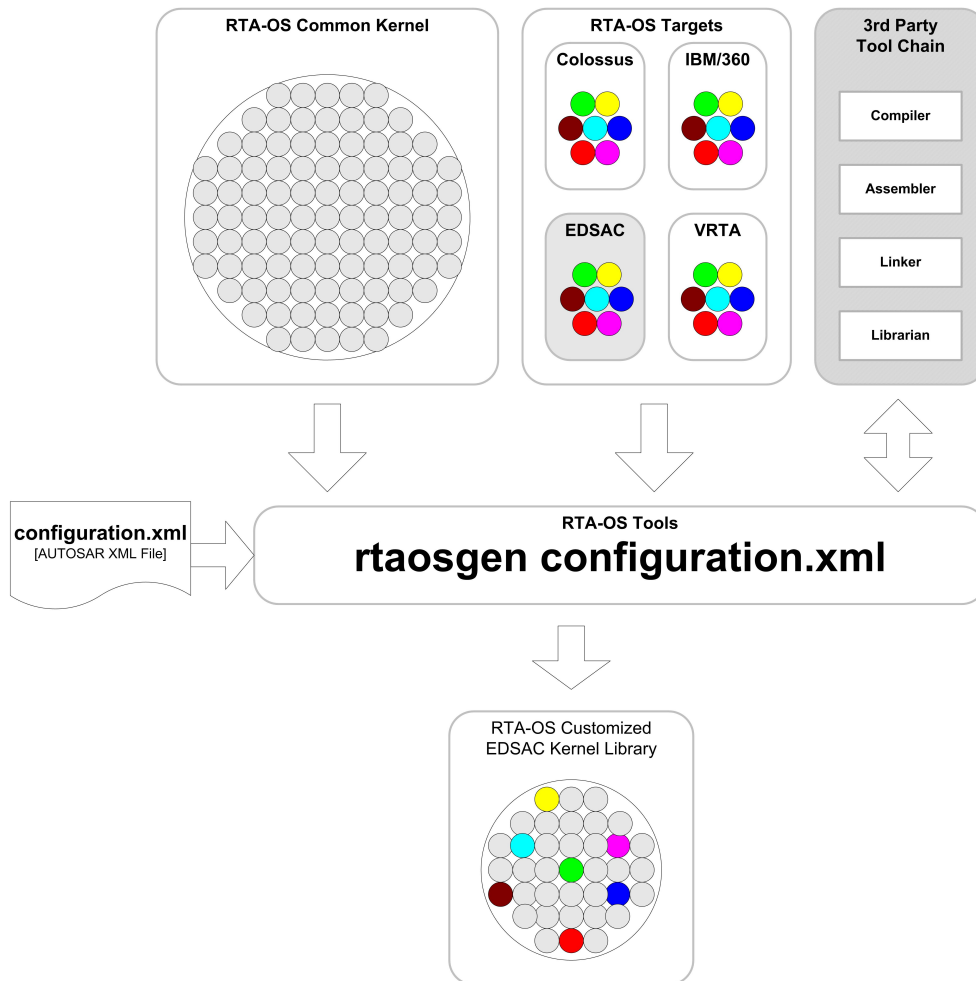


Figure 3.12: Building an RTA-OS3.x Kernel Library

Typing `C:\>set` at the command prompt will list every Windows environment variable. You should see your compiler's executable directory on the path. If you do not, then you can add your compiler to the path by typing:

```
C:\> set PATH=PATH;<Path to your compiler executable>
```

To check whether you are using a compatible version of the compiler tool chain you should consult the *Target/Compiler Port Guide* for your port which will tell you which version (or versions) are compatible.

3.3.2 Understanding AUTOSAR Dependencies

RTA-OS3.x is an AUTOSAR basic software module³ and as such it must conform to the AUTOSAR basic software module build concept. If you know how this works already then skip ahead to Section 3.3.4.

In AUTOSAR, all basic software modules provide a single include file called `<BSW Short name>.h`. For the OS this is `Os.h`. Each of these header files has dependencies on other AUTOSAR include files. The AUTOSAR include hierarchy is shown in Figure 3.13

Std_Types.h provides all the portable (i.e. target hardware invariant) type definitions for AUTOSAR. `Std_Types.h` includes a further two AUTOSAR header files:

Platform_Types.h defines the AUTOSAR standard types (`uint8`, `uint16`, `boolean`, `float32` etc.) for the target hardware.

Compiler.h defines a set of macros that are used internally by basic software modules to mark functions, data and pointers according to the mode by which they can be addressed.

The memory and pointer classes used by `Compiler.h` are defined by each basic software module in a file called `Compiler_Cfg.h`. A minimum set of names are defined by the AUTOSAR standard and each name is prefix with the name of the basic software module. For the OS, all section name macros start with `OS_`.

The system integrator must merge the `Compiler_Cfg.h` files for all basic software modules to create a 'master' `Compiler_Cfg.h` *before the system is compiled*. In RTA-OS3.x, the OS module's `Compiler_Cfg.h` is called `Os_Compiler_Cfg.h` and it contains the complete list of the names used by RTA-OS3.x. The file can be easily **#included** into the system-wide `Compiler_Cfg.h`.

You should take particular note of the RTA-OS3.x section called `OS_APPL_CODE`. This should be used to place all application code

³One of many - there are other modules for communication stacks (CAN, LIN, FlexRay etc.), non-volatile memory handling, peripherals drivers, etc. that are available from third-parties.

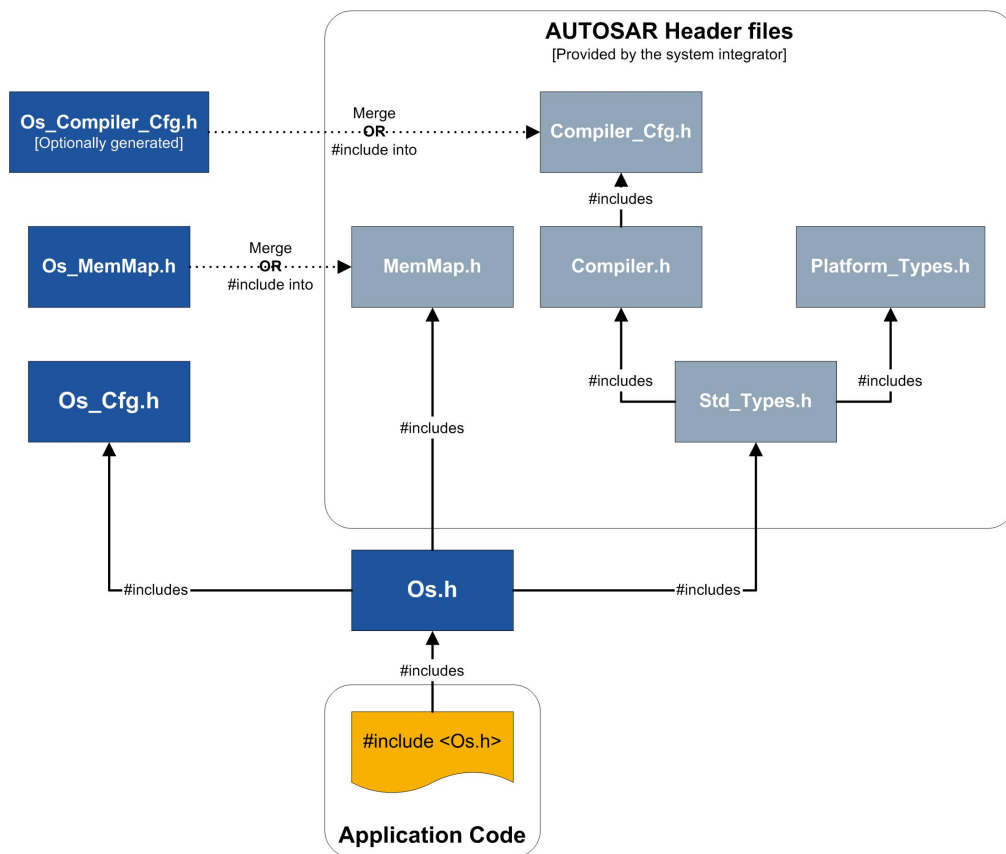


Figure 3.13: AUTOSAR Header File Hierarchy

that is required by the kernel, i.e. all your hooks and callback routines. Code can be placed in this section using the directive `FUNC(<typename>, OS_APPL_CODE)`. For example the following code shows how to place the `ErrorHook()` into `OS_APPL_CODE`:

```
FUNC(void, OS_APPL_CODE) ErrorHook(StatusType Error){
    /* Handle error */
}
```

The RTA-OS3.x documents use this form when defining callbacks and hook routines.

MemMap.h defines how data and code is mapped to memory sections and uses the compiler's primitives for placing code and data into different types of memory section according to the following process:

1. each basic software module defines a series of section names using macros in `Compiler_Cfg.h`
2. the vendor of the basic software module uses these macros to place code in the virtual sections during implementation, for example:

```
#define OS_START_SEC_CODE
#include "MemMap.h"
/* Some OS code here */
#define OS_STOP_SEC_CODE
#include "MemMap.h"
```

3. the system integrator develops a `MemMap.h` file that maps the basic software's virtual section names onto system-wide section names and from there onto primitives of the compiler for section placement, for example:

```
/* Map OS code into the section containing all BSW
code */
#ifdef OS_START_SEC_CODE
    #undef OS_START_SEC_CODE
    #define START_SECTION_BSW_CODE
#endif
...
/* Name the system section with a compiler primitive
*/
#ifdef START_SECTION_BSW_CODE
    #pragma section code "bsw_code_section"
#endif
```

As with `Compiler_Cfg.h`, each basic software module must also provide a module-specific version of `MemMap.h`. In RTA-OS3.x, the module-specific version of `MemMap.h` is called `Os_MemMap.h`. The `Os_MemMap.h`

file can be either merged or **#included** into a 'master' MemMap.h *before the system is compiled*

To build an RTA-OS3.x library it follows that all the standard AUTOSAR header files are required as inputs to the build process and these are outside the scope of the OS. However, RTA-OS3.x can generate *sample* versions of the AUTOSAR standard header files if required.



*You must enhance or replace the sample AUTOSAR standard header files generated by **rtaosgen** for use in production software.*

3.3.3 Running **rtaosgen**

rtaosgen is a command line tool. You can invoke it from the Windows command prompt, from a make script, Ant script, in fact from anywhere where you can call a Windows application. The **rtaosgen** tool can be run from the **rtaoscfg** Builder if you prefer to use a graphical environment.

rtaosgen takes one or more configuration files as input. Configuration files can be:

- AUTOSAR XML
- RTA-OS3.x project files
- A mixture of both

Like **rtaoscfg**, **rtaosgen** will merge the contents of all files passed on the command line into an in-memory OS configuration before generating the kernel.

3.3.4 Building the library

To build an RTA-OS3.x library it follows that all the standard AUTOSAR header files are required as inputs to the build process. You must include the path to the location of the AUTOSAR standard headers files when invoking **rtaosgen**. For example, to build a library for the *Hello World* example application for an RTA-OS3.x target you can type:

```
C:\>rtaosgen --include:PathToAutosarHeaderFiles HelloWorld.rtaos
```

If you do not have access to AUTOSAR include files (for example, if you are using RTA-OS3.x outside of a full AUTOSAR system), then **rtaosgen** can generate them automatically for you.

```
C:\>rtaosgen --samples:[Includes] --include:Samples\Includes  
HelloWorld.rtaos
```

Note that **rtaosgen** does not force you to use a specific extension - you can use any extension you like.

rtaosgen generates four classes of messages during execution:

Information. These messages tell you useful things about the configuration, for example how many tasks you have configured. **rtaosgen** will generate output files.

Warning. These messages warn you that your configuration will result in an OS that might not behave as you expect. **Rtaosgen** will generate output files.

Error. These messages tell you that there is something wrong with your configuration. **rtaosgen** will stop processing your configuration at a convenient point and no output files will be generated.

Fatal. You will get at most one fatal message. It tells you that there is something fundamentally wrong with either your configuration or **rtaosgen**. **rtaosgen** stops immediately.

You can do other things from the command line like change the output directory for generated files, suppress messages, etc. For more details, see the *Reference Guide*.

Building from **rtaoscfg**

It is also possible to build RTA-OS3.x from within **rtaoscfg** in the “Builder” workspace. This allows you to run **rtaosgen** with command line you specify **rtaoscfg**. If you specify a command line, then it will be saved in the RTA-OS3.x project file.

The same tool, **rtaosgen** is used to build the kernel irrespective of whether it is called directly from the command line or internally from **rtaoscfg**.

Note that if you build from within **rtaoscfg** then you still need to ensure that your compiler tool chain is on your Windows path.

Any settings that you configure in the “Builder” are stored in your RTA-OS3.x project file.

3.3.5 Generated Files

When **rtaosgen** runs and terminates without reporting any errors or fatal messages then it will have generated the following files:

Filename	Contents
Os.h	The main include file for the OS.
Os_Cfg.h	Declarations of the objects you have configured. This is included by Os.h.
Os_MemMap.h	AUTOSAR memory mapping configuration used by RTA-OS to merge with the system-wide MemMap.h file.
RTA0S.<lib>	The RTA-OS library for your application. The extension <lib> depends on your target.
RTA0S.<lib>.sig	A signature file for the library for your application. This is used by rtaosgen to work out which parts of the kernel library need to be rebuilt if the configuration has changed. The extension <lib> depends on your target.

Generating sample code

The **rtaosgen** code generator can generate sample code that can be used as a basis for your application. You have already seen one case of this in Section 3.3.4 when `--samples[Includes]` was used to generate sample AUTOSAR standard header files.

The set of samples provided is port-dependent, but you can get a list of provided samples using the following command line:

```
C:\>rtaosgen --target:YourTarget --samples:?
```

Most ports will provide samples that show how to write AUTOSAR OS hook functions like the `ErrorHook()`, `StartupHook()` etc. For example, to generate a default `ErrorHook()` you could use the following command line:

```
C:\>rtaosgen --samples:[ErrorHook]
--include:PathToAutosarHeaderFiles HelloWorld.rtaos
```

3.4 Integration

3.4.1 Accessing the OS in your Source Code

To access RTA-OS3.x in your source code you simply include **#include** `<Os.h>` in every C compilation unit (i.e. every C source code file) where you need to access RTA-OS3.x. The header file is protected against multiple-inclusion. RTA-OS3.x does not place any restrictions on how you organize your source code - you can put all of your code into a single source file or place each task and interrupt implementation into its own source file as you (or your configuration control process) requires.

3.4.2 Implementing Tasks and ISRs

Tasks

For each task that you declare at configuration time you must provide an implementation of the task. Each task needs to be marked using the TASK(x) macro. Tasks typically have the following structure:

```
#include <0s.h>
TASK(MyTask){
    /* Do something */
    TerminateTask();
}
```

Category 2 ISRs

Each Category 2 ISR that you declare needs to be implemented. This is also marked, this time by ISR(x):

```
#include <0s.h>
ISR(MyISR){
    /* Do something */
}
```



A Category 2 ISR handler does not need a return from interrupt call - RTA-OS3.x does this automatically. Depending on the behavior of interrupt sources on your target hardware, you may need to clear the interrupt pending flag. Please consult the hardware documentation provided by your silicon vendor for further details.

Category 1 ISRs

Each Category 1 ISR that you declared also needs to be implemented. Your compiler will use a special convention for marking a C function as an interrupt. RTA-OS3.x provides a macro that expands to the correct directive for your compiler. Your Category 1 handler will therefore look something this:

```
#include <0s.h>
CAT1_ISR(MyCat1ISR) {
    /* Do something */
}
```

3.4.3 Starting the OS

RTA-OS3.x does not take control of your hardware so you need to start the OS manually using the StartOS() API call, usually in your main() program. RTA-OS3.x provides a macro called OS_MAIN() which expands to the correct type of main() definition for your compiler toolchain⁴.

⁴On many compilers this will be **void** main(**void**), but there are compilers that insist upon the main() program returning an integer or other (non **void**) type.

```

#include <Os.h>
OS_MAIN() {
    /* Initialize target hardware */
    /* Do any mode management, pre-OS functions etc. */
    StartOS();
    /* Call does not return so you never reach here */
}

```

3.4.4 Interacting with the RTA-OS3.x

You interact with RTA-OS3.x by making kernel API calls. You can find a complete list of calls in the *Reference Guide*.

3.4.5 Compiling and Linking

When you compile your code you must make sure that `Os.h` and `Os_Cfg.h` are reachable on your compiler include path. When you link your application you must link against `RTAOS.<lib>`, and the library must be on your linker's library path.

3.5 Memory Images and Linker Files

When you build your application, the various pieces of code, data, ROM and RAM that were placed into the sections defined in `MemMap.h` need to be located at the right place in memory. This is typically done by your linker⁵ which resolves references made by user-supplied code to the RTA-OS3.x library, binds together the relevant object modules and allocates the resultant code and data to addresses in memory before producing an image that can be loaded onto the target.

The linker needs to know what parts of the program to place in which types of memory, where the ROM and RAM are on the microcontroller, and how map the parts of the program to the correct sort of memory.

3.5.1 Sections

Code and data output by compilers and assemblers is typically organized into *sections*. Some sections will contain just code, some code and data and some will contain data only. You might see a piece of assembler that says something like that shown in Code Example 3.1.

⁵An historical note: Technically this job is that of the locator which locates sections into memory by mapping virtual to physical addresses and these tools used to be called linker/locators. In modern times the locator part has dropped out of common usage and the tools are commonly referred to as 'linkers'.

```

.section CODE
.public MYPROC
mov    r1, FRED
add    r1, r1
ret
.end CODE
.section DATA
.public FRED
.word 100, 200, 300, 400
.end DATA
.section BSS
.public WORKSPACE
.space 200
.end BSS

```

Code Example 3.1: Example Assembler Output Showing Sections

This means that the code for MYPROC should be assembled and the object code should assume that it will be located in a section of memory called CODE whose location we will specify later in the linker control file. Similarly, the data labeled FRED will be placed in a section called DATA, and a space of 200 bytes labeled WORKSPACE will be allocated in section BSS.

C compilers typically output your code into a section called code or text, constants that must go into ROM in a section called something like **const**, and variables into data. There will usually be more - consult the reference manual for your toolchain for more details on what the sections are called and familiarize yourself with where they need to go.

Under AUTOSAR, your MemMap.h will define the actual names of the sections that need to be located. for example So far we have yet to map these onto addresses in 'real' memory. We must therefore look at how these sections are mapped into a memory image.

'Near' and 'Far' Address Spaces

On some processors there exist regions of memory space that can be addressed economically (typically with shorter, smaller instructions that have simpler effective-address calculations), are located on-chip rather than off-chip, or that are fabricated in a technology such that they are more cycle-efficient to access. RTA-OS3.x terms this memory *near* space and on these processors places some key data in these areas. On such platforms you will be supplied with information on where you must locate *near* space in ROM and/or RAM, and told in the *Target/Compiler Port Guide* what data is placed in it. *Far* space refers to the whole of memory.

Program and Data Space on Harvard Architectures

Most of the discussion about memory so far has assumed the conventional *von Neumann architecture*, in which data and code occupy one address space with ROM and RAM located at different offsets inside this. Some processors (typically very small microcontrollers like PICs, or high-performance Digital Signal Processors) adopt a *Harvard architecture*, in which there are distinct address spaces for code and data (there are some performance advantages to this that offset the programming disadvantages). On a Harvard-architecture processor, RTA-OS3.x may use data space (typically RAM) to store data that would normally be ROM constants on a von Neumann architecture processor, and the startup code will typically contain code to fetch a copy of the constant data into data space. If you are using a Harvard architecture processor, the *Target/Compiler Port Guide* will contain information on any use of RAM used to store copies of constants.

3.5.2 The Linker Control File

The linker control file governs the placement of code, data and reserved space in the image that is downloaded to the target microcontroller. Linker files vary considerably between platforms and targets, but typically include at least the following:

- declarations of where ROM and RAM are located on chip - these may vary across different variants in a CPU family.
- Lists of sections that can be placed into each memory space
- Initialization of the stack pointer, reset address, interrupt vectors etc.

Code Example 3.2 shows a hypothetical linker control file:

```
ONCHIPRAM start 0x0000 {
    Section .stack size 0x200 align 16 # system stack
    Section .sdata align 16           # small data
    Section bsw_near align 16         # near data
}

def __SP = start stack               # initialize stack ptr

RAM start 0x4000 {
    Section .data align 16            # compiler data
    Section .bss align 16            # compiler BSS
    Section bsw_zero_init align 16   # Basic Software zeroed RAM
    Section bsw_startup_init align 16 # Basic Software initialized
    RAM
    Section swc_startup_init align 16 # Application initialized RAM
}
```

```

ROM start 0x8000 {
    Section .text                # compiler code
    Section .const              # compiler constants
    Section swc_data align 16   # Application static data
    Section swc_init align 16   # Application initial data
    Section bsw_init align 16   # Basic Software initial data
}

VECTBL start 0xFF00 {
    Section 0sVectorTable      # RTA-OS's vector table
}

def __RESET = __main          # reset to __main

```

Code Example 3.2: A Linker Control File

The file above defines four separate parts of memory - ONCHIPRAM, RAM, ROM, and VECTBL. Into each section are placed the appropriate data, as described by the comments.

The example applications supplied with RTA-OS3.x embedded ports will contain a fully-commented linker control file; consult this and the *Target/Compiler Port Guide* for details of how to locate the sections correctly for your target platform.

3.6 Summary

- There are 5 steps to integrate RTA-OS3.x with your application:
 1. Configure the features of the OS you want to use
 2. Generate a customized RTA-OS3.x kernel library
 3. Write application code that uses the OS
 4. Compile your application code and linking with the RTA-OS3.x library
 5. Run your application on your target
- There are two offline tools: **rtaoscfg** to configure RTA-OS3.x and **rtaosgen** to generate and build the kernel library
- RTA-OS3.x is an AUTOSAR Basic Software module and has dependencies to AUTOSAR header files. These can be generated by **rtaosgen** if required.
- Linking and locating of RTA-OS3.x depends on the content of the MemMap.h file with which **rtaosgen** builds the kernel library.

4 Tasks

A system that has to perform a number of different activities at the same time is known as concurrent. These activities may have some software part, so the programs that provide them must execute concurrently. The programs will have to cooperate whenever necessary, for example, when they need to share data.

Each concurrent activity in a real-time system is represented by a task. The majority of the application code exists within tasks. If you have a number of tasks that must be executed at the same time, you will need to provide a means to allow concurrency. One way for you to do this is to have a separate processor for each task. You could use a parallel computer, but this solution is too expensive for many applications.

A much more cost effective way for you to achieve concurrent behavior is to run one task at a time on a single processor. You can then switch between tasks, so that they appear to be executing at the same time.

4.1 Scheduling

RTA-OS3.x provides a *scheduler* that switches between tasks according to a fixed priority which is assigned at configuration time. A priority is just a reflection of the relative urgency of tasks. There are many schemes that you can use to assign priorities to tasks, but common ones you may have heard of are:

Deadline Monotonic Assignment higher priorities are allocated to tasks with shorter deadlines.

Rate Monotonic Assignment higher priorities are allocated to tasks that need to run the most frequently.

However you choose to assign priorities, the sequence in which your tasks execute is determined by a *scheduling policy*. The scheduling policy determines when tasks actually run.

AUTOSAR OS supports two scheduling policies:

1. Preemptive Scheduling.

The fixed-priority preemptive scheduling algorithm is simple: run the highest priority task that is ready to run. If a task is running and a higher priority task becomes ready to run, then the higher priority task preempts the running task. This is called a task switch. When the higher priority task has finished then the preempted task resumes.

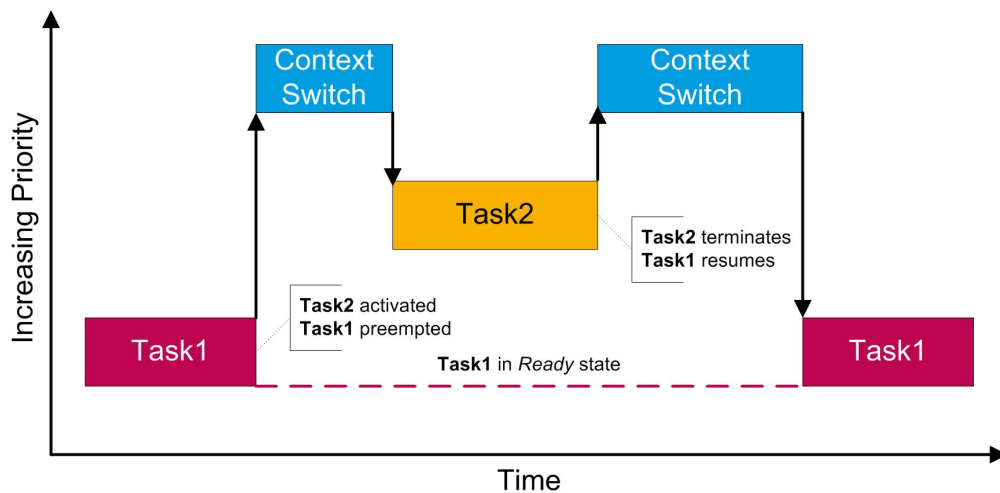


Figure 4.1: Preemptive scheduling of tasks

For a system where all tasks need to meet their deadlines at runtime, preemptive scheduling is the most efficient scheduling policy and will guarantee the shortest time between a task being activated (made ready to run) and terminating. This time is called the response time for the task. Preemptively scheduled systems need to consider the effect of preemption on shared data and may need to introduce mechanisms for concurrency control (see Chapter 6).

2. Non-Preemptive scheduling.

The OS runs the highest priority task that is ready to run, as with preemptive scheduling. However, unlike preemptive scheduling, if a higher priority task becomes ready, then it remains ready to run until the running task terminates - it does not preempt. What this means is that a non-preemptive task that starts running will always run to completion and then terminate.

Non-preemptive scheduling results in a less responsive system than preemptive scheduling (i.e. tasks will usually have longer response times), but the system does not need to worry about concurrency problems that arise for accessing shared data because the scheduling model doesn't allow concurrent access to shared data.

Actually, AUTOSAR OS provides support for a third type of scheduling called cooperative scheduling because it allows a non-preemptive task to tell the OS when it could be preempted. The reason we said AUTOSAR OS supports 2 policies is that there is only configuration for two - the third you have to build yourself.

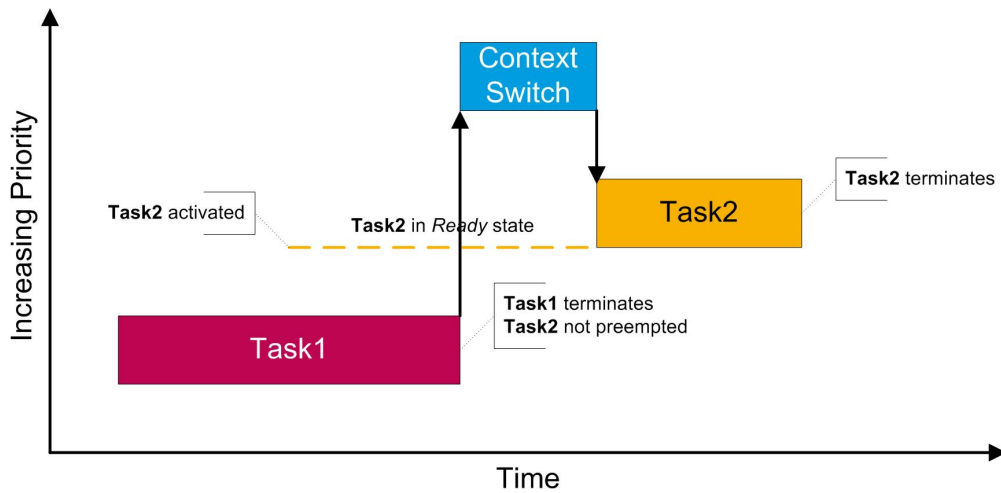


Figure 4.2: Non-preemptive scheduling of tasks

3. Cooperative scheduling.

The OS runs the highest priority task that is ready to run. If a higher priority task becomes ready, then it remains ready to run until either: the running task terminates (just like non-preemptive scheduling); or the running task makes a `Schedule()` API call to tell the OS it can be preempted. When the `Schedule()` call is made then the higher priority task preempts the running task and a task switch is said to have occurred (just like preemptive scheduling). When the higher priority task has finished then the preempted task resumes.

With careful design, the co-operative model provide can provide systems that, while not as responsive as fully preemptive systems, do not suffer the lack of responsiveness found with non-preemptive scheduling.

With all these types of scheduling it is important to realize that any task, whether preemptive or not, can be interrupted (preempted) by an interrupt service routine. Chapter 5 provides more information about how RTA-OS3.x deals with interrupts.

4.2 Basic and Extended Tasks

RTA-OS3.x OS supports two types of task:

1. Basic tasks.

Basic tasks start, execute and terminate (this is often called a single-shot tasking model). A basic task only releases the processor if it terminates, or if it is preempted by a higher priority task. This behavior

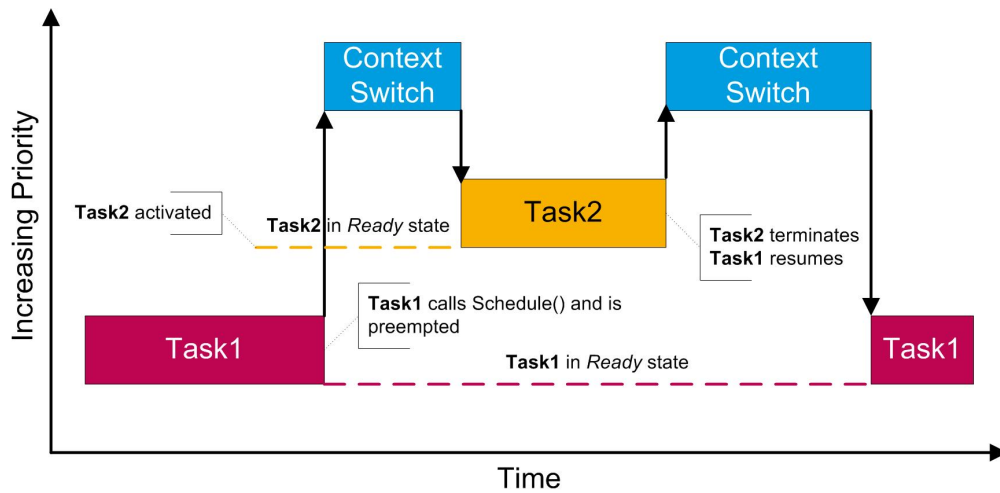


Figure 4.3: Co-operative scheduling of tasks

makes them highly suitable for embedded control functionality. Basic tasks are fast and efficient.

2. Extended tasks.

Extended tasks start, execute, wait for events and (optionally) terminate. The ability for an extended task to voluntarily suspend itself during execution provides a way for the task to have synchronization points. This feature makes extended tasks more suitable for functionality requiring mid-execution synchronization (for example, waiting for user interaction) than basic tasks.

4.2.1 Task States

Basic tasks operate on a 3-state model. A basic task can exist in the following states:

1. Suspended.
2. Ready.
3. Running.

Extended tasks can have an extra state which they enter when waiting for events:

4. Waiting.

Figure 4.4 shows the 3 and 4 state task models.

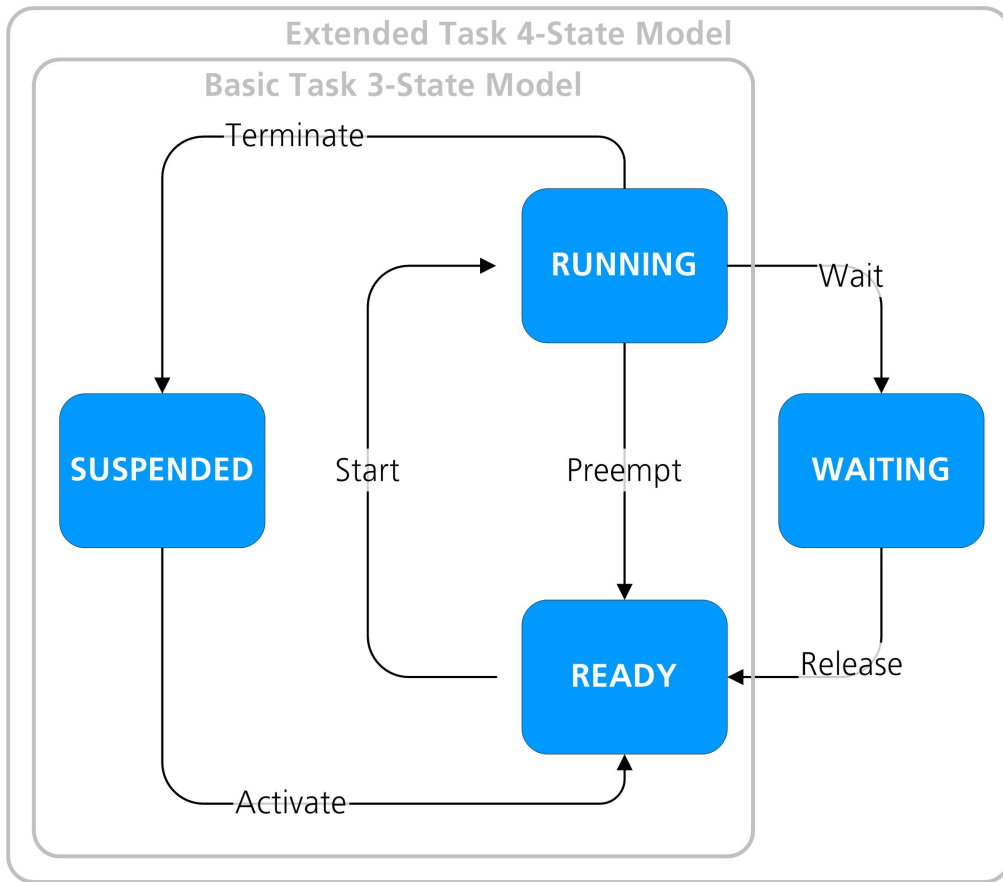


Figure 4.4: Task State Model

The default state for all tasks is suspended. A task is moved into the ready state by the process of activation. It is important to understand that activation does not cause the task to run - it makes it ready to run. Activation can happen a number of ways, for example by calling the `ActivateTask()` API in your code or as a result of some trigger, such as the expiry of an alarm (see Chapter 9) or a schedule table expiry point (see Chapter 10).

When a task becomes the highest priority task in the system, RTA-OS3.x moves the task into the running state and starts task execution at the first statement in the task. This is often referred to as dispatching the task. A task may be preempted during execution by other higher priority tasks that become ready.

If a higher priority task becomes ready to run, the currently executing task is preempted and is moved from the running state into the ready state. This means that only one task can be in the running state at any one time.

A task returns to the suspended state by terminating. A task can be made ready again later and the whole process can repeat.

Basic and extended tasks behave identically with respect to the ready, running and suspended states. Extended tasks, however, can also enter the waiting state. An extended task moves from the running to the waiting state when it voluntarily suspends itself by waiting on an event.

An event is simply an OS object that is used to provide an indicator for a system event. Examples of events include data becoming ready for use or sensor values being read. You can find out more about events in Chapter 7.

When an extended task enters the waiting state, then the OS will dispatch the highest priority task that is ready to run. When an event is set, the task is moved from the waiting to the ready state. Note that extended tasks return to the ready state and not the running state. This is because, during the time that the extended task was in the waiting state, some other higher priority task may have been activate and then dispatched.

4.2.2 Task Priorities

AUTOSAR OS allows tasks to share priorities. When tasks have the same priority, each task at the shared priority will run in mutual exclusion from each other. This means that if one task is running, then its execution will be serialized with all other tasks that share the same priority.

When tasks share priorities they are released from the ready state in first-in, first-out (FIFO) order.



When shared priorities and queued task activation are used together, RTA-OS3.x maintains an internal queue at the priority level. You should avoid this type of configuration if you want a fast and efficient OS.

If you need to serialize the execution of a set of tasks, then this is best achieved using unique priorities and AUTOSAR OS's internal resources (see Section 6.5) rather than sharing task priorities. Using internal resources guarantees serialization, just like sharing priorities, and the uniqueness of task priorities means that when multiple tasks become ready at the same time the OS has a statically defined dispatch ordering



Sharing priorities between tasks is bad real-time programming practice because it prevents you from performing schedulability analysis on your system. This is because, in the general case, sharing priorities makes the release point for a task (i.e. the point from where a response time is measured) computationally impossible to calculate. If it is impossible to work out when the release occurs then it is impossible to decide if the task will meet its deadline!

4.2.3 Queued Task Activation

Under most circumstances you will only activate a task when it is in the suspended state. In fact AUTOSAR OS treats the activation of a task while it is in the ready, running or waiting states as an error case.

However, there are some situations where you may need to implement a system where the same task must be activated a number of times but the shortest time between successive activations can be less than the time needed to run the task. For example, you might be unpacking CAN bus frames in a task and need to handle transient bursting of frames on the network.

This means you need to queue task activations at run time. AUTOSAR OS allows you to queue the activation of basic tasks to help you build this kind of application. Like other things in AUTOSAR OS the size of the task queue is statically configured. You must specify the maximum number of activations that can be pending for the task.

If the queue is already full when you try and activate the task then this will be handled as an error and the activation will be ignored.

Of course, you might have tasks that share priorities and use queued activation. In this case, tasks are queued in FIFO order in a queue with a length equal to the sum of the queue lengths for each task that shares the same priority. However, each task can only use up to its own number of entries.

4.3 Conformance Classes

You now know that tasks can:

- Be basic or extended
- Can share priorities
- Can queue activations.

However, AUTOSAR OS places some restrictions on what kind of features be used together. These are called Conformance Classes and are used to group task features for ease of understanding, enable partial implementations of the standard and provided scalability for different classes of application.

AUTOSAR OS has four conformance classes:

BCC1 - Basic tasks, unique priority and no queued activation.

BCC2 - Basic tasks, shared priorities and/or queued activation.

ECC1 - Extended tasks, unique priority and no queued activation. An ECC1 task is like a BCC1 task, but it can wait on events.

ECC2 - Extended tasks, shared priorities and no queued activation. Note that, unlike BCC2 tasks, ECC2 tasks cannot queue activations.

The following table gives a quick summary of the types tasks that can be used in different classes of AUTOSAR OS system:

System Class	Basic Tasks	Extended Tasks	Shared Task Priorities	Queued Task Activation
BCC1	✓	X	X	X
BCC2	✓	X	✓	✓
ECC1	✓	✓	X	X
ECC2	✓	✓	✓	✓ ¹

Each conformance class requires more resources - a system that is BCC1 will be much faster and smaller than a system which is ECC2. You do not need to be concerned about which conformance class to use - RTA-OS3.x supports all conformance classes and will calculate the conformance class from your OS configuration.

¹But only for basic tasks within the ECC2 system. Activations of extended tasks cannot be queued.

4.4 Maximizing Performance and Minimizing Memory

RTA-OS3.x is designed to be very aggressive at minimizing code and data usage on the target application. It will analyze the characteristics of the application and generate a system containing only the features that are required.

Your choice of task characteristics has a major influence on the final application size and speed. There is “no such thing as a free lunch”, so as you add tasks to your application that use more advanced types of tasks, the system will inevitably become slightly larger and slower.

A system with one or more BCC2 tasks has a greater overhead than one with only BCC1 tasks. A system without shared priorities, even if multiple activations are allowed, will be more efficient than one with shared priorities.

A system with ECC1 tasks has an even greater overhead still and a system with one or more ECC2 tasks has the largest overhead of all.

To make RTA-OS3.x as efficient as possible you should use basic tasks only and not share priorities.

4.5 Task Configuration

Unlike other real-time operating systems that you might have seen, the tasks in AUTOSAR OS (and, therefore, RTA-OS3.x) are defined statically. This technique is used because it saves RAM and execution time.

Tasks cannot be created or destroyed dynamically. Most of the information about a task can be calculated offline, allowing it to be stored in ROM.

The maximum number of tasks supported by RTA-OS3.x depends upon your port and you should consult the *Target/Compiler Port Guide* for further details. For all ports, RTA-OS3.x can provide a highly optimized system if you limit your number of tasks to the native word size of your microcontroller.

Device Type	Maximum	Optimal
8-bit	256	16 or fewer
16-bit	256	16 or fewer
32-bit	1024	32 or fewer

When you configure your task properties, you will most likely use the [rtaoscfg](#) configuration tool. Figure 4.5 shows the task configuration entry.

An AUTOSAR task has 5 attributes:

Name. The name is used to refer to, or provide a handle to, C code that you will write to implement the task functionality.

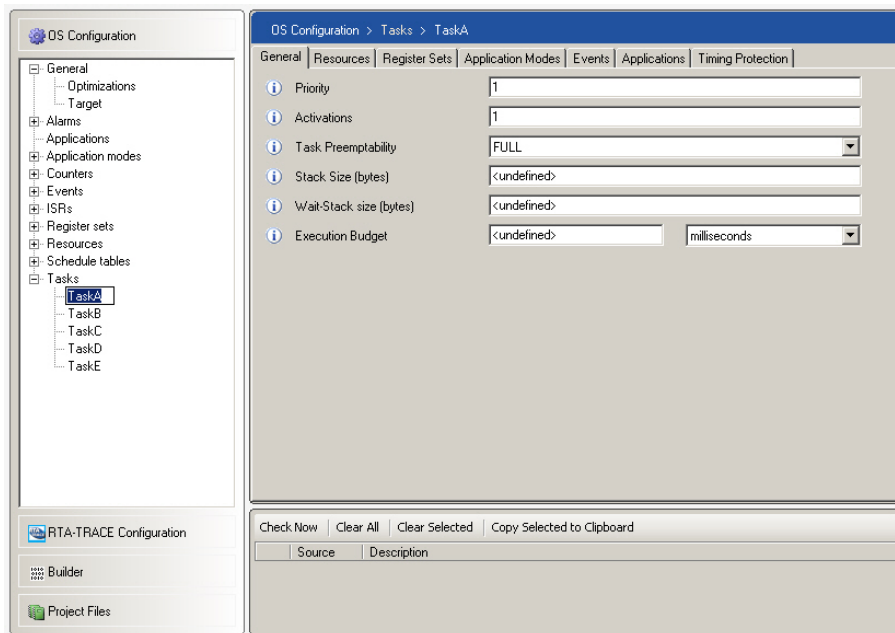


Figure 4.5: Task Configuration

Priority. The priority is used by the scheduler to determine when the task runs. Priorities cannot be changed dynamically. Zero is the lowest possible task priority in RTA-OS3.x. Higher task priorities are represented by larger integers. Tasks can share priorities, but if you are building a real-time system, then you should not do this because it cannot be analyzed.

Scheduling. A task can run fully preemptively or non-preemptively. In general, fully preemptive scheduling should be selected over non-preemptive scheduling for best application performance.

Activations. The maximum number of task activations that can be queued in the ready state. For a BCC1, ECC1 and ECC2 tasks the number of activations is always one. This means that these types of task can only be activated if they are in the suspended state. Any attempt to activate a such a task when it is not suspended will result in an error. A value greater than one indicates that the OS will queue activations (for example to smooth out transient peak loads in your application).

Autostart. This controls whether the task is started automatically when you start the OS.



The number of tasks that can be defined is fixed for each target (it is usually 256 or 1024, depending on the target processor). The Target/Compiler Port Guide for your target will contain further information.

4.5.1 Scheduling Policy

A fully preemptable task can be preempted by a task of higher priority. That means that when a higher priority task is made ready to run, it will run in preference.

You can prevent a task from being preempted by declaring it to be non-preemptable at configuration time. Tasks that are declared as non-preemptive cannot be preempted by other tasks. When a non-preemptive task moves to the running state it will run to completion and then terminate (unless it makes a `Schedule()` call, as explained in Section 4.10). Making tasks non-preemptive therefore means that if a lower priority task is started before a higher priority task, then the higher priority task will be prevented from executing for the time that the lower priority task runs. This is called blocking. Systems that use non-preemptive tasks will, in general, be less responsive than systems that run preemptively.

Even if a task is non-preemptive, it can still be interrupted by ISRs.

You will often find that it is unnecessary to use non-preemptable tasks because there are other, more suitable methods, which you can use to achieve the same effect. If you use these other techniques, it will usually result in a more responsive system. You will find out more about these techniques later, but they include:

- Using standard resources to serialize access to data or devices.
- Using internal resources to specify exactly which other tasks cannot cause preemption.

4.5.2 Queued Activation

Under most circumstances you will only activate a task when it is in the suspended state. However, you may need to implement a system where the same task must be activated a number of times and where the shortest time between successive activations is less than the time needed to run the task.

If this happens you will be activating the task while it is in the ready state or the running state. This means that activations will be lost.

To prevent loss of activations, you must specify the maximum number of multiple activations required for the task.



In accordance with the AUTOSAR OS standard, this feature is only available for basic tasks. You cannot specify multiple activations for extended tasks.

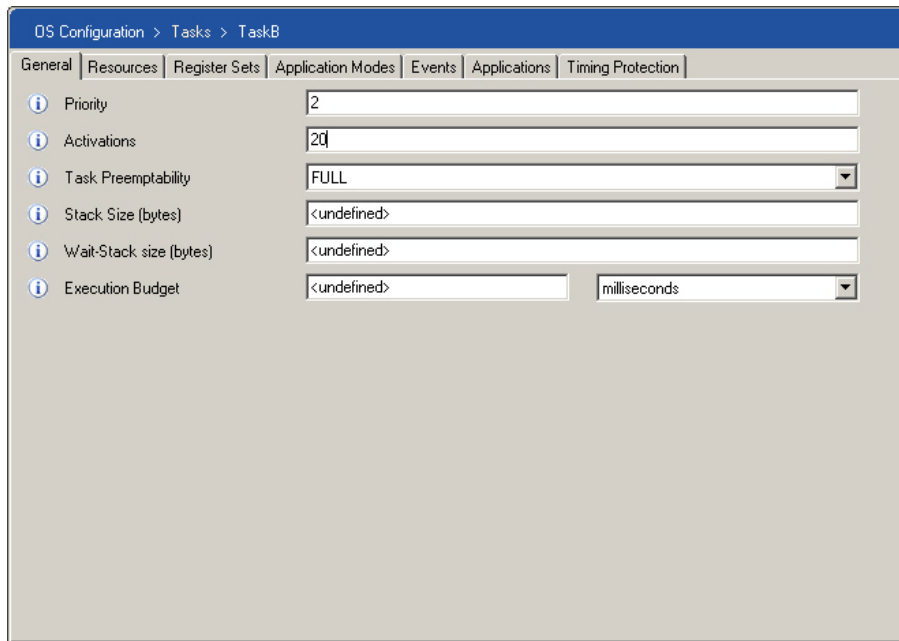


Figure 4.6: Specifying the number of queued activations

You will use `rtaoscfg` to specify the maximum number of simultaneous task activations. Figure 4.6 shows that, for the task in this example, the maximum number of activations has been set to twenty.

When multiple activations are specified, RTA-OS3.x automatically identifies that the task is BCC2. When you build your application, RTA-OS3.x will calculate the maximum size of the multiple activation queue needed for each BCC2 task.

When BCC2 tasks share priorities, RTA-OS3.x uses a FIFO queue to hold pending activations. If a BCC2 task has a unique priority in your AUTOSAR OS application then RTA-OS3.x automatically optimizes the queuing strategy to counted activation. Counted activation is significantly more efficient than FIFO activation and should be used wherever possible.

4.5.3 Auto-starting Tasks

Tasks can be auto-started, which means that when the operating system starts, they are activated automatically during `StartOS()`.

For basic tasks, which start, run and then terminate, auto-starting a task will make it run exactly once before it will return to the suspended state (from where it can be activated again). Auto-starting is mainly useful for starting extended tasks that wait on events because it removes the need to write code to activate the tasks.

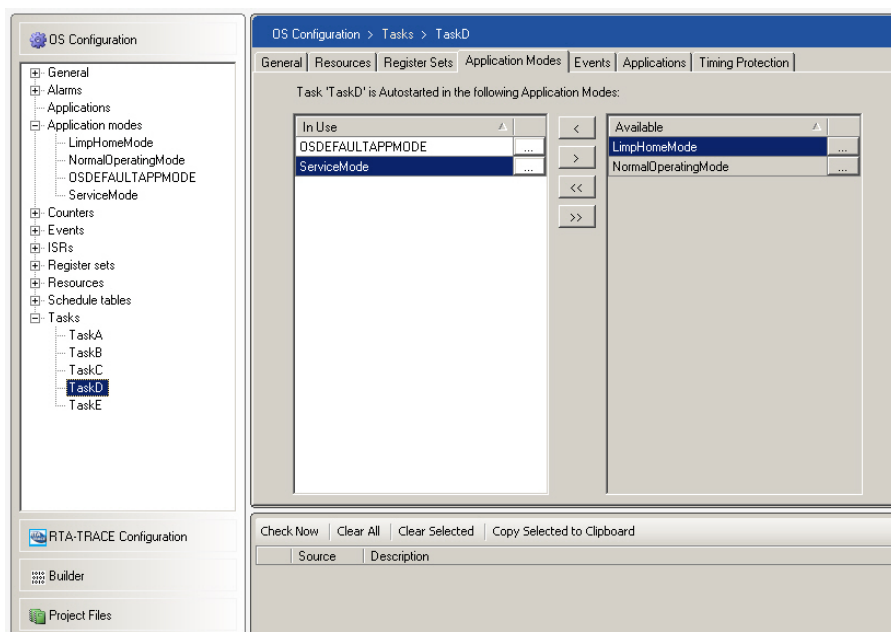


Figure 4.7: Configuring auto-started tasks

rtaoscfg can be used to specify that a task is only auto-activated in specific application modes, choose the application mode in question and select the tasks that you want to auto activate.

In Figure 4.7, TaskD is auto-started in the OSDEFAULTAPPMODE and ServiceMode application modes and is not autostarted in LimpHomeMode and NormalOperatingMode

4.6 Stack Management

RTA-OS3.x uses a single-stack model which means that all tasks and ISRs run on a single stack. The single stack is simply the C stack for the application.

As a task runs, its stack usage grows and shrinks as normal. When a task is preempted, the higher priority task's stack usage continues on the same stack (just like a standard function call). When a task terminates, the stack space it was using is reclaimed and then re-used for the next highest priority task to run (again, just as it would be for a standard function call). Figure 4.8 shows how the single stack behaves as tasks are stated, preempted and terminate.

In the single stack model, the stack size is proportional to the number of priority levels in the system, not the number of tasks/ISRs. This means that tasks which share priorities, either directly, or by sharing internal resources, or through being configured as non-preemptive, can never be on the stack at

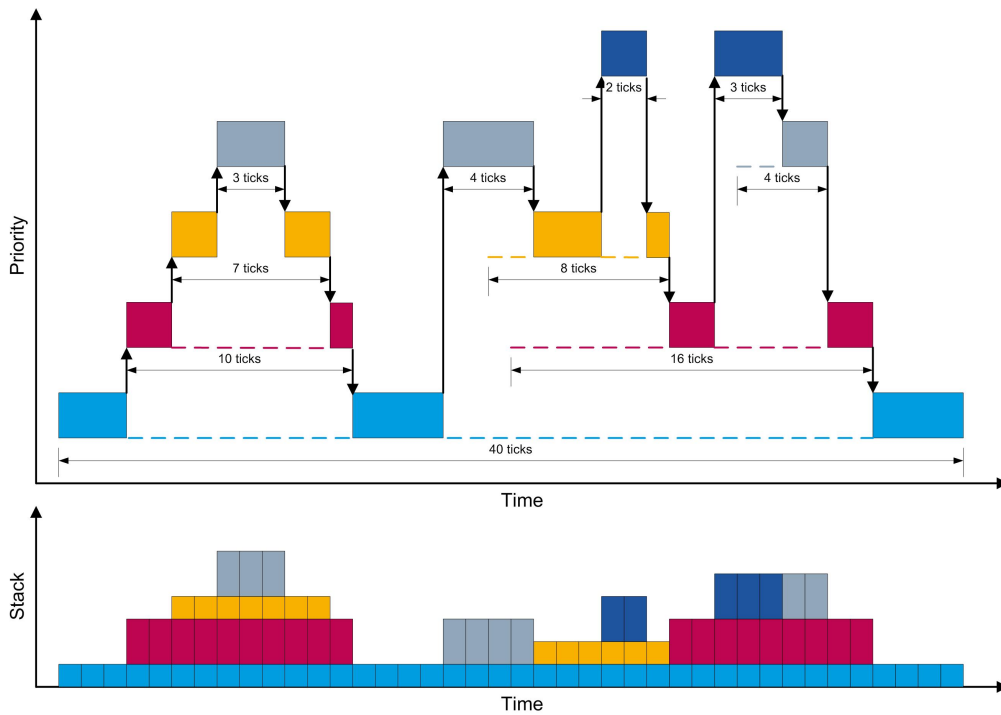


Figure 4.8: Single-stack behavior

the same time. The same is true of ISRs that share priorities in hardware. This means that you can trade system responsiveness, i.e. how long it takes for a task or ISR to complete, for stack space by simple changes to configuration.

Figure 4.9 shows the execution of the same task set, with the same arrival pattern as Figure 4.8 but this time the tasks are scheduled non-preemptively. You can see that the response times for the higher priority tasks are much longer than when they were preemptively scheduled but the overall stack consumption is much lower.

The single stack model also significantly simplifies the allocation of stack space at link time as you need only allocate a single memory section for the entire system stack, in exactly the same way as if you were not using an OS at all.

4.6.1 Working with Extended Tasks

RTA-OS3.x uniquely extends the single stack model to provide support for extended tasks without any impact on the performance of basic tasks.

In RTA-OS3.x, the lifecycle of an extended task is as follows:

Suspended → Ready The task is added to the ready queue.

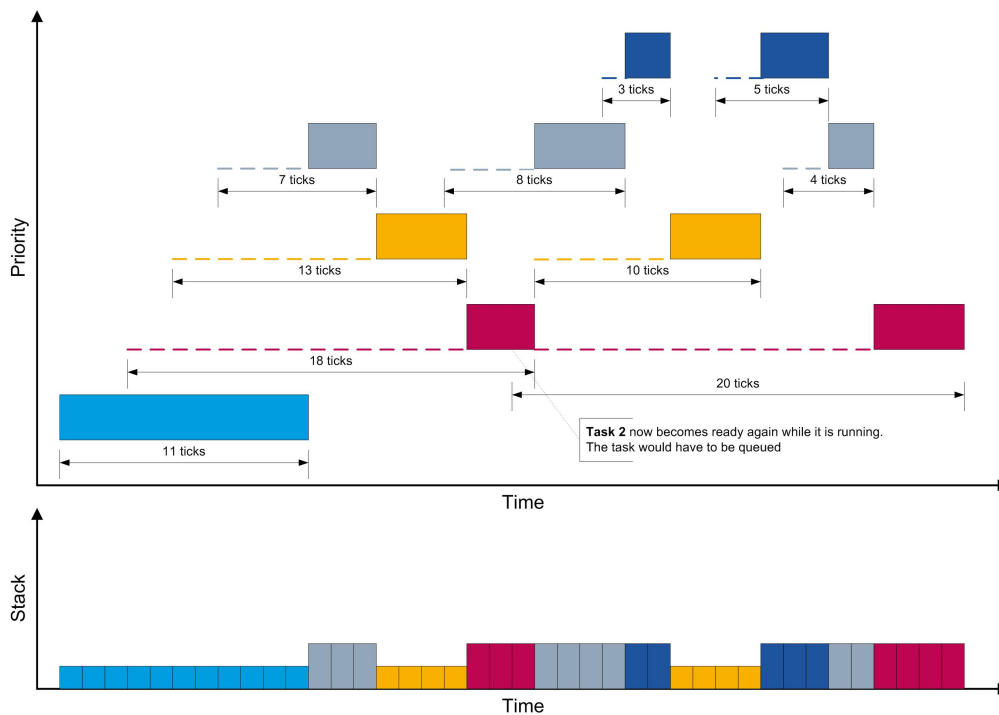


Figure 4.9: Single-stack behavior with non-preemptive tasks

Ready → Running The task is dispatched but, unlike a basic task where the context is placed in the top of the stack, the context is placed in the stack space at the pre-calculated worst case preemption depth of all lower priority tasks.

Running → Ready The extended task is preempted. If the preempting task is a basic task, then it is dispatched on the top of the stack as normal. If the preempting task is an extended task, then it is dispatched at the pre-calculated worst case preemption depth of all lower priority tasks.

Running → Waiting The task's *Wait Event Stack* context, comprising the OS context, local data, stack frames for function calls, etc, is saved to an internal OS buffer

Waiting → Ready The task is added to the ready queue.

Running → Suspended The task's "Wait Event Stack" context is copied from the internal OS buffer back onto the stack at the pre-calculated worst case preemption depth of all lower priority tasks.

This process allows the additional cost of managing extended tasks to apply only to extended task themselves. Basic tasks in system including extended tasks have the same performance as they would have in a basic task only system.

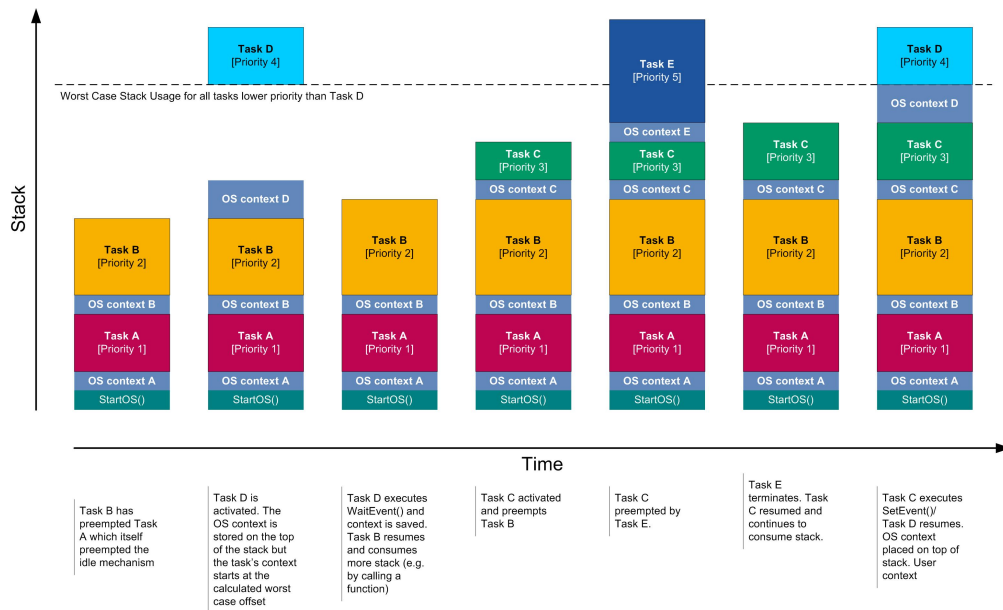


Figure 4.10: Single-stack management with Extended Tasks

The key parts of this lifecycle are the dispatch/resume at the worst case preemption depth and the copy on and off the stack. The dispatch at the worst case preemption point guarantees that whenever an extended task resumes after waiting, it can resume with its local variables at exactly the same location in memory. It is guaranteed that every possible preemption pattern of lower priority tasks will never exceed the dispatch point of the extended task. The dispatch-wait-resume cycle for an extended task D is illustrated in Figure 4.10.

The copy off and on allows the extended tasks stack context to be restored. This is necessary because higher priority tasks and/or ISRs may occur while the extended task is waiting. These may consume stack space greater than the worst case preemption point (remember that the worst case point is for lower priority objects only), thereby overwriting the context of the extended task. However, fixed priority preemptive scheduling guarantees that no higher priority task can be ready to run at the point the extended task is resumed (it could not be resumed if this was the case).

Extended task management requires RTA-OS3.x to be told how much stack is used by tasks and ISRs. The following sections describe the various configuration parameters.

4.6.2 Mandatory Stack Information

The calculated worst case dispatch points define the number of bytes, relative to the address of the stack pointer when StartOS() is called, at which

an extended task needs to start. These offsets are stored as ROM data in the extended task control blocks and are added to the base address of the stack at runtime.

This means that RTA-OS3.x needs to be told various parameters about use of the stack. The values that are captured are port-specific and you should be read the *Target/Compiler Port Guide* for your port for additional guidance.

RTA-OS3.x provides runtime features for measuring the worst-case stack values for tasks and ISRs. See Section 14.3 for further information.

Typically, all ports will allow you to specify the following values, all figures are in bytes:

Stack used for C-startup (SpPreStartOS)

The amount of stack space in use at the point StartOS() is called.

This value is *not* required for the calculation of worst case dispatch points and can safely be set to zero. However, the value is used in the calculation of worst-case stack usage if a 'Stack Usage' report is requested.

Stack used when idle (SpStartOS)

The maximum amount of stack space used by Os_Cbk_Idle() if it is declared. If you do not declare an Os_Cbk_Idle() then this value can be set to zero.

Stack overheads for ISR activation (SpPreemption)

The number of additional bytes of stack used when returning from an ISR into a different task than the one preempted.

This is the stack size difference between resuming the task which was preempted and starting a new (higher priority) task.

Stack overheads for ECC tasks (SpECC)

The number of additional bytes of stack required to activate an extended task compared to a basic task.

This value can be obtained by measuring the stack value :

1. before an (upwards) activation of a basic task and immediately in the entry function of the activated task.
2. before an (upwards) activation of an extended task and immediately in the entry function of the activated task.

and then subtracting the second value from the first value.

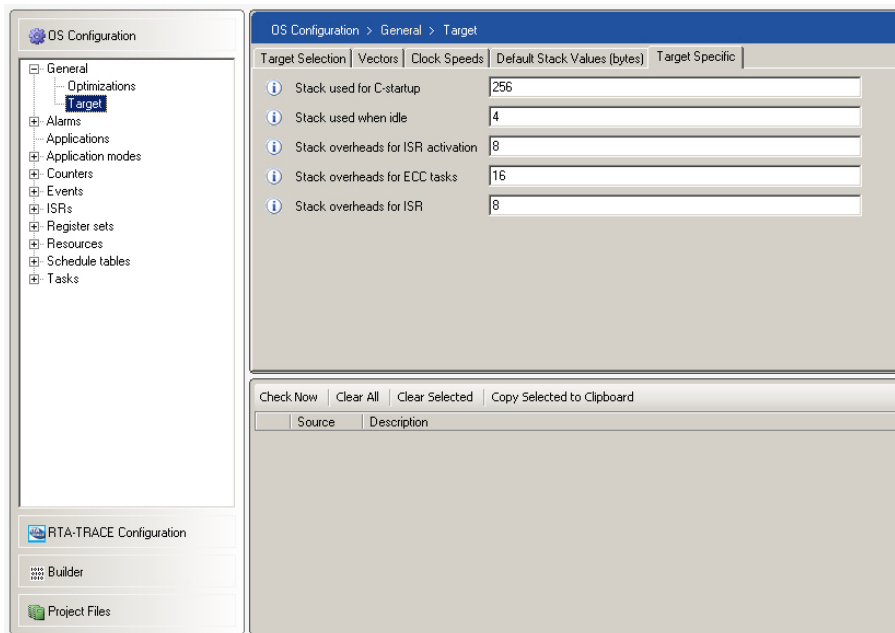


Figure 4.11: Essential Stack Values

Stack overheads for ISR (SpIDisp)

The number of bytes that are placed on the stack between a task or ISR being preempted and entering the Category 2 handler.

This value can be obtained by measuring the stack value before an interrupt occurs and immediately in the Category 2 ISR's entry function during your testing.

All of the mandatory stack values except Stack overheads for ISR (SpIDisp) are added to stack values during checking. This means that specifying larger values than actually occur is safe. However, Stack overheads for ISR (SpIDisp) is subtracted from the current value of the of stack pointer on entry to the ISR to check whether the preempted task or ISR has exceeded its stack usage. Therefore, specifying a large value here may result in an error being reported where it has not occurred (i.e. RTA-OS3.x stack management will report a "false positive").



Figure 4.11 shows the configuration of the essential stack values.

4.6.3 Specifying Task Stack Allocation

In systems that contain only basic tasks it is not necessary to tell RTA-OS3.x any stack allocation unless you are doing stack monitoring (see Section 14.1). You simply need to allocate a stack section large enough for your application in your linker/locator. This is one of the benefits of the single stack architecture.

For applications that use extended tasks, you allocate your linker section as before, but you must also tell RTA-OS3.x the stack allocation for every task in your configuration that is lower priority than the highest priority extended task, even if they are basic tasks. RTA-OS3.x uses the stack allocation information to calculate the worst case preemption point for each extended task off-line.

The stack allocation you specify is the entire stack used for the task and includes:

- the OS context
- space for local variables in the task body
- the space required for any functions called from the task body (and their locals)

You can use RTA-OS3.x's stack measurement feature to obtain accurate values for the stack allocation. See Section 14.1 for further details.



RTA-OS3.x only uses the stack information you provide to calculate the worst case preemption point. RTA-OS3.x does not reserve any stack space. You must still specify the stack application stack space in the same way you would do for a normal application.

Figure 4.12 shows how stack allocation is configured.

While RTA-OS3.x uses a single-stack model, on some ports this does not necessarily mean that just one *physical* stack is used. It may be the case that either the compiler or the hardware forces data onto different stacks automatically. For example, some devices place interrupts on to a dedicated interrupt stack.

Even with multiple physical stacks, RTA-OS3.x still provides the benefits of the single-stack architecture - the stack space required on every physical stack can be overlaid when tasks and/or ISRs share a priority level. However, for the stack allocation to work correctly you will need to specify the space needed on each stack. RTA-OS3.x will automatically ask you for multiple stack values if you configure a target for which this information is required. Figure 4.13 shows a dialogue box from such a configuration where there are two stacks: 'Supervisor' and 'Context'.

4.6.4 Optimizing the Extended Task context save

Recall from Section 4.6.1 that each time an extended task enters the waiting state, RTA-OS3.x saves the task's "Wait Event Stack" context and that the context is restored when the task re-enters the running state.

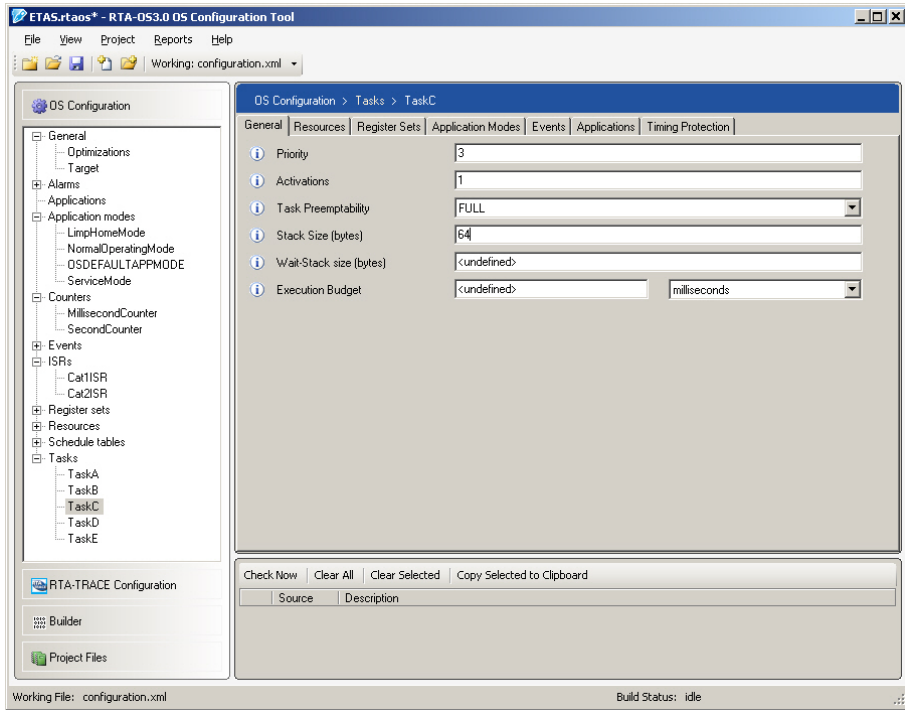


Figure 4.12: Stack Allocation Configuration

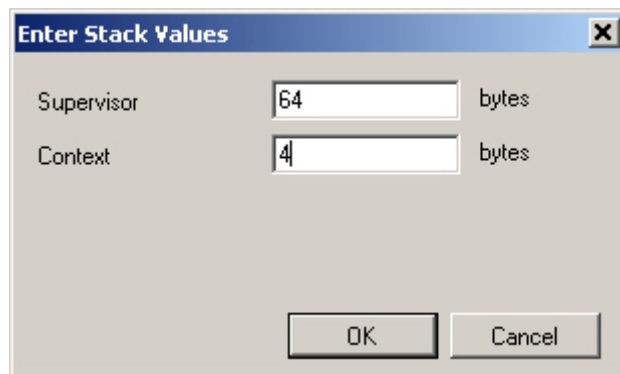


Figure 4.13: Stack Allocation Configuration for multiple stacks

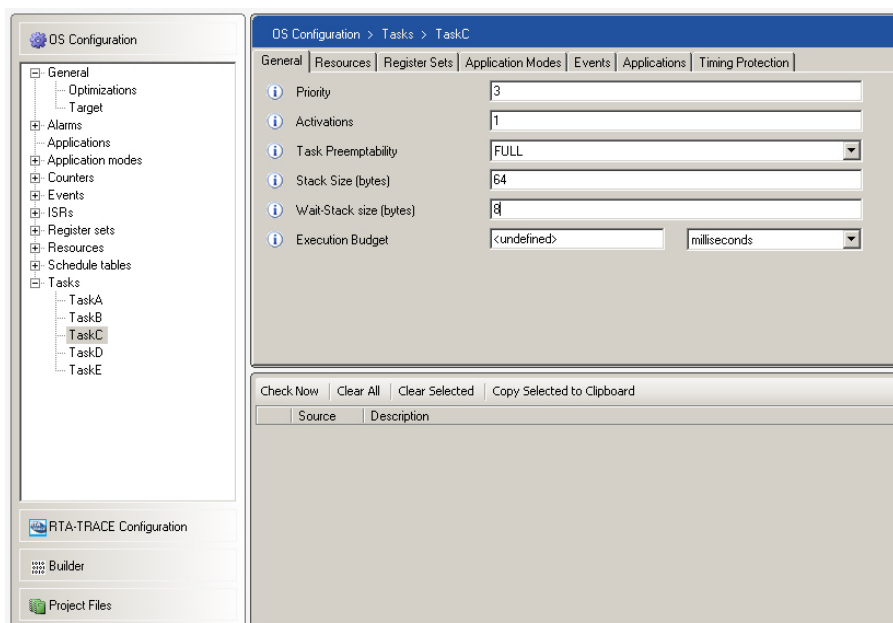


Figure 4.14: Specifying a WaitEvent () Stack allocation

RTA-OS3.x saves the “Wait Event Stack” context in an internal buffer. By default, RTA-OS3.x allocates a buffer equal to the worst case stack allocation you specify for the task. Assuming that your stack allocation is correct, this should always be enough to hold the worst case stack usage when you call WaitEvent ().

This sounds expensive because it appears that RTA-OS3.x needs to allocate twice the RAM you would expect for each extended task: once on the stack and once for the task’s save/restore buffer! However, RTA-OS3.x needs to save the context *only* when WaitEvent () is called. This means that you can significantly optimize the RAM size required by RTA-OS3.x when using extended tasks by allocating only enough buffer space to save the worst case “Wait Event Stack” context, rather than the absolute worst case space required by the task.

Typically, most applications that use extended tasks only call WaitEvent () from the task’s entry function where only a small amount of local data is on the stack so this optimization can be applied in most extended task systems.

You can control exactly how many bytes of stack are saved by RTA-OS3.x by specifying the worst case stack depth at the point you call WaitEvent () as shown in Figure 4.14.



If you leave the `WaitEvent()` Stack allocation as 'undefined' then RTA-OS3.x will default to use the number of bytes you specified for the stack allocation.

Using Default Values

While you should set a stack value for each task for memory efficiency, RTA-OS3.x allows you to set a global default value that is used by all tasks. This can be found in **General → Default Stack Values**.

If a Stack Allocation is not configured for a task, then RTA-OS3.x will use the default value for:

- Calculating the worst case stack offset
- Configuring the `WaitEvent()` save/restore area
- Stack Monitoring (when configured)

The specification of a task/ISR-specific stack allocation overrides the default value.

4.6.5 Handling Stack Overrun

If the stack allocation figures you provided to RTA-OS3.x are wrong (i.e. they are too small) then this is a potential source of errors at runtime. There are three things that can go wrong:

1. the extended task cannot start because the current value of the stack pointer is higher than the calculated worst case dispatch point when RTA-OS3.x tries to dispatch an extended task. This means one (or more) of the lower priority tasks that are on the stack has consumed too much space (Stack monitoring, described in Section 14.1, can be used to identify which task is at fault).
2. the extended task cannot resume from the waiting state because the stack pointer is higher it should be. This may occur when `SetEvent()` has been called for an event on which the extended task was waiting and the extended task is now the highest priority task in the system.
3. the extended task cannot enter the waiting state because the current amount of stack the task is using is greater than the size of the 'WaitEvent() stack' that was configured.

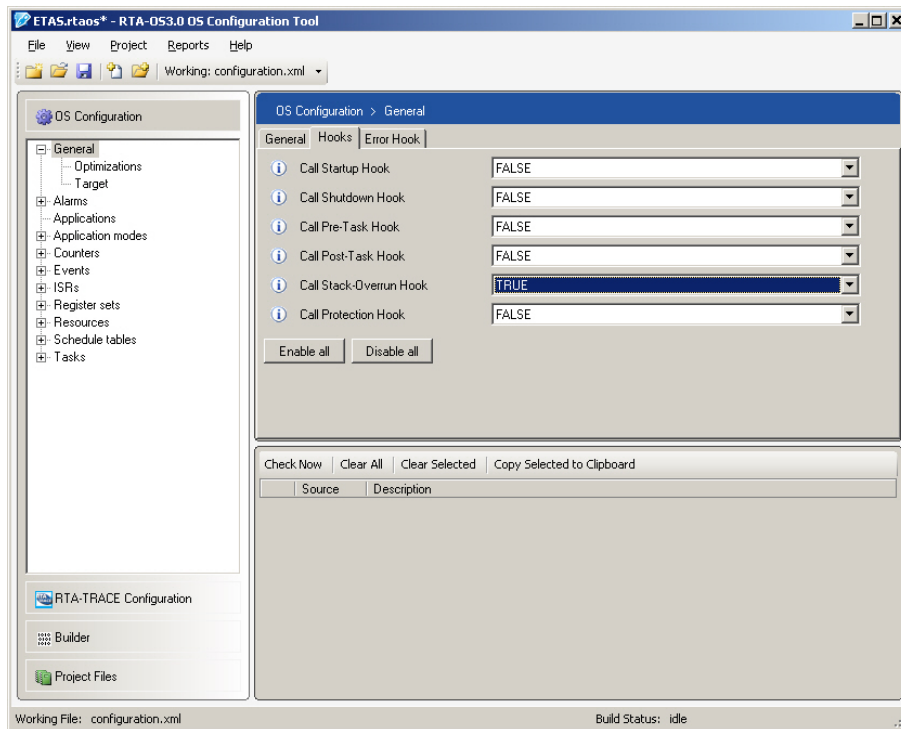


Figure 4.15: Enabling the `Os_Cbk_StackOverrunHook()`

When RTA-OS3.x detects a problem with extended task stack management it will call `ShutdownOS()` with the error code `E_OS_STACKFAULT`.

If you want to debug the problem then you can enable the stack fault hook as shown in Figure 4.15.

When configured, RTA-OS3.x will call the user-provided callback `Os_Cbk_StackOverrunHook()` instead of `ShutdownOS()` when a stack fault occurs. The callback is passed two parameters:

1. Overrun tells you the number of bytes of the overrun
2. Reason tells you what caused the overrun

For an extended task system without stack monitoring enabled, the overrun can be either:

- `OS_ECC_START` - the extended task could not start (or resume from waiting) because the current stack pointer exceeds the worst case dispatch point calculated at build time. The cause of this fault is that one (or maybe more) of the lower priority tasks has exceeded the configured

stack allocation. To fix this problem you need to identify which task is in error. Chapter 14 explains how to do this using RTA-OS3.x's stack monitoring feature.

- OS_ECC_WAIT - the extended task could not enter the waiting state because the amount of stack space it has consumed exceeds the configured WaitEvent() stack size. To fix this problem, you should increase the WaitEvent() stack size by at least the number of bytes indicated by the Overrun parameter.

Code Example 4.1 shows a simple example.

```
#ifndef OS_STACKOVERRUNHOOK
FUNC(void, OS_APPL_CODE) Os_Cbk_StackOverrunHook(Os_StackSizeType
    Overrun, Os_StackOverrunType Reason) {
{
    /* Identify problem */
    for(;;) {
        /* Do not return! */
    }
}
#endif /* OS_STACKOVERRUNHOOK */
```

Code Example 4.1: Minimum recommended Os_Cbk_StackOverrunHook()

4.7 Implementing Tasks

Tasks are similar to C functions that implement some form of system functionality when they are called by RTA-OS3.x.



You do not need to provide any C function prototypes for task entry functions. These are provided through the Os.h header file generated by RTA-OS3.x.

When a task starts running, execution begins at the task entry function. The task entry function is written using the C syntax in Code Example 4.2.

```
TASK(task_identifier)
{
    /* Your code */
}
```

Code Example 4.2: A Task Entry Function

Remember that basic tasks are single-shot. This means that they execute from their fixed task entry point and terminate when completed.

Code Example 4.3 shows the code for a basic task called BCC_Task.


```

#include <Os.h>
TASK(BCC_Task) {
    do_something();
    /* Task must finish with TerminateTask() or equivalent. */
    TerminateTask();
}

```

Code Example 4.3: A Basic Task

Now, compare the example in Code Example 4.3 with Code Example 4.4. Code Example 4.4 shows that extended tasks need not necessarily terminate and can remain in a loop waiting for events.

```

#include <Os.h>
TASK(ECC_Task) {
    InitializeTheTask();
    while (WaitEvent(SomeEvent)==E_OK) {
        do_something();
        ClearEvent(SomeEvent);
    }
    /* Task never terminates. */
}

```

Code Example 4.4: Extended Task Waiting for Events

4.8 Activating Tasks

A task can only run after it has been activated. Activation either moves a task from the suspended state into the ready state or it adds another entry to the queue of ready tasks (if the task supports multiple activation). The task will run once for each of the activations. It is an error to exceed the activation count and your application will generate `E_OS_LIMIT` errors when this happens (even in the Standard build status).

Tasks can be activated from both tasks and (Category 2) ISRs.

Activating a task does not cause the task to begin executing immediately, it just makes it ready to run. However, RTA-OS3.x needs to check whether the activated task has a higher priority than the currently running task and, if it does, cause a context switch to occur so the new task can preempt the currently running task.

When you activate a task RTA-OS3.x from another task, the exact behavior depends upon the relative task priorities. If the activated task has higher priority than the currently running task, then the newly activated task will preempt the current task. Otherwise, the task will remain on the ready queue until it becomes the highest priority ready task.

In a well-designed real-time system, it is unusual for a task to activate a higher priority task. Normally ISRs capture system triggers and then activate the tasks to do any associated processing. In turn, these tasks may activate lower priority tasks to implement trigger responses that have longer deadlines.

Observing this fact leads to one of the major optimizations in RTA-OS3.x. If you specify that your tasks never activate higher priority tasks, RTA-OS3.x can eliminate the internal code that tests whether a context switch is needed following each activation. This is configured by selecting the “Disallow Upwards Activation” optimization.

This is similar to the behavior when activating a task from an ISR. All ISRs have a priority that is strictly higher priority than the highest task priority. When a task is activated from an ISR it can never enter the running state immediately so it is never necessary to check for a context switch. Such a check is only necessary when leaving the ISR.

4.8.1 Direct Activation

Tasks can be activated in a number of different ways. The basic mechanism for task activation is the `ActivateTask()` API call, which directly activates a task. The `ActivateTask(TaskID)` call places the named task into the ready state. The `ChainTask(TaskID)` call terminates the calling task (see Section 4.11) and places the named task into the ready state.

API Call	Description
<code>ActivateTask()</code>	A task or ISR can make this call to activate the task directly.
<code>ChainTask()</code>	A task can make this call to terminate the currently running task and to activate the task indicated.

4.8.2 Indirect Activation

Besides directly activating tasks, it is possible to use other AUTOSAR OS mechanisms to indirectly activate a task. These methods are described in more detail in later chapters of this user guide.

Activation by an Alarm. For each alarm in the system, you can specify a task that is activated each time the alarm expires.

Activation by a Schedule Table. For each schedule table in the system, you can specify a task that is activated on one or more expiry points on the table.

4.9 Controlling Task Execution Ordering

In many cases, you will need to constrain the execution order of specific tasks. This is particularly true in data flow based designs where one task needs to perform some calculation before another task uses the calculated value. If the execution order is not constrained, a race condition may occur and the application behavior will be unpredictable. Task execution ordering can be controlled in the following ways:

- Direct activation chains (see Section 4.9.1).
- Priority levels (see Section 4.9.2).
- Non-preemptable tasks (see Section 2).

4.9.1 Direct Activation Chains

When you use direct activation chains to control the execution order, tasks make `ActivateTask()` calls on the task(s) that must execute following the task making the call.

Consider the following; there are three tasks Task1, Task2 and Task3 that must execute in the order Task1, then Task2, then Task3. Code Example 4.5 shows example task bodies.

```
#include <Os.h>
TASK(Task1) {
    /* Task1 functionality. */
    ActivateTask(Task2);
    TerminateTask();
}

TASK(Task2) {
    /* Task2 functionality. */
    ActivateTask(Task3);
    TerminateTask();
}

TASK(Task3) {
    /* Task3 functionality. */
    TerminateTask();
}
```

Code Example 4.5: Using Direct Activation Chains

Figure 4.16 shows how these tasks would execute assuming that Task1 has the highest priority and Task 3 has the lowest priority.

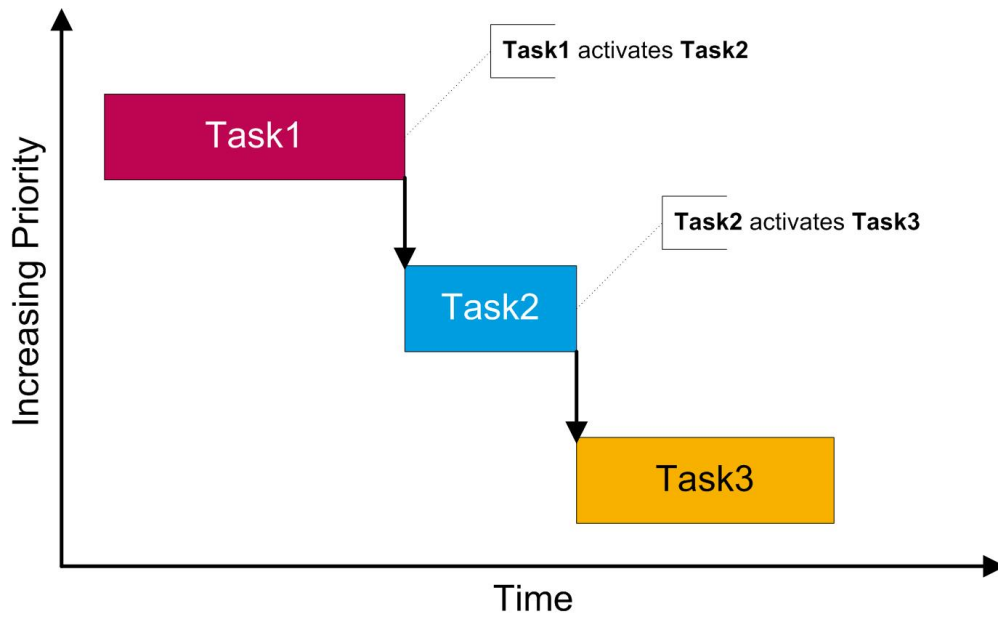


Figure 4.16: Direction activation to control task execution order

4.9.2 Using Priority Levels

The priority level approach to constraining task execution ordering can be used to exploit the nature of the preemptive scheduling policy to control activation order.

Recall from Section 4.1 that, under fixed priority preemptive scheduling, the scheduler always runs the highest priority task. If a number of tasks are released onto the ready queue, they will execute in priority order. This means that you can use task priorities to control execution order.

Following on from our previous example, in Code Example 4.5, let's assume that Task1 has the highest priority and Task3 has the lowest priority. This means that the task bodies can be rewritten to exploit priority level controlled activation. This can be seen in Code Example 4.6.

```
#include <Os.h>
TASK(Task1) {
    /* Task1 functionality. */
    ActivateTask(Task2); /* Runs when Task1 terminates. */
    /* More Task1 functionality. */
    ActivateTask(Task3); /* Runs when Task2 terminates. */
    TerminateTask();
}

TASK(Task2) {
    /* Task2 functionality. */
```

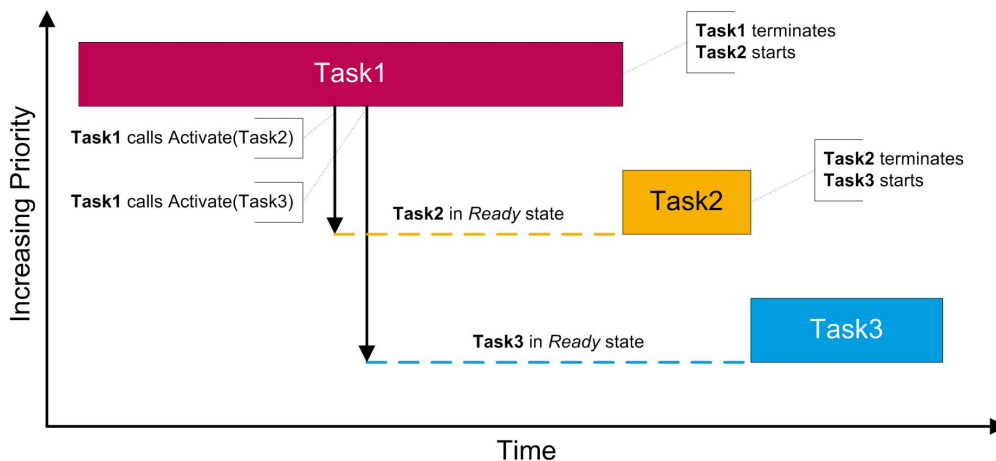


Figure 4.17: Using priority to control task execution order

```

    TerminateTask();
}

TASK(Task3) {
    /* Task3 functionality. */
    TerminateTask();
}

```

Code Example 4.6: Using Priority Level Controlled Activation

```
/* Task1 functionality. */
```

Figure 4.17 shows how these tasks would execute.

4.10 Co-operative Scheduling in RTA-OS3.x

When a task is running non-preemptively, it prevents any task (including those of higher priority) from executing. Sometimes, however, it is useful for non-preemptive tasks to offer explicit places where rescheduling can take place. This is more efficient than simply running non-preemptively because higher priority tasks can have shorter response times to system stimuli. A system where tasks run non-preemptively and offer points for rescheduling is known as a co-operatively scheduled system.

The `Schedule()` API call can be used to momentarily remove the preemption constraints imposed by both the non-preemptive tasks and the tasks using internal resources.

When `Schedule()` is called, any ready tasks that have a higher priority than the calling task are allowed to run. `Schedule()` does not return until all higher priority tasks have terminated.

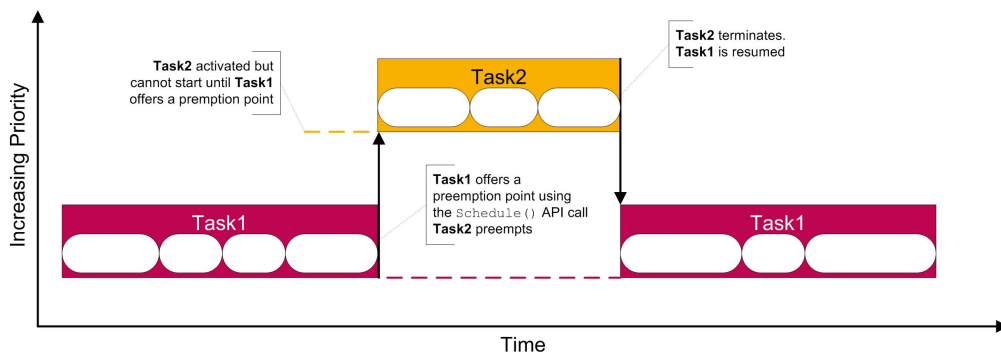


Figure 4.18: Co-operative tasks

In the following code example, the non-preemptive task Cooperative includes a series of function calls. Once started, each function runs to completion without preemption, but the task itself can be preempted between each function call.

```
#include <Os.h>
TASK(Cooperative){
    Function1();
    Schedule();/* Allow preemption */
    Function2();
    Schedule();/* Allow preemption */
    Function3();
    Schedule();/* Allow preemption */
    Function4();
    TerminateTask();
}
```

Figure 4.18 shows how two tasks, Task1 and Task2, which are co-operative would interact. The white sections represent non-preemptible sections of code.

4.10.1 Optimizing out the Schedule() API

Schedule() is of no use in a fully preemptive system. If you do not intend to use it, you can disallow calls to Schedule() in **rtaoscfg** using the "Optimizations, RTA-OS, Disallow Schedule()". If you disallow calls to Schedule() then you will see that the worst-case stack requirement for the system is reduced.

4.11 Terminating Tasks

Tasks that terminate in AUTOSAR OS must make an API call to tell the OS that this is happening. The AUTOSAR OS standard defines two API calls for task termination. One of these must be used to terminate any task. These API calls are:

- TerminateTask()
- ChainTask(TaskID)

When a task has finished, it must make one of these API calls. This ensures that RTA-OS3.x can correctly schedule the next task that is ready to run.

TerminateTask() forces the calling task into the suspended state. RTA-OS3.x will then run the next highest priority task in the ready state.

ChainTask(TaskID) terminates the calling task and activates the task TaskID. The API is therefore like executing a TerminateTask() followed immediately by ActivateTask(TaskID). Chaining a task places the named task into the ready state.

4.11.1 Optimizing Termination in RTA-OS3.x

The AUTOSAR OS standard allows task termination API calls to be called by a task at any point, including within a deeply nested set of function calls. This is bad programming practice - equivalent to the use of **goto**. At runtime, RTA-OS3.x must store information that allows it to clear the stack when the task terminates somewhere other than the entry function. This is normally done using a setjmp/longjmp pair.

Code Example 4.7 shows a task that makes nested calls to other functions. When Task1 runs, it calls Function1(). Function1() then calls Function2(). Function2() contains codes that can terminate the calling task (in this example, this is Task1).

```

/* Include Header file generated by \RTAOS */
#include <Os.h>

void Function1(void) {
    ...
    Function2();
    ...
}

void Function2(void) {
    if (SomeCondition) {
        TerminateTask();
    }
}

TASK(Task1) {
    /* Make a nested function call. */
    Function1();
    /* Terminate the task in the entry function*/
}

```

```
    TerminateTask();  
}
```

Code Example 4.7: Terminating a Task

However, one of the key benefits of the a single-stack architecture is that a task which terminates in its entry function can simply return - `TerminateTask()` does not need to do anything. If all your tasks either do not terminate or only terminate in their entry function, then the context that RTA-OS3.x saves to allow a return from anywhere does not need to be stored.

RTA-OS3.x allows you to exploit good application design using the fast termination optimization (**Optimizations → Fast Terminate**). You can enable this optimization when all tasks that execute the `TerminateTask()` or `ChainTask()` APIs only do so in their entry function. The optimization tells RTA-OS3.x not to generate code to save unnecessary context and, as a result, save stack space.

4.12 The Idle Mechanism

Any preemptive operating system must have something to do when there are no tasks or ISRs to run. In AUTOSAR OS this is achieved by an idle mechanism. In RTA-OS3.x the OS will sit in a busy wait loop doing nothing when there are no tasks or ISRs to run.

However, you can override the default behavior by providing your own implementation of the idle mechanism by declaring a callback called `Os_Cbk_Idle`.

The `Os_Cbk_Idle` behaves in the same way as a task except that:

- it cannot be activated
- it cannot be terminated
- it cannot wait for events
- it cannot be chained
- it cannot use internal resources

The `Os_Cbk_Idle` has the lowest priority of any task in the system, so it runs only when there are no tasks (or ISRs) that are ready to run. The idle mechanism therefore gives you an “extra task” that is almost entirely free from system overheads.

Code Example 4.8 shows an implementation of `Os_Cbk_Idle` that is used to control RTA-TRACE (see Chapter 17).


```

#include <Os.h>
FUNC(boolean, OS_APPL_CODE) Os_Cbk_Idle() {
    #ifdef OS_TRACE
        CheckTraceOutput();
        UploadTraceData();
    #endif /* OS_TRACE */
    return TRUE;
}

OS_MAIN()
{
    /* System hardware initialization. */
    StartOS(OSDEFAULTAPPMODE);
    /* The call never returns */
}

```

Code Example 4.8: An Idle Mechanism

Os_Cbk_Idle returns a boolean on exit that tells RTA-OS3.x whether or not to call Os_Cbk_Idle again. When TRUE is returned then RTA-OS3.x immediately calls Os_Cbk_Idle again. When FALSE is returned then RTA-OS3.x stops calling Os_Cbk_Idle and enters the default behavior of sitting in a busy wait loop.

4.13 Pre and Post Task Hooks

Suppose that you need to execute some code before each task starts and/or after each task ends, for example to profile a trace of execution. You can do this using the PreTask and PostTask hooks provided by AUTOSAR OS.

The PreTask Hook is called by RTA-OS3.x whenever a task moves into the running state. This means that the PreTask Hook will also be called whenever a task is resumed after preemption.

The PostTask Hook is called by RTA-OS3.x whenever a task moves out of the running state. The PostTask Hook will be called when the task terminates and each time a task is preempted.

Figure 4.19 shows where the PreTask and PostTask Hooks are called relative to task preemption.

Both of these hooks are only called when configured. Figure 4.20 shows how to enable the hooks.

Code Example 4.9 shows how the hooks should appear in your code.

```

FUNC(void, OS_APPL_CODE) PreTaskHook(void) {
    /* PreTask hook code. */
}

```

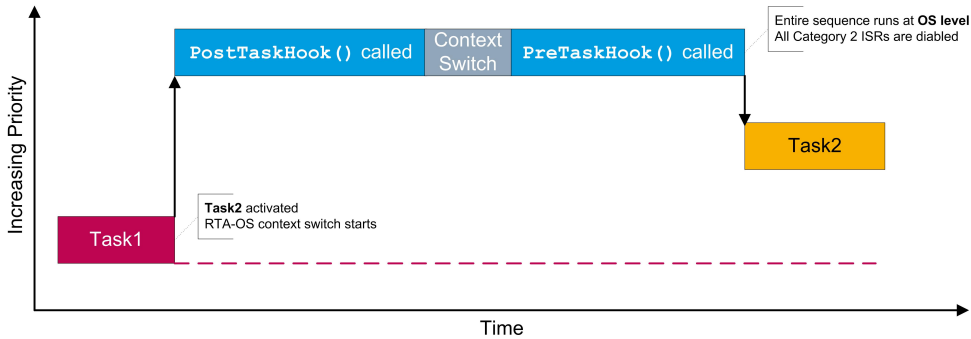


Figure 4.19: The `PreTaskHook()` and `PostTaskHook()` Relative to Task Pre-emption

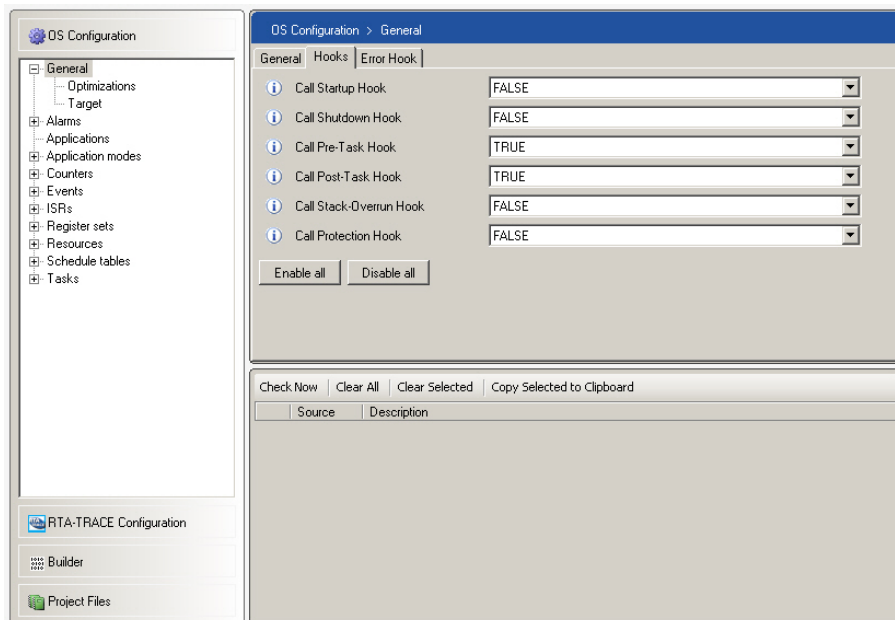


Figure 4.20: Enabling the `PreTaskHook()` and `PostTaskHook()`

```

FUNC(void, OS_APPL_CODE) PostTaskHook(void) {
    /* PostTask hook code. */
}

```

Code Example 4.9: The PreTaskHook and PostTaskHook

The PreTask and PostTask Hooks are called on entry and exit of tasks and for each preemption/resumption. This means that it is possible to use these hooks to log an execution trace of your application. Since the same PreTask and PostTask Hooks must be used for all of the tasks in the application, it is necessary to use the `GetTaskID()` API call to work out which task has been or will be running when the hook routine is entered.

RTA-OS3.x defines a set of macros that are only defined if the corresponding hook is enabled. These macros are called:

- `OS_PRETASKHOOK`
- `OS_POSTTASKHOOK`

This allows you write code where the hooks can be conditionally compiled as shown in Code Example 4.10.

```

#ifdef OS_PRETASKHOOK
FUNC(void, OS_APPL_CODE) PreTaskHook (void)
{
    /* Your code */
}
#endif /* OS_PRETASKHOOK */

```

Code Example 4.10: Conditional Compilation of PreTaskHook

4.14 Saving Hardware Registers across Preemption

RTA-OS3.x saves as little context as necessary on a context switch - only the context for the correct operation of the OS is saved. However, you may find that you need to save and restore additional application-dependent context at runtime. For example, you may have tasks that use floating-point registers and therefore will require the floating-point context for your microcontroller to be saved across context switches.

You could choose to implement this by hand using the PreTask and PostTask hooks and an application-managed stack. However, it becomes difficult to optimize this type of implementation without making it fragile to changes in OS configuration. You can either:

- always save the context on every switch into a task and then restore on every switch out.

This model means you might be making unnecessary saves and restores (for example, saving a register set when switching into a task that doesn't use it); or

- calculate the saves required offline and then write a more complex pair of hooks that use `GetTaskID()`/`GetISRID()` to work out if a save/restore is needed.

This model is fragile because changes to the configuration, for example adding new tasks/ISRs or modifying priorities, will mean that re-work is necessary.

To avoid these issues, RTA-OS3.x provides a simple general-purpose mechanism for saving user-specific context together with the OS context. RTA-OS3.x is able to exploit its knowledge of the priority space to calculate exactly which tasks need to save register sets at runtime so that unnecessary saves are optimized away automatically, saving both the time and stack required for the context switch. For example:

- if you only have one task or ISR that saves a given register set then no save or restore is needed.
- if multiple tasks use the same register set but cannot execute at the same time (because they are non-preemptable, share an internal resource or share priority) then RTA-OS3.x does not need to save the register set.
- a context switch into the lowest priority task that uses a register set does not need to do a save because it can be guaranteed that no other task can be using the set (because the lowest priority task could not be running if a higher priority task was using the register set).
- similarly, a context switch from the highest priority task that uses a register set does not need to do a save because no higher priority task uses the register set and therefore cannot corrupt the context.

Figure 4.21 shows a register set that is shared by tasks 1, 3 and 5. You can see that when a save is not needed (when switching into a task that does not use the register set) then no context save is made.

Each register set you need to save needs to be declared to RTA-OS3.x at configuration time. `rtaosgen` uses the declaration to define two callback functions that you must provide to save and restore the register set. Figure 4.22 shows the definition of three register sets.

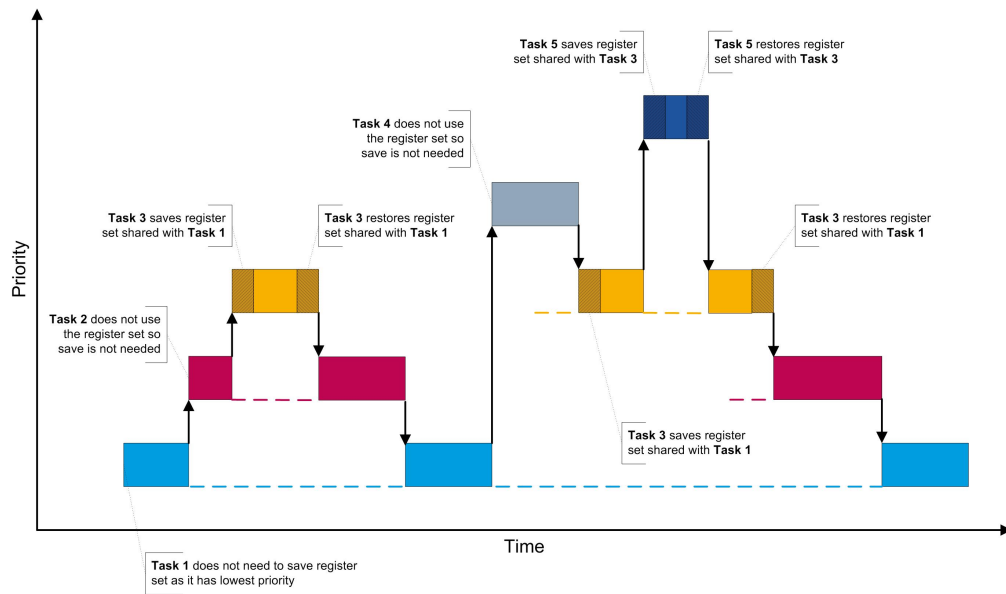


Figure 4.21: Register saving in action

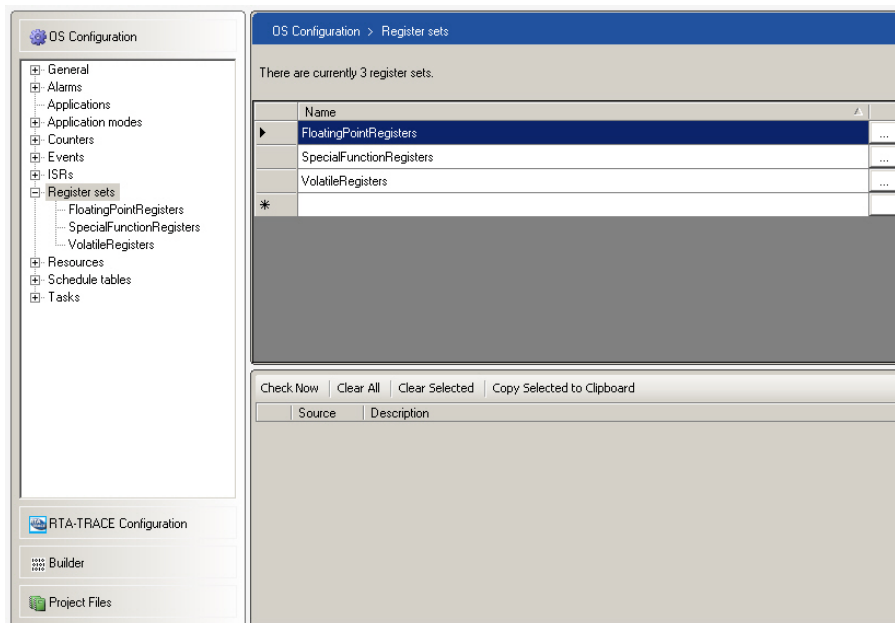


Figure 4.22: Register Set Definition

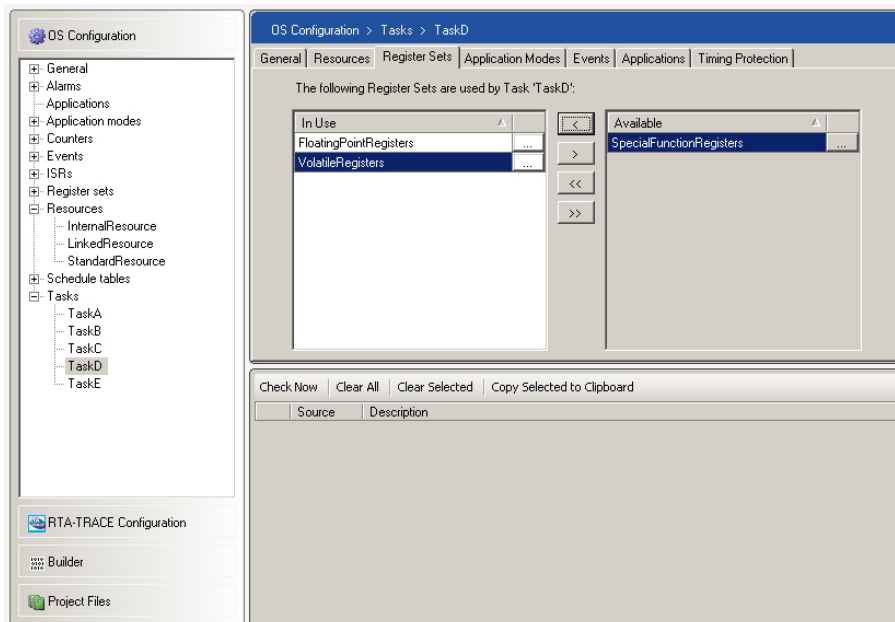


Figure 4.23: Using a register set in a task

Each task that uses a register set needs to declare this at runtime so that **rtaosgen** can calculate the maximum number of sets that need to be saved. Figure 4.23 shows one how this is done for a task.

RTA-OS3.x does not know how or where to save and restore the register sets you declare - it just knows how many saves are necessary and when to save and restore them. For each register set you define, RTA-OS3.x generates a macro called `OS_REGSET_<RegisterSetName>_SIZE` that defines the worst-case number of register set saves required. You should use this in your application code to define an array of size `OS_REGSET_<RegisterSetName>_SIZE` where each element of the array holds the saved register set.

You will also need to provide callback functions for the save and restore operations:

- `Os_Cbk_RegSetSave_<Name>(Os_RegSetDepthType Depth)` is called by RTA-OS3.x whenever it is necessary to save a register set.
- `Os_Cbk_RegSetRestore_<NAME>(Os_RegSetDepthType Depth)` is called by RTA-OS3.x whenever it is necessary to restore a register set.

Both of the callbacks are passed a `Depth` value that indicates the register set to save or restore. Code Example 4.11 shows how the callbacks might should appear in your code.

```

typedef volatile uint32 RegType;

#define VOLATILEREREGISTER (*(RegType*)(0xFECAFECA))

uint32 VolatileRegisterSaveArea[OS_REGSET_VolatileRegister_SIZE];

FUNC(void, OS_APPL_CODE)
    Os_Cbk_RegSetSave_VolatileRegister(Os_RegSetDepthType Depth) {
    VolatileRegisterSaveArea[Depth] = VOLATILEREREGISTER;
}

FUNC(void, OS_APPL_CODE)
    Os_Cbk_RegSetRestore_VolatileRegister(Os_RegSetDepthType
    Depth) {
    VOLATILEREREGISTER = VolatileRegisterSaveArea[Depth];
}

```

Code Example 4.11: Register Set Save And Restore

4.15 Summary

- A task is a concurrent activity.
- There are two classes of tasks: basic and extended.
- Tasks can share priorities, though it is recommended that you do not do this.
- Tasks are scheduled according to priority. When a higher priority task is made ready to run it will preempt lower priority tasks but it will not preempt any task that has been configured as non-preemptive.
- Tasks exist in states: ready, running, suspended or waiting (however, only extended tasks can enter the waiting state).
- If a task terminates, it must call `TerminateTask()` or `ChainTask(TaskID)` to do so.
- Systems where all tasks that terminate do so in their entry functions can use the “fast termination” optimization to minimize stack usage and context switching time.
- Tasks can only be activated when they are in the suspended state unless you specify multiple activations.
- The `PreTask` and `PostTask` Hooks allow you to execute code before your task starts and after it ends. This can be used to profile your application at run-time.

5 Interrupts

Interrupts provide the interface between your application and the things that happen in the real-world. You could, for example, use an interrupt to capture a button being pressed, to mark the passing of time or to capture some other stimulus.

When an interrupt occurs, the processor usually looks at a predefined location in memory called a vector. A vector usually contains the address of the associated interrupt handler. Your processor documentation and the *Target/Compiler Port Guide* for your target will give you further information on this. The block of memory that contains all the vectors in your application is known as the vector table.

5.1 Single-Level and Multi-Level Platforms

Target processors are categorized according to the number of interrupt priority levels that are supported. You should make sure that you fully understand the interrupt mechanism on your target hardware.

There are two different types of target:

Single-level. On single-level platforms there is a single interrupt priority. If an interrupt is being handled, all other pending interrupts must wait until current processing has finished.

Multi-level. On multi-level platforms there are multiple interrupt levels. If an interrupt is being handled, it can be preempted by any interrupt of higher priority. This is sometimes called a *nested* interrupt model.

5.2 Interrupt Service Routines

AUTOSAR operating systems capture interrupts using Interrupt Service Routines (ISRs). ISRs are similar to tasks; however, ISRs differ because:

- They cannot be activated by RTA-OS3.x API calls.
- They cannot make `TerminateTask()` and `ChainTask()` API calls.
- They start executing from their entry point at the associated interrupt priority level.
- Only a subset of the RTA-OS3.x API calls can be made.

The *Reference Guide* tells you the permitted calling context for every API call. You can refer to this to see whether or not you can use an API call in an ISR.

5.3 Category 1 and Category 2 Interrupts

AUTOSAR operating systems classify interrupts into two categories called Category 1 and Category 2. The category indicates whether or not the OS is involved with handling the interrupt.

5.3.1 Category 1 Interrupts

Category 1 interrupts do not interact with RTA-OS3.x. They should always be the highest priority interrupts in your application. It is up to you to configure the hardware correctly, to write the handler and to return from the interrupt.

You can find out more about Category 1 interrupt handlers in Section [5.6.1](#).

The handler executes at or above the priority level of RTA-OS3.x. However, you can make RTA-OS3.x API calls for enabling/disabling and resuming/suspending interrupts.

5.3.2 Category 2 Interrupts

With Category 2 interrupts, the interrupt vector points to internal RTA-OS3.x code. When the interrupt is raised, RTA-OS3.x executes the internal code and then calls the handler that you have supplied.

The handler is provided as an ISR bound to the interrupt (which you can think of as a very high priority task). Execution starts at the specified entry point of the ISR and continues until the entry function returns. When the entry function returns, RTA-OS3.x executes another small section of internal code and then returns from the interrupt.

Figure [5.1](#) shows the state diagram for a Category 2 interrupt handler.

Figure [5.2](#) shows how the internal RTA-OS3.x code wrappers can be visualized.

5.4 Interrupt Priorities

Interrupts execute at an *Interrupt Priority Level* (IPL). RTA-OS3.x standardizes IPLs across all target microcontrollers, with IPL 0 indicating user level, where all tasks execute, and an IPL of 1 or more indicating interrupt level. It is important that you do not confuse IPLs with task priorities. An IPL of 1 is higher than the highest task priority used in your application.

The IPL is a processor-independent description of the interrupt priority on your target hardware. The *Target/Compiler Port Guide* for your port will tell you more about how IPLs are mapped onto target hardware interrupt priorities.

On a single-level platform there are two IPLs, 0 and 1. IPL 0 means that the target is not interrupted and tasks run in priority order. IPL 1 means that

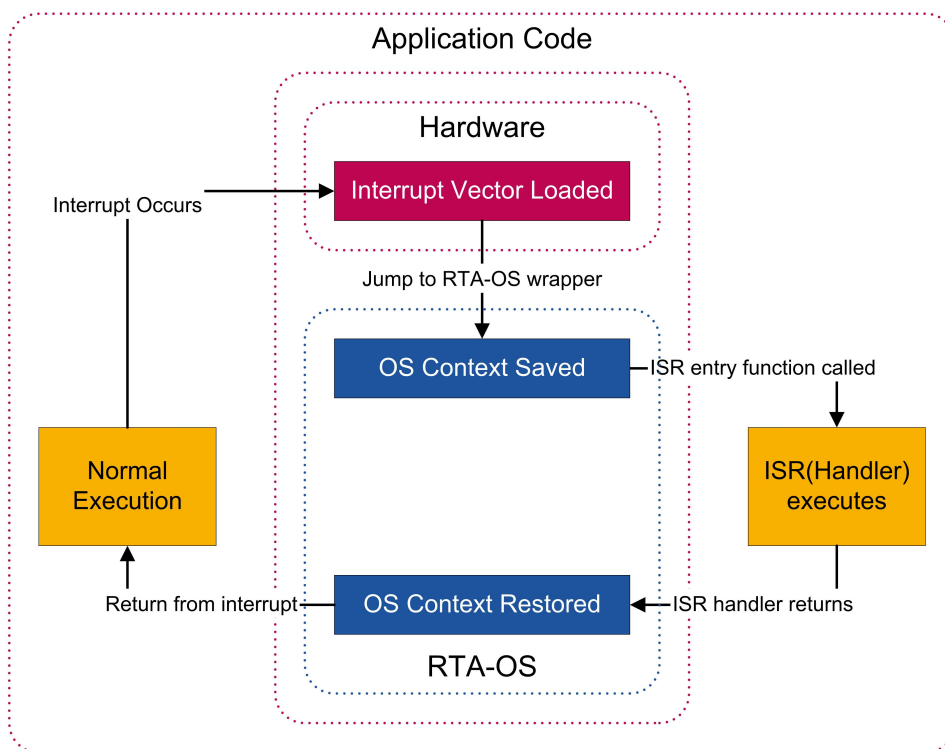


Figure 5.1: Category 2 Interrupt Handling State Diagram

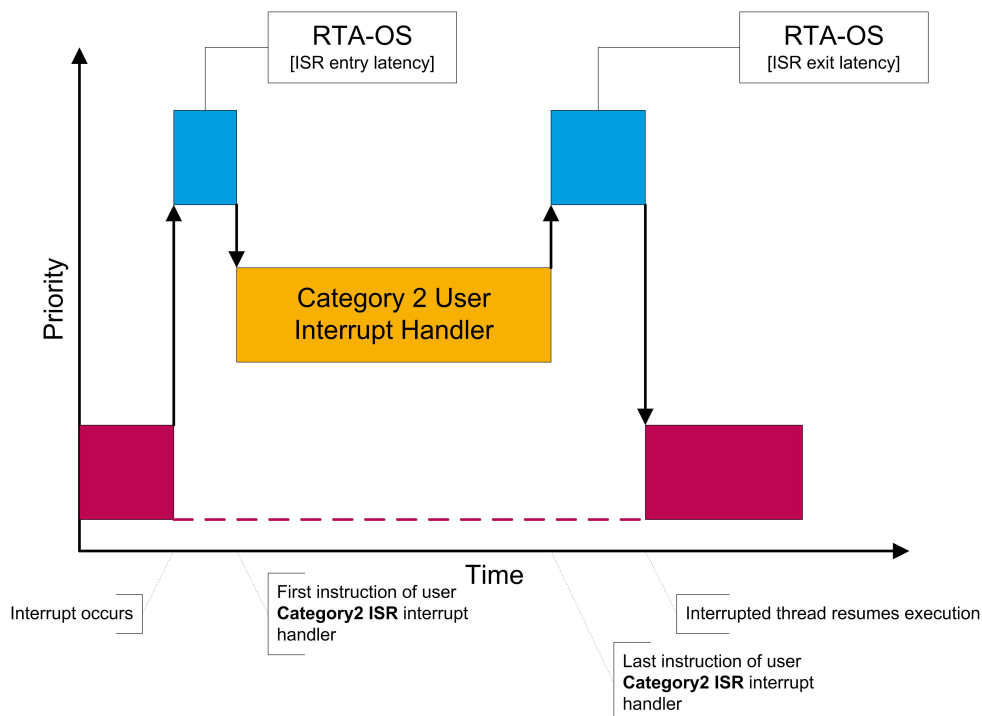


Figure 5.2: Visualizing RTA-OS3.x Category 2 Wrappers

the target is servicing an interrupt. As there is only one non-zero IPL, all interrupts, both Category1 and Category 2, run at the same priority. This means that all interrupts are serialized.

On multi-level platforms, higher priority interrupts can preempt lower priority interrupts and, therefore, the ISR handlers can be nested. So, for example, a higher priority ISR can interrupt the execution of a low priority ISR. However, an ISR can never be preempted by a task.

A Category 1 ISR must never be interrupted by a Category 2 ISR. This is because it is possible for a Category 2 ISR to activate a task and the OS therefore needs to check for a context switch when leaving the ISR - this is what the OS is doing in the second part of the 'wrapper' function shown in Figure 5.2. As ISRs can nest on a multi-level platform, this check must happen as each interrupt exit. Now, if a Category 1 ISR could be preempted by a Category 2 ISR, on exit from the Category 1 ISR no checking for a context switch would occur and the originally preempted task would resume instead of the activated higher priority task. This is priority inversion and can cause unknown side-effects in your system.

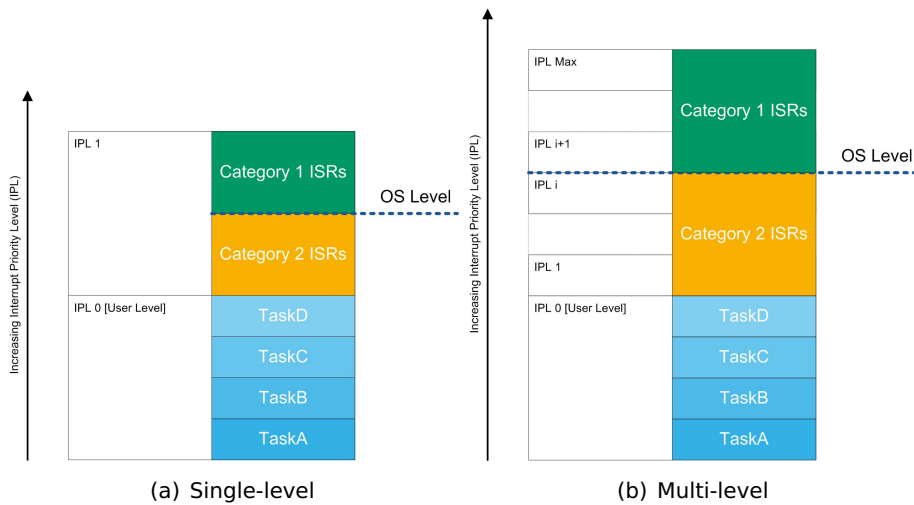


Figure 5.3: Interrupt Priority Hierarchies

This issue means that all Category 2 ISRs must have an IPL that is no higher than the lowest priority Category 1 ISR. RTA-OS3.x automatically checks this at build time and will generate an error if this is the case.

The interrupt priority hierarchies for single and multi-level platforms are shown in Figure 5.3.

5.4.1 User Level

User level is the lowest interrupt priority level that allows all interrupts to be handled. All tasks start executing at user level from their entry point.

A task will sometimes need to run above user level, for example it may need to access data shared with an ISR. While the data is being accessed it must prevent the interrupt being serviced. The simplest way to do this is for the task to disable interrupts while the data is being accessed. This is discussed in Section 5.7. An alternative mechanism is to use AUTOSAR OS's resource mechanism. This is discussed in Chapter 6.

An ISR may preempt a task even when the task is running with interrupt priority level above user level. It can only do this, however, if the ISR has a higher interrupt priority level than the current level.

5.4.2 OS Level

The priority of the highest priority Category 2 ISR defines OS level. If execution occurs at OS level, or higher, then no other Category 2 interrupts can occur.

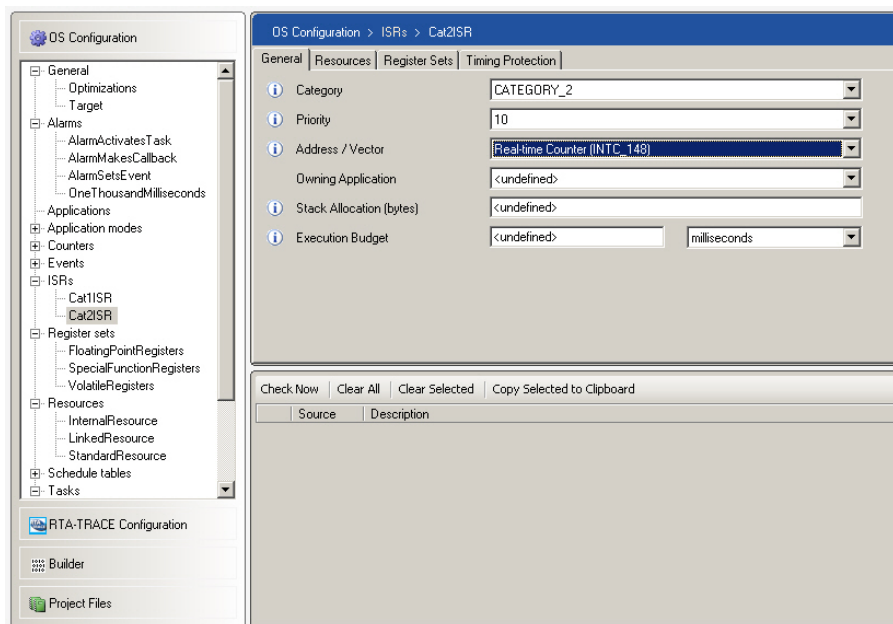


Figure 5.4: Configuring an Interrupt using in [rtaoscfg](#)

RTA-OS3.x uses OS level to guard against concurrent access to internal OS data structures. Any RTA-OS3.x API that manipulates the internal state of the OS will perform some (if not all) of its execution time at OS level. OS hooks (for example the Error Hook, PreTask and PostTaskHook and OS callbacks also run at OS level. If a task executes at OS level, then no RTA-OS3.x operations will take place (except for calls made by the task).

5.5 Interrupt Configuration

In RTA-OS3.x, interrupts are configured statically using [rtaoscfg](#). Figure 5.4 shows how an interrupt has been constructed.

At the simplest level, an interrupt has the following attributes:

An interrupt name. The name is used to refer to C code that you will write to implement the handler functionality (you will learn how to do this in Section 5.6).

An interrupt category. This is either Category 1 if the handler does not need to execute RTA-OS3.x API calls and Category 2 otherwise.

An interrupt priority. The priority is used by the scheduler to determine when the interrupt runs (in a similar way to a task priority being used for tasks). Priority is a microcontroller specific parameter so an RTA-OS3.x target must be selected before you can set a priority. Note that some targets only support a single interrupt priority.

An interrupt vector. RTA-OS3.x uses the specified vector to generate the vector table entry for the interrupt. Like the interrupt priority, interrupt vector configuration is microcontroller specific so a target must be selected before the interrupt vector can be configured.

On microcontrollers where the IPL is user-programmable then it is your responsibility to ensure that the programmed priority level of an interrupting device matches the level you have configured for RTA-OS3.x. Since this configuration must occur before the OS is started, RTA-OS3.x is not able to do this for you since there may be Category 1 ISRs that need to execute. RTA-OS3.x may be able generate appropriate configuration data for you to use. You should consult your Target/Compiler Port Guide for specific instructions.



The RTA-OS3.x GUI allows different targets to be selected (e.g. to allow you to quickly migrate one OS configuration to a new microcontroller). When the target is changed, all target-specific configuration is removed, including the interrupt priority and interrupt vector settings. You will therefore need to provide new target settings as appropriate.



5.5.1 Vector Table Generation

In most cases, RTA-OS3.x can generate the vector table automatically¹. **rtaosgen** will create a vector table with the correct vectors pointing to the internal wrapper code and place this in the generated library.

If you want to write your own vector table then you must make sure that RTA-OS3.x does not generate a vector table. You can prevent a vector table being generated by disabling vector table generation (**Target → Disable Vector Table Generation**) as shown in Figure 5.5.

When you write your own vector table you will need to make sure that all interrupt vectors that are associated with Category 2 ISRs branch to the RTA-OS3.x interrupt wrapper that sets up the context in which the ISR executes.

You must not branch directly to your interrupt handler implementation. Doing so will bypass RTA-OS3.x and any interaction you try to make with the kernel in the context of the handler is likely to result in unrecoverable corruption of the kernel state.



Typically your own vector table will need to branch to labels of the form `0s_Wrapper_VECTOR` where `VECTOR` is the hexadecimal address of the vector. However, the exact details are port-specific. You should consult the *Target/Compiler Port Guide* for your port to obtain specific details of how to provide your own vector table.

¹It may be the case that the compiler for your port generates the vector table. You should consult the *Target/Compiler Port Guide* for your port to obtain specific details.

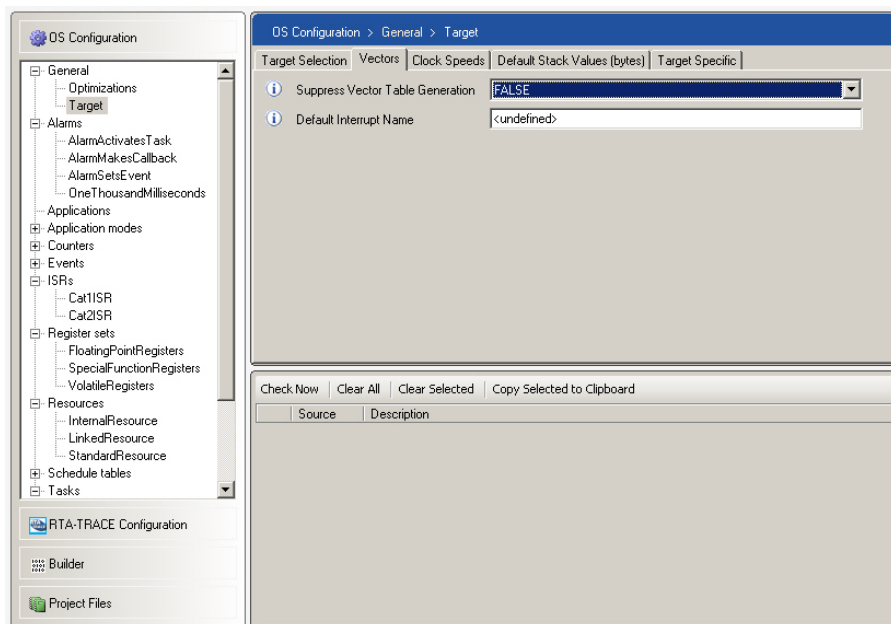


Figure 5.5: Preventing RTA-OS3.x from Automatically Generating a Vector Table

5.6 Implementing Interrupt Handlers

You will now learn about interrupt handlers for Category 1 and Category 2 interrupts.

5.6.1 Category 1 Interrupt Handlers

The format for writing a Category 1 ISR is non-portable. The compiler for the microcontroller typically defines a compiler-specific extension to ANSI C that allows a function to be marked as an interrupt. Some compilers, however, cannot do this. When this happens you will need to write an assembly language handler.

You must make sure that the name of a Category 1 ISR entry function is the same as the name that you specified for the ISR during configuration.

For Category 1 ISRs, there is usually a compiler-specific keyword (sometimes called a *pragma* or a *directive*) that has to be used when defining entry functions. RTA-OS3.x provides a macro called `CAT1_ISR` that expands to the correct directive for your compiler toolchain which you should use to mark your function as a Category 1 ISR.

An entry function for a Category 1 ISR is shown in Code Example 5.1.

```
CAT1_ISR(Interrupt1) {
    /* Handler body. */
}
```

```
    /* Return from interrupt. */  
}
```

Code Example 5.1: Entry Function for a Category 1 ISR

5.6.2 Category 2 Interrupt Handlers

You saw earlier that Category 2 interrupts are handled under the control of RTA-OS3.x. A Category 2 ISR is similar to a task. It has an entry function that is called by RTA-OS3.x when the interrupt handler needs to run. A Category 2 interrupt handler is written using the C syntax in Code Example 5.2.

```
#include <Os.h>  
ISR(isr_identifer){  
    /* Handler body. */  
}
```

Code Example 5.2: Entry Function for a Category 2 ISR

You do not need to provide any C function prototypes for Category 2 ISR entry functions. These are provided in the `Os.h` header file that is generated by **rtaosgen**.



You must not place a 'return from interrupt' command in your Category 2 ISR. Returning from the interrupt is handled by RTA-OS3.x.

5.6.3 Dismissing Interrupts

When the hardware detects an interrupt, it will typically set a *pending bit* which tells the interrupt controller that an interrupt has occurred. The interrupt controller will then branch to the handler through the interrupt vector table.

The handling of the pending bit is target dependent but there are two basic models:

1. the pending bit is cleared automatically after the interrupt is handled (i.e. when the branch to the interrupt handler occurs). When the handler exits it will be automatically re-triggered if an interrupt has become pending while the current interrupt was being handled;
2. the pending bit must be cleared manually by user code in the interrupt handler. The body of the interrupt handler, whether Category 1 or Category 2, will need to include the code to clear the pending bit and signal to the hardware that the interrupt has been handled.

If you need to clear the pending bit, it is good practice to do this immediately on entry to the handler because this minimizes the time between the pending bit being set by a second instance of the interrupt occurring and then subsequently cleared. This helps to prevent issues where the interrupt becomes pending multiple times but this cannot be recognized by the hardware. Code example 5.3 shows how the recommended structure of a Category 2 ISR handler.

```
#include <Os.h>
ISR(Interrupt1) {
    /* Dismiss the interrupt where required */
    /* Rest of the handler */
}
```

Code Example 5.3: Dismissing the interrupt

You will need to consult your hardware reference manual to find out what you need to do on your target hardware.

5.6.4 Writing Efficient Interrupt Handlers

Each interrupt handler you write will block all interrupts of equal or lower priority for the time that it takes your code to execute. When you write an interrupt handler it is good practice to make the handler as short as possible. A long running handler will add additional latency to the servicing of lower priority interrupts.

By minimizing the execution time of your interrupt handlers you can maximize overall system responsiveness.

If you need to execute a long-running piece of code in response to the interrupt occurring, then you can put that code into a task and then activate the task from a Category 2 ISR. Code Example 5.4 and Code Example 5.5 show how these techniques differ.

With Category 2 handlers you can move the required functionality to a task, a simply use the interrupt handler to activate the task and then terminate.

```
#include <Os.h>
ISR(InefficientHandler) {
    /* Long handler code. */
}
```

Code Example 5.4: Inefficient interrupt handler

```
#include <Os.h>
ISR(EfficientHandler) {
    ActivateTask(Task1);
}
```

```

TASK(Task1) {
    /* Long handler code. */
    TerminateTask();
}

```

Code Example 5.5: More efficient interrupt handler

5.7 Enabling and Disabling Interrupts

Interrupts will only occur if they are enabled. By default, RTA-OS3.x ensures that all interrupts are enabled when `StartOS()` returns.



AUTOSAR OS uses the term `Disable` to mean masking interrupts and `Enable` to mean un-masking interrupts. The `enable` and `disable` API calls do not therefore enable or disable the interrupt source; they simply prevent the processor from recognizing the interrupt (usually by modifying the processor's interrupt mask).

You will often need to disable interrupts for a short amount of time to prevent interrupts occurring in a critical section of code in either tasks or ISRs. A critical section is a sequence of statements that accesses shared data.

You can enable and disable interrupts using a number of different API calls:

- `DisableAllInterrupts()` and `EnableAllInterrupts()`
Disable and enable all interrupts that can be disabled on the hardware (usually all those interrupts that can be masked).
These calls cannot be nested.
- `SuspendAllInterrupts()` and `ResumeAllInterrupts()`
Suspend and resume all interrupts that can be disabled on the hardware (usually all those interrupts that can be masked).
These calls can be nested.
- `SuspendOSInterrupts()` and `ResumeOSInterrupts()`
Suspend and resume all Category 2 interrupts on the hardware.
These calls can be nested.



You must make sure that there are never more 'Resume' calls than 'Suspend' calls. If there are, it can cause serious errors and the behavior is undefined. Subsequent 'Suspend' calls may not work. This will result in unprotected critical sections.

Code Example 5.6 shows you how the interrupt control API calls are used and nested correctly.

```

#include <Os.h>
TASK(Task1) {
    DisableAllInterrupts();
    /* First critical section */
    /* Nesting not allowed */
    EnableAllInterrupts();
    SuspendOSInterrupts();
    /* Second critical section */
    /* Nesting allowed. */
    SuspendAllInterrupts();
    /* Third critical section */
    /* Nested inside second */
    ResumeAllInterrupts();
    ResumeOSInterrupts();
    TerminateTask();
}

```

Code Example 5.6: Nesting Interrupt Control API Calls

In the case of Category 1 ISRs, you must make sure that no RTA-OS3.x API calls are made (except for other Suspend/Resume calls) for the entire time that the interrupts are disabled.

If a Category 2 ISR raises the interrupt level above OS level by calling `DisableAllInterrupts()` then it may not make any other RTA-OS3.x API calls, except for the `EnableAllInterrupts()` call to restore the interrupt priority. When executing an ISR, you are not allowed to lower the interrupt priority level below the initial level.

5.8 Saving Register Sets

Recall from Section 4.14 that RTA-OS3.x provides a mechanism for saving register sets across context switches and that `rtaosgen` can optimize the amount of saving that is required to improve runtime performance.

The same mechanism can also be used by Category 2 ISRs by simply selecting which ISRs use the configured register set as shown in Figure 5.6.

5.9 The Default Interrupt

If you are using RTA-OS3.x to generate a vector table, then you may want to fill unused vector locations with a default interrupt.

Figure 5.7 shows how the default interrupt is defined.



The default interrupt is not supported by all ports.

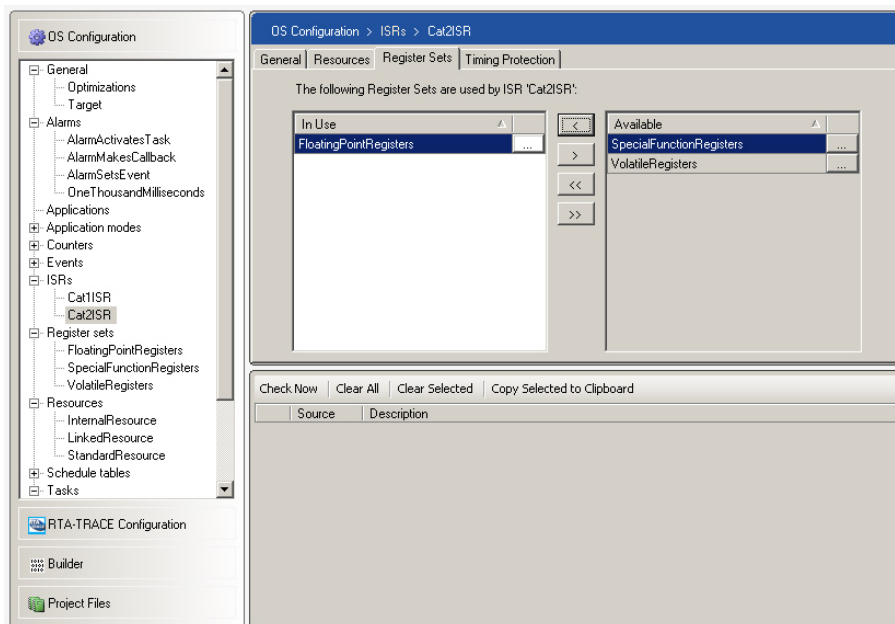


Figure 5.6: Using a register set in a Category 2 ISR

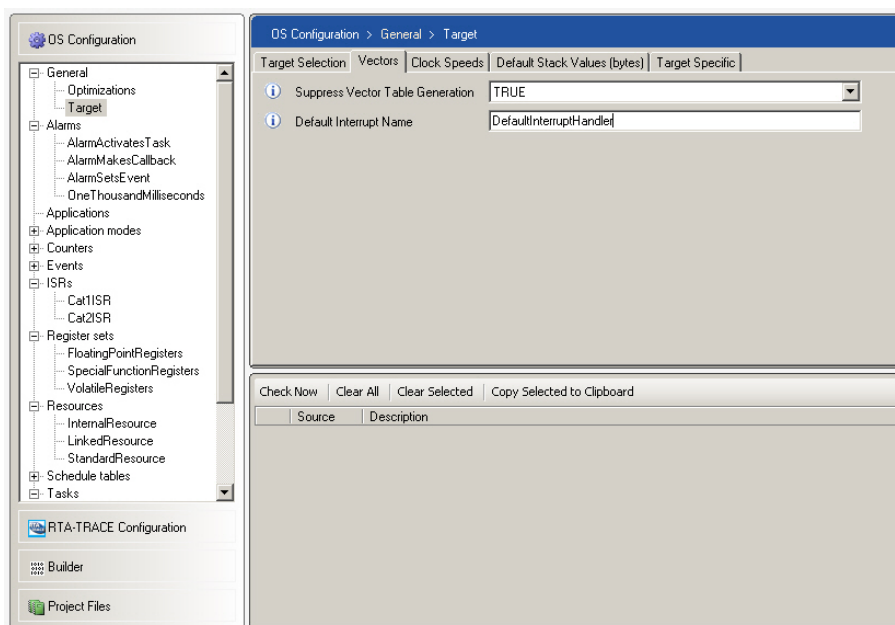


Figure 5.7: Placing a Default Interrupt in the Vector Table

The name allocated to the default interrupt at configuration time is the name that must be used in your application code when you write the handler. Code Example 5.7 shows a default handler that would work with the configuration shown in Figure 5.7.

The default interrupt is slightly different to other interrupts. It is used to fill every location in the vector table for which you have not defined an interrupt. This feature has been provided as a debugging aid and as a means of providing a 'fail-stop' in the event of erroneous generation of interrupts in production systems. If you actually want to attach interrupt handlers to vectors to do useful work, you should explicitly create them as ISRs.

There are limitations on the use of the default interrupt handler. It cannot make any OS calls, and system behavior is undefined if it ever returns.



Do not make any RTA-OS3.x API calls from the default interrupt and you must not return from the handler.

The default interrupt is implemented like an OSEK Category 1 interrupt and must therefore be marked as an interrupt with the CAT1_ISR macro. The last statement in your default interrupt handler should be an infinite loop. Code Example 5.7 shows how this can be done.

```
CAT1_ISR(DefaultInterruptHandler) {
    /* invoke target-specific code to lock interrupts */
    asm('di'); /* or whatever on your platform */
    for (;;) {
        /* Loop forever */
    }
    /* Do NOT return from default handler. */
}
```

Code Example 5.7: The Default Interrupt Handler

5.10 Summary

- RTA-OS3.x supports two categories of interrupts: Category 1 and Category 2.
- Category 1 ISRs are normal embedded system interrupts that bypass RTA-OS3.x. As a result they cannot interact with the OS and are forbidden from making (most) RTA-OS3.x API calls. They should be marked using the CAT1_ISR macro.
- Category 2 ISRs are OS managed interrupts that run in a wrapper provided by RTA-OS3.x. These interrupts can make RTA-OS3.x API calls. They must be marked using the ISR macro.

- All interrupts run at an Interrupt Priority Level (IPL) which is always strictly higher than the highest task priority.
- IPLs standardize the interrupt priority model across all hardware devices - higher IPLs mean higher priority.
- RTA-OS3,0 can generate an interrupt vector table or you can choose to write your own. When generating a vector table, RTA-OS3.x can plug unused locations with a user-configured default interrupt.

6 Resources

Access to hardware or data that needs to be shared between tasks and ISRs can be unreliable and unsafe. This is because task or ISR preemption can occur while a lower priority task or ISR is part way through updating the shared data. This situation is known as a race condition and is extremely difficult to test for.

A sequence of statements that accesses shared data is known as a critical section. To provide safe access to code and data referenced in the critical section you need to enforce mutual exclusion. In other words, you must make sure that no other task or Category 2 ISR in the system is able to preempt the executing task during the critical section.

In Chapter 4 you saw that you can declare tasks to be non-preemptive and that this prevents problems with mutual exclusion. However, this method is 'brute-force' because it prevents preemption problems by preventing preemption - rather like preventing car accidents by getting rid of cars!

The OS provide alternative mutual exclusion mechanisms based on resources. A resource is just a binary semaphore. When a task or Category 2 ISR gets a resource, no other task or ISR can get the resource. This prevents any other task or ISR entering the same critical section at the same time. When the critical section is finished, the task or ISR releases the resource and the critical section can be entered by another task/ISR.

When a high priority task is being prevented from executing by a lower priority task this is called priority inversion because the higher priority task takes longer to complete its execution than the lower priority task. The lower priority task appears to be running in precedence to the higher priority task, contrary to what would be expected from their actual priority assignment. The high priority task is said to be blocked by the low priority task.

Binary semaphores in traditional operating systems often get a bad name because priority inversion can introduce unbounded blocking in the system. For example, if the low priority task is preventing the high priority task from executing but is itself preempted by a medium priority task that does not need access to the shared resource then the high priority task will be blocked by execution of the medium priority task as well. As the low priority task might be preempted multiple times while it holds the shared resource, the blocking suffered by the high priority task can be unbounded, posing a significant problem if you need to determine the longest time it takes a task to respond¹. In extreme cases, tasks can reach a state called 'deadlock' where each task is waiting to enter a critical section that is being used by some other task.

¹Because the response time of the task depends on a factor that you cannot calculate.

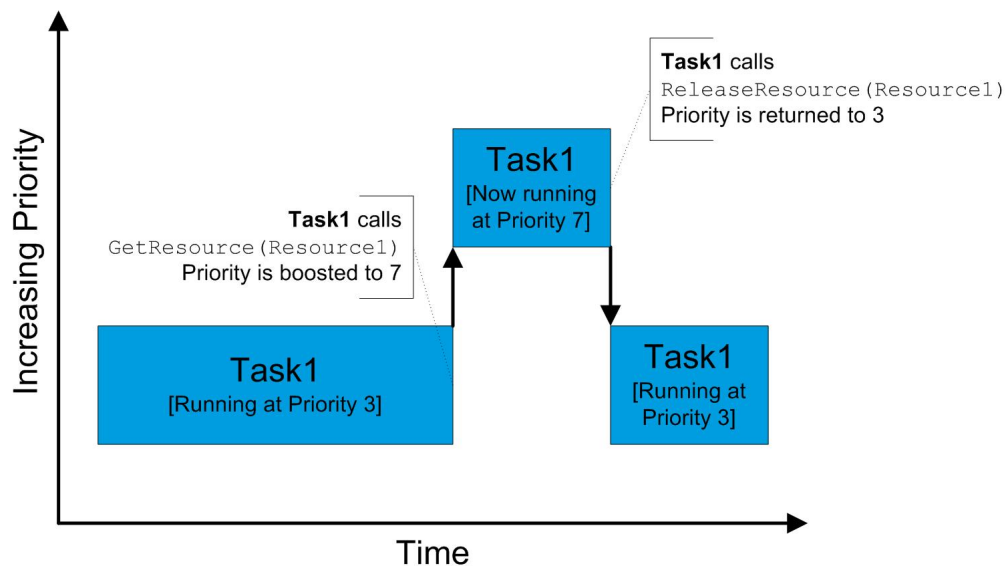


Figure 6.1: Raising to ceiling priority

In AUTOSAR OS, the problems typically associated with priority inversion and deadlock are avoided because resources are locked according to a locking protocol. This locking protocol is called priority ceiling protocol, in particular a version called immediate inheritance priority ceiling protocol (or alternatively stack resource protocol).

Priority ceiling protocol uses the concept of a *ceiling priority*. Each resource in the system is allocated a ceiling priority that is equal to the highest priority of any task or ISR that needs access to the resource. When a task or ISR gets a resource, the running priority of the task/ISR is increased to the ceiling priority of the resource (if and only if this is higher than the task/ISR's current running priority). When the resource is released, the priority of the task or reverts to the priority immediately prior to the task or ISR making the call. This is shown in Figure 6.1.

Immediate inheritance priority ceiling protocol provides two major benefits:

1. Priority inversion is minimized.

Each time a high priority task or ISR becomes ready, its execution can only be delayed at most once by a single lower priority task or ISR that already holds the a resource. This means there is no cumulative blocking so it is possible to place an upper bound on the blocking that a task suffers - the maximum blocking time is the longest time that a lower priority task/ISR holds the shared resource. Furthermore, this blocking always occurs at the start of execution. A consequence of this is that a

resource is always free at the point it needs to be locked. There is no need in AUTOSAR OS to wait for a resource to be released.

2. It is guaranteed to be deadlock free.

A task or ISR must be executing in order to make the lock. This can be proved by contradiction. Assume that a task (or ISR) tries to get a resource. If another task or ISR already had the resource then, because that task or ISR must be running at the ceiling priority, the task making the request not be executing (it would not be the highest priority task or ISR in the system) and, therefore, could not be attempting to lock the resource.

6.1 Resource Configuration

At the most basic level, resources only need to be named and assigned a type. There are three types of resource in AUTOSAR OS:

1. Standard resources are normal OS semaphores. Configuring a standard resource creates a resource with the specified name.
2. Linked resources allow you to alias a standard (or another linked) resource so that nested locking of the same resource is possible. These are discussed in more detail in Section 6.4.
3. Internal resources are resources that are locked automatically on entry to a task and released automatically on termination. These are discussed in more detail in Section 6.5.

Figure 6.2 shows how a standard resource is configured in the `rtaoscfg`.

RTA-OS3.x needs to know which tasks and ISRs use which resources. It can then calculate the ceiling priorities used by the priority ceiling protocol.

Additional resource usage information for each task or ISR can be configured during task or ISR configuration.

Figure 6.2 shows that a resource called Resource1 has been declared. When you refer to this resource in your program you must use the same name.

6.2 Resources on Interrupt Level

Resources that are shared between tasks and interrupts are optional in OSEK. This optional feature is supported by RTA-OS3.x.

RTA-OS3.x will automatically identify the resources that are combined resources, so you don't need to do any special configuration.

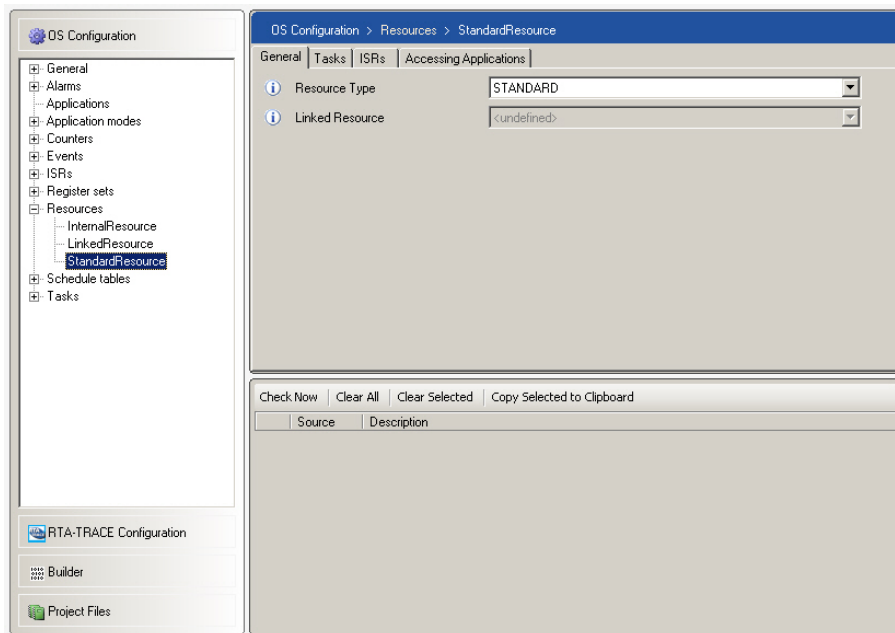


Figure 6.2: Configuring Resources using the `rtaoscfg`

When a task gets a resource shared with an ISR, RTA-OS3.x will mask all interrupts with interrupt priority less than or equal to the highest priority interrupt that shares the resource.

This is simply an extension of priority ceiling protocol.

Sharing resources between tasks and ISRs means provides greater control over interrupt masking than the Enable/Disable and Suspend/Resume API calls because they make it possible to mask a subset of interrupts up to a particular priority level. Resources on interrupt level are therefore especially useful when using an RTA-OS3.x port that supports nested interrupts.

6.3 Using Resources

You can get a resource using the `GetResource()` API call. You can then release a resource using the `ReleaseResource()` call. A task or ISR must not terminate until it has released all resources that it locked.

A task or ISR can only use the resources that you specify during RTA-OS3.x configuration. Code Example 6.1 shows you how resources are used in `Task1`.

```
#include <Os.h>
TASK(Task1) {
    ...
    GetResource(Resource1);
    /* Critical section. */
    ReleaseResource(Resource1);
}
```

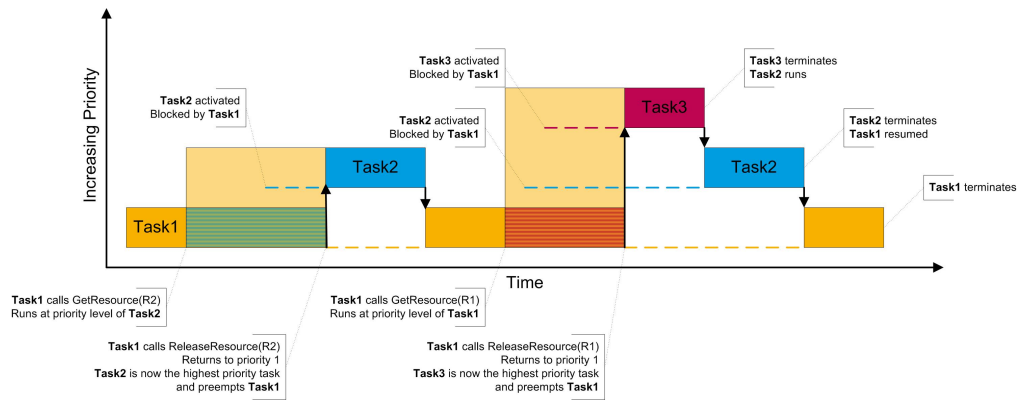


Figure 6.3: Execution of tasks with resource locks

```

...
    TerminateTask();
}

```

Code Example 6.1: Using Resources

Calls to `GetResource()` and `ReleaseResource()` must be matched. You cannot get a resource that is already locked. You cannot release a resource you have not already locked.

When a `GetResource()` is made, it boosts the priority of the calling task or ISR to the ceiling priority of the resource. The resource's ceiling priority is the highest priority of any task or ISR that shares the resource and is automatically calculated by RTA-OS3.x. If any task with a priority less than the ceiling priority is made ready to run, then it is prevented from executing (it is **blocked**) until the priority of the running task returns to normal.

Figure 6.3 shows this effect with the following configuration:

Task	Priority	Locks Resource R1	Locks ResourceR2
3	High	✓	✗
2	Medium	✗	✓
1	Low	✓	✓

The first activation of Task 2 is blocked because Task 1 has locked R1. The second activation of Task 2 is also blocked, but this time because Task 1 has locked R1. The first activation of Task 3 is similarly blocked because of Task 1 holding R1. When Task 1 releases R1, the OS runs the highest priority ready task which is Task 3. On termination of Task 3, Task 2 executes and finally, when Task 2 terminates and Task 1 resumes.

6.3.1 Nesting Resource Calls

You can get more than one resource concurrently, but the API calls must be strictly nested. Let's look at two examples; one showing incorrectly nested calls and the other showing the API calls nested correctly. Code Example 6.2 shows Resource1 and Resource2 being released in the wrong order.

```
GetResource(Resource1);
    GetResource(Resource2);
ReleaseResource(Resource1); /* Illegal! */
    /* You must release Resource2 before Resource1 */
    ReleaseResource(Resource2);
```

Code Example 6.2: Illegal Nesting of Resource Calls

A correctly nested example is shown in Code Example 6.3. All of the resources are held and then released in the correct order.

```
GetResource(Resource1);
    GetResource(Resource2);
        GetResource(Resource3);
            ReleaseResource(Resource3);
        ReleaseResource(Resource2);
    ReleaseResource(Resource1);
```

Code Example 6.3: Correctly Nested Resource Calls

6.4 Linked Resources

In AUTOSAR OS, GetResource() API calls for the same resource cannot be nested. However sometimes, there are cases where you may need make nested resource locks.

Your application may, for instance, use a function shared amongst a number of tasks. What happens if the shared function needs to get a resource used by one of the tasks, but not by the others? Have a look at Code Example 6.4.

```
#include <Os.h>
void SomeFunction(void) {
    GetResource(Resource1);    /* !!! Not allowed if caller is
        Task1 !!! */
    ...
    ReleaseResource(Resource1); /* !!! Not allowed if caller is
        Task1 !!! */
}

TASK(Task1) {
    GetResource(Resource1);
        /* Critical section. */
        SomeFunction();
    ReleaseResource(Resource1);
```

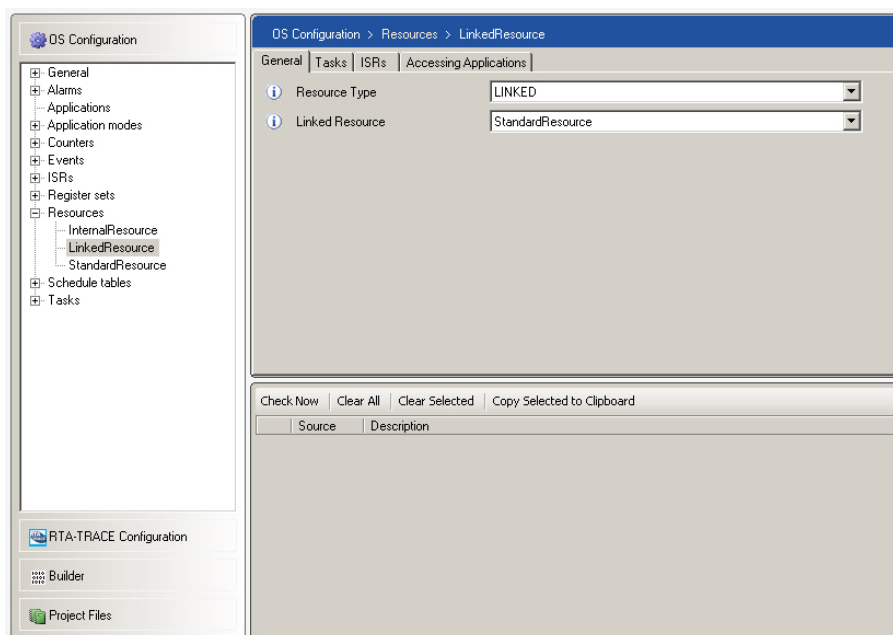


Figure 6.4: Configuring a Linked Resource

```

}

TASK(Task2) {
    SomeFunction();
}

```

Code Example 6.4: Illegal locking of previously locked resource

In these cases, the nesting of a (potentially) held resource must use linked resources. A linked resource is an alias for an existing resource and protects the same, shared, object.

Figure 6.4 shows how linked resources are declared using `rtaoscfg`.

With the linked resource, Code Example 6.4 would be re-written as shown in Code Example 6.5.

```

#include <Os.h>
void SomeFunction(void) {
    GetResource(LinkedToResource1); /* Okay */
    ...
    ReleaseResource(LinkedToResource1); /* Okay */
}

TASK(Task1) {
    GetResource(Resource1);
    /* Critical section. */
}

```

```

        SomeFunction();
    ReleaseResource(Resource1);
}

TASK(Task2) {
    SomeFunction();
}

```

Code Example 6.5: Using Linked Resources

Linked resources are held and released using the same API calls for standard resources (these are explained in Section 6.3). You can also create linked resources to existing linked resources.

6.5 Internal Resources

If a set of tasks share data very closely, then it may be too expensive, in terms of runtime cost, to use standard resources to guard each access to each item of data. You may not even be able to identify all the places where resources need to be held.

You can prevent concurrent access to shared data by using internal resources. Internal resources are resources that are allocated for the lifecycle of a task.

Internal resources are configured offline using **rtaoscfg**. Unlike normal resources, however, you cannot get and release them. Conceptually, RTA-OS3.x locks the internal resource immediately before starting the task and releases the resource immediately after the task terminates.



In AUTOSAR OS R3.x internal resources are only available to tasks. However, there is no reason why internal resources cannot be shared by Category 1 and 2 ISRs as well. RTA-OS3.x provides an extension to AUTOSAR OS R3.x that allows ISRs to use internal resources. When the a task locks an internal resource that is shared with an ISR, then the task executes at the IPL of the interrupt and all interrupts of equal or lower priority will be blocked for the duration of the task.

The implementation of internal resources in RTA-OS3.x does not incur a runtime cost when the task enters the running state because **rtaosgen** calculates the priority at which the task will run offline and simply dispatches the task at this priority. The set of tasks that share an internal resource is statically defined at configuration time using **rtaoscfg**.

Figure 6.5 shows the declaration of an internal resource (called `IntResource1`) which is shared between tasks `t1` and `t3`.

If a task uses an internal resource, RTA-OS3.x will automatically get the internal resource before calling the task's entry function. The resource will then

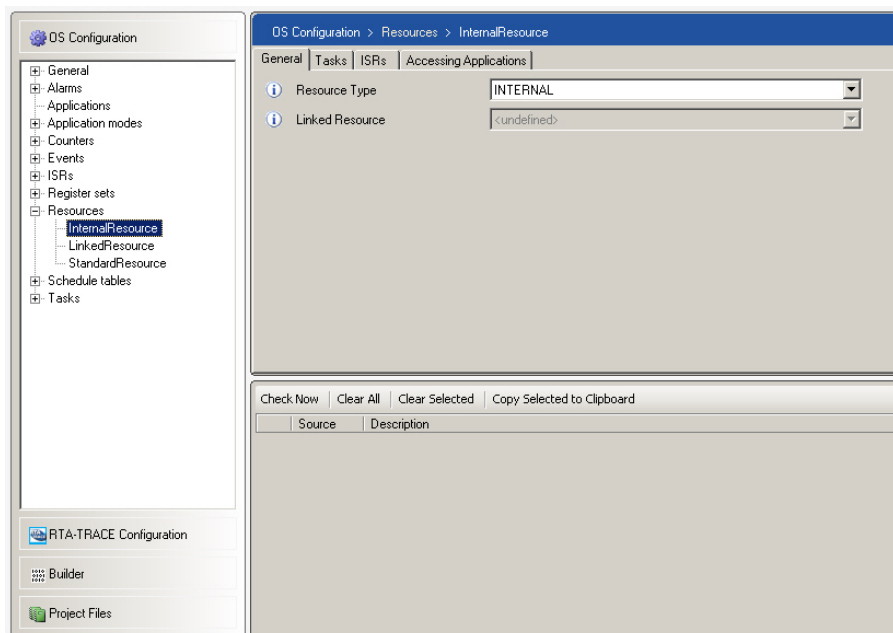


Figure 6.5: Declaring an Internal Resource using `rtaoscfg`

be automatically released after the task terminates, makes a `Schedule()` or a `WaitEvent()` call.

During task execution, all other tasks sharing the internal resource will be prevented from running until the internal resource is released. Figure 6.6 shows the execution of three tasks that share the same internal resource.

It is important to note that the OS makes a scheduling decision based on the normal (base) priority of the ready tasks when a task that holds an internal resource terminates. If a task is running and multiple tasks that share the same internal resource have become active then, on termination of the running task, the highest priority ready tasks is selected to run and then is dispatched at the ceiling priority of the internal resource.

Any tasks with a priority lower than the ceiling priority of the internal resource, including those that do not share the internal resource, will be blocked if a task sharing the internal resource is executing. You can see an illustration of this in Figure 6.6 where Task1 with Priority 1 shares an internal resource with a Task that has Priority 3. If Task1 starts running before Task2 or Task3 start, then both of these tasks will be delayed (blocked) until Task1 completes.

However, preemption is still possible by all higher priority tasks that do not share the internal resource. Figure 6.7 shows that initially Task 1 is running at priority 3 because it shares an internal resource with a task of priority 3. While

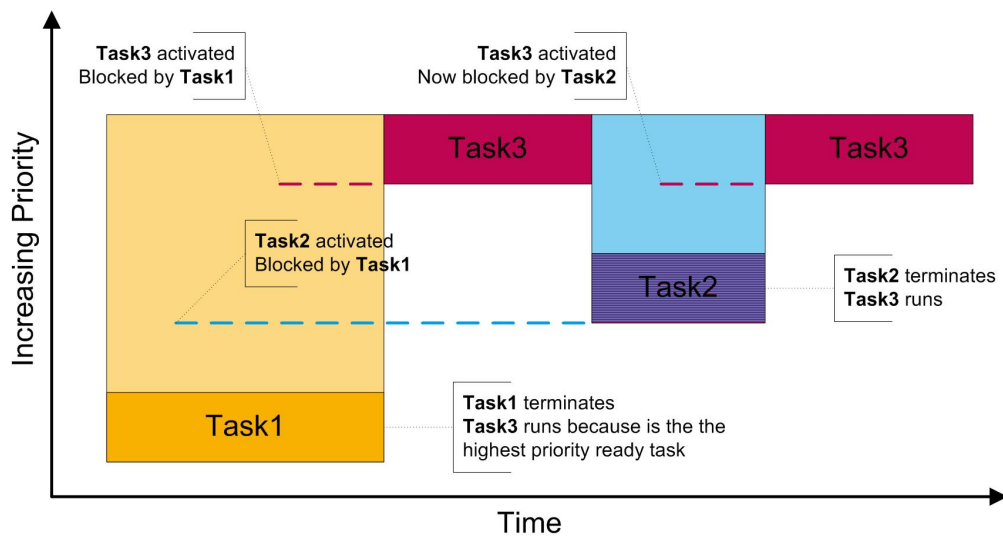


Figure 6.6: Execution with internal resources

Task 1 is running, Task 2 becomes ready to run. Task 2 is lower priority than the active priority of Task1 so it cannot preempt. When Task4 is activated, it can preempt Task1 because its priority is 4 i.e. it is higher priority than the active priority of Task 1. Task 2 can only run when Task 1 terminates.

From this behavior it should be clear that a task which locks an internal resource will prevent *any* task with a higher priority than itself but lower priority than the ceiling priority of the internal resource from running *for the entire duration of the task*. When a lower priority task prevents a higher priority task from executing this is called *blocking*.

Tasks that share an internal resource run non-preemptively with respect to each other. Once a task in the set sharing the internal resource gets access to the CPU, it will run without being preempted by any other task in the set. The consequence of this is that it may take longer for higher priority tasks to get access to the CPU than would be the case in a fully preemptive system.

6.6 Using Resources to Minimize Stack Usage

The primary role of resources in an application is to provide mutual exclusion over critical sections. However, the single-stack model of RTA-OS3.x means that resources have a useful secondary role - minimizing stack usage. Recall that tasks which share resources do not preempt each other. In the single-stack model used by RTA-OS3.x this means that their stack usage is effectively overlaid.

It is possible to exploit this feature to trade off time in the system against stack usage. The following sections describe how simple modifications to an

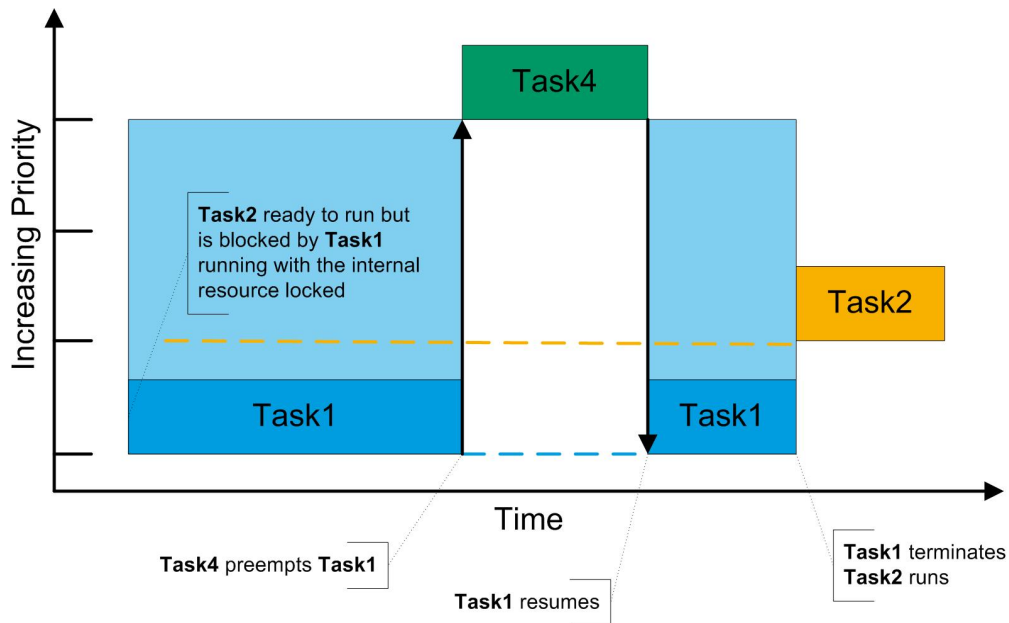


Figure 6.7: Internal resources blocking tasks that do not share the resource

application can reduce stack usage. All of these modifications will introduce additional blocking factors into the system.

The impact of these blocking factors depends on the system. Recall that the priority ceiling protocol ensures that a task or ISR is blocked at most once during execution. The worst-case blocking time is the maximum time that any lower priority task or ISR can hold the same resource.

This means that if the additional blocking factors are less than or equal to the current worst-case blocking suffered by a task/ISR, then there will be no impact on response times and the reduced stack usage will be free. If the additional blocking factors are longer than current worst-case blocking then response times will be longer. Providing that response times remain inside the required deadlines for tasks/ISRs, the system will still behave correctly.

6.6.1 Internal Resources

Given a set of tasks that share an internal resource, the worst case stack used by RTA-OS3.x is equal to the maximum stack space required by the task that uses the most stack. In conventional operating systems, the maximum stack space would be equal to the sum of the task's stacks, not their maximum.

If you need to minimize stack space then you can exploit this benefit of RTA-OS3.x's single-stack architecture by sharing internal resources between tasks which consume lots of stack. The first stack in Figure 6.8 shows the worst-case stack consumption for 5 preemptive tasks, A, B, C, D and E. By sharing

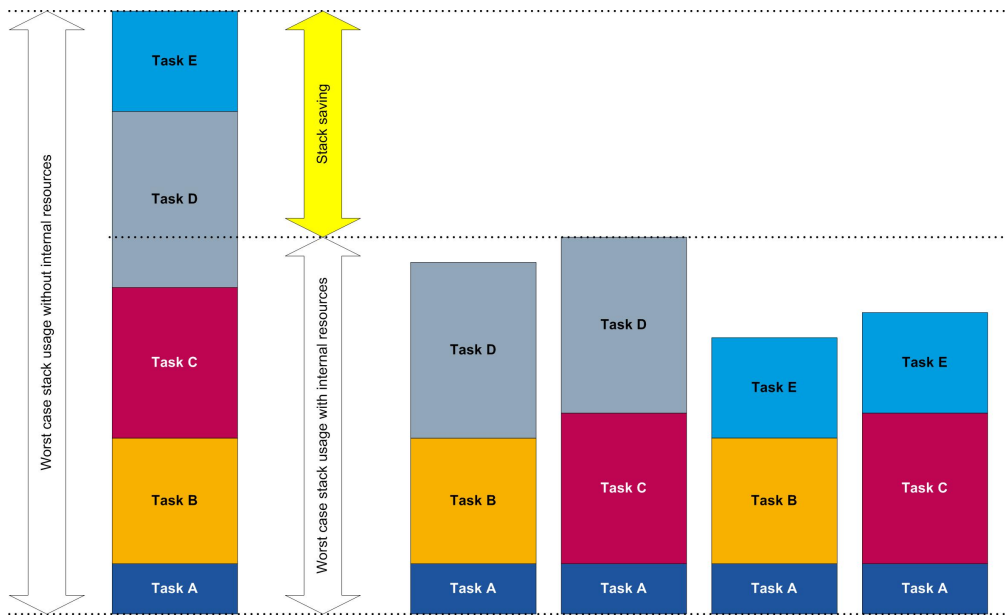


Figure 6.8: Saving Stack Space Using Internal Resources

an internal resource between tasks B and C, and between tasks D and E a significant saving of stack space can be made. The other four stacks in Figure 6.8 show the cases that can now occur - the worst case is A preempted by the worst of B or C preempted by the worst of D and E. You can see from the figure that A preempted by C preempted by D gives the worst case and that this is significantly less stack than when internal resources were not used.

6.6.2 Standard Resources

If a task calls a function that uses a lot of stack then you could consider locking a resource around the function call and sharing the resource with the tasks of higher priority. The tasks do not need to lock the resource in code or call the function - the sharing is simply to force the execution of the task to run at a higher priority. This will prevent higher priority tasks preempting the task while it is using lots of stack and will therefore reduce the total stack requirement.

Disabling interrupts around the function call has a similar effect - effectively overlaying the function call's stack usage with the ISRs that are temporarily masked.

6.7 The Scheduler as a Resource

A task can hold the scheduler if it has a critical section that must be executed without preemption from any other task in the system (recall that the scheduler is used to perform task switching). A predefined resource called

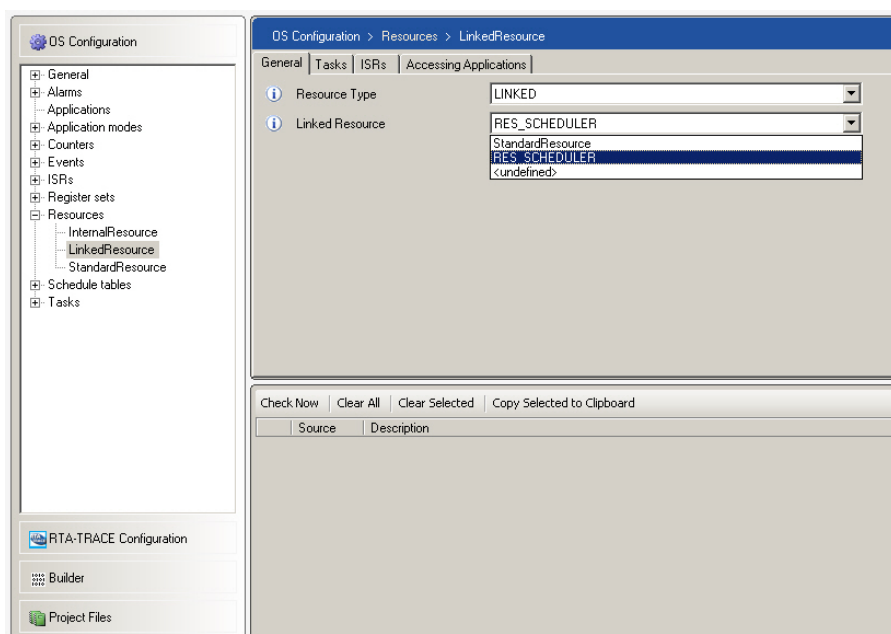


Figure 6.9: Linking to RES_SCHEDULER

RES_SCHEDULER is available to all tasks for this purpose. RES_SCHEDULER is a convenient way for tasks to share data without you needing to declare a resource that is shared between all tasks manually.

When a task gets RES_SCHEDULER, all other tasks will be prevented from preempting until the task has released RES_SCHEDULER. This effectively means that the task becomes non-preemptive for the time that RES_SCHEDULER is held. This is better than making the entire task non-preemptive, particularly when a task only needs to prevent preemption for a short part of its total execution time.

You must specify whether your application uses RES_SCHEDULER or not. This is configured in **General → Optimizations**. If you configure RES_SCHEDULER then RTA-OS3.x will automatically generate a standard resource called RES_SCHEDULER and share it between every task in your configuration. As RES_SCHEDULER behaves like a standard resource, you can create linked resources that link to RES_SCHEDULER as shown in Figure 6.9.

Using RES_SCHEDULER can improve response times of low priority tasks that might otherwise suffer multiple preemptions by other tasks in the application, but at the cost of longer response times for higher priority tasks.

If you have no need to use RES_SCHEDULER in your application then you can save ROM and RAM space by disabling its generation as shown in Figure 6.10.

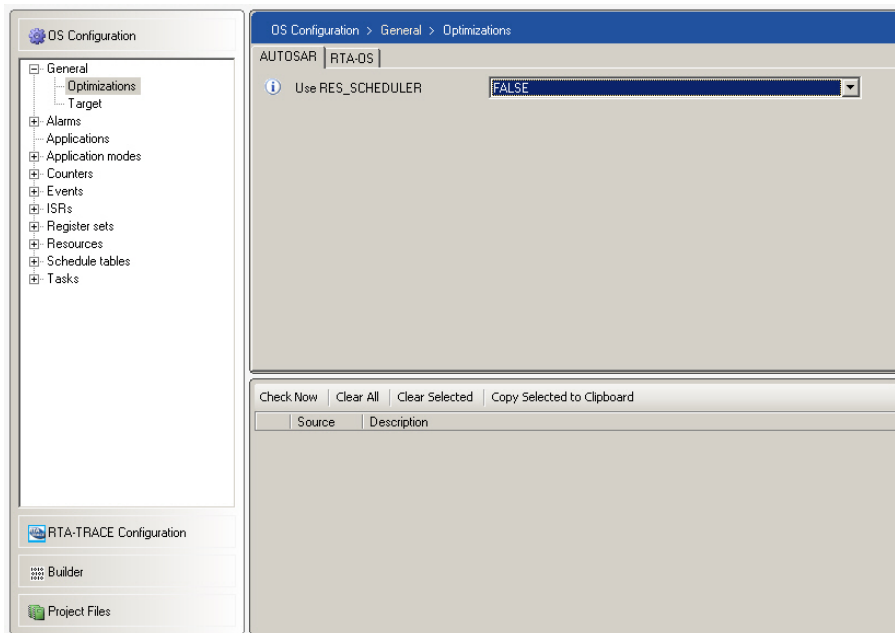


Figure 6.10: Disabling RES_SCHEDULER

6.8 Choosing a Preemption Control Mechanism

If code that does not require locks appears between a pair of `GetResource()` and `ReleaseResource()` calls, the system responsiveness can potentially be reduced.

With this in mind, when you use resources in your application, you should place `GetResource()` calls as closely as possible around the section of code you are protecting with the resource.

However, there is an exception to this rule. This exception occurs when you have a short running task or ISR that makes many `GetResource()` and `ReleaseResource()` calls to the same resource. The cost of the API calls may then make up a significant part of the overall task execution time and, therefore, potentially the response time.

You may find that placing the entire task or ISR body between `GetResource()` and `ReleaseResource()` calls actually shortens the worst-case response time.

You should avoid using non-preemptive tasks and getting `RES_SCHEDULER` wherever possible. System responsiveness and schedulability is improved when resources are held for the minimum amount of time and when this affects the smallest number of tasks.

6.9 Avoiding Race Conditions

The AUTOSAR OS standard specifies that resources must be released before a `TerminateTask()` call is made. In some circumstances, this can introduce a race condition into your application. This can cause task activations to be missed (you learned about race conditions at the beginning of this chapter).

Code Example 6.6 shows the type of system where race conditions can become a problem. Assume that two BCC1 tasks exchange data over a bounded buffer.

```
#include <Os.h>
TASK(Write)
    /* Highest priority */
    WriteBuffer();
    GetResource(Guard);
    BufferNotEmpty = True;
    ReleaseResource(Guard);
    ChainTask(Read);
}

TASK(Read)
    /* Lowest priority */
    ReadBuffer();
    GetResource(Guard);
    if( BufferNotEmpty ) {
        ReleaseResource(Guard);
        /* !!! Race condition occurs here !!! */
        ChainTask(Read);
    } else {
        ReleaseResource(Guard);
        /* !!! Race condition occurs here !!! */
        TerminateTask();
    }
}
```

Code Example 6.6: A System where a Race Condition can Occur

In Code Example 6.6, between the resource being released and the task terminating, `Read` can be preempted by `Write`. When task `Write` chains task `Read`, the activation will be lost. This is because `Read` is still running. In other words a task is being activated, but it is not in the suspended state.

To solve this problem, you can allow queued activations of the `Read` task. This means that you should make the task BCC2. See Section 4.5.2 for more details.

6.10 Summary

- Resources are used to provide mutual exclusion over access to shared data or hardware resources.
- Tasks and ISRs can share any number of resources.
- All `GetResource()` and `ReleaseResource()` calls must be properly nested.
- All resources must be released before the task or ISR terminates.
- The scheduler can be used as a resource, but internal resources should be used in preference, if possible.
- Internal resources provide a cost free mechanism for controlling preemption between a group of tasks and ISRs

7 Events

In an AUTOSAR OS system, events are used to send signal information to tasks. This chapter explains what events are, how to configure them and how to use them at runtime.

Events can be used to provide multiple synchronization points for extended tasks. A visualization of synchronization is shown in Figure 7.1.

An extended task can wait on an event, causing the task to move into the waiting state. When an event is set by a task or ISR in the system, the waiting task is transferred into the ready state. When it becomes the highest priority ready task it will be selected to run by RTA-OS3.x.

Events are owned by the extended task with which they are associated. Usually, an extended task will be an infinite loop that contains a series of guarded wait calls for the events it owns. The event mechanism therefore allows you to build event driven state machines using OSEK.

If timing behavior is important in your system, all of your extended tasks (in other words, any task that waits for an event) must be lower priority than the basic tasks.

7.1 Configuring Events

Events are configured using `rtaoscfg`. The maximum number of events that can exist in your application is determined by your target hardware. You should consult the *Target/Compiler Port Guide* for your port to find out how many events you can have per task.

When an event is declared it must have:

- A name.
Names are used only to indicate the purpose of an event at configuration time.
- At least one task that uses it.
- An event mask.

The event name that is specified in `rtaoscfg` is used as a symbolic name for the event mask at run-time. A mask is an N-bit vector with a single bit set, where N is the maximum number of events on which a task can wait. The set bit identifies a particular event.

The event name is used at run-time as a symbolic name for the mask. The mask is configured by selecting the bit which indicates the event. Figure 7.2

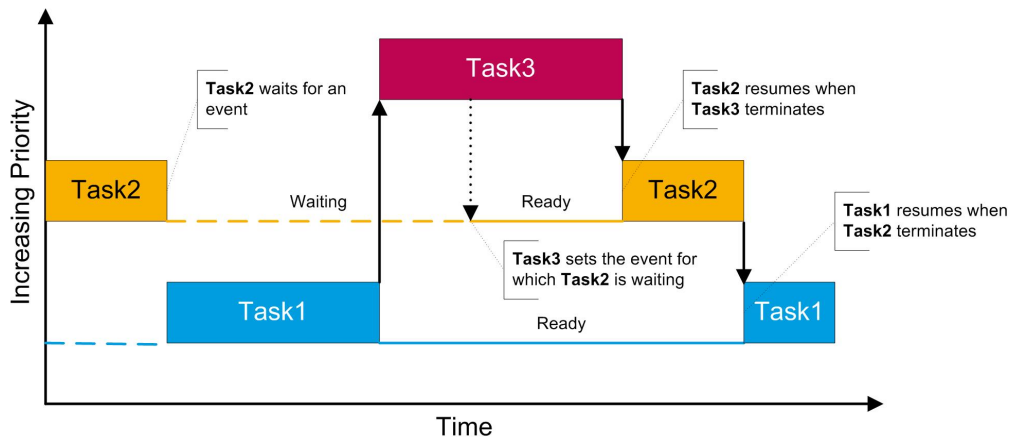


Figure 7.1: Visualizing Synchronization

shows that an event called Event1 has been declared which will be using bit nine in the event mask.

If an event is used by more than one task, each task has its own individual copy. When an event is set, a task must be specified at the same time. So, for example, if you set an event called Event2 for a task called t3, this has no effect on Event2 for the task t4.

7.1.1 Defining Waiting Tasks

Waiting tasks are selected using **rtaoscfg**. If you declare a task that waits on an event, it automatically means that it will be treated as an extended task.

Figure 7.3 shows that an event Event1 has been declared and that the tasks t1 and t2 have been configured to wait on the event.

An extended task that waits on an event will usually be auto-started and the task will never terminate. When the task starts executing, all the events it owns are cleared by RTA-OS3.x.

7.2 Waiting on Events

A task waits for an event using the `WaitEvent(EventMask)` API call. The EventMask must correspond to the one that is declared in **rtaoscfg**.

The `WaitEvent()` takes an event as its sole parameter. When the call executes there are two possibilities:

1. The event has not occurred

In this case the task will enter the waiting state and RTA-OS3.x will run the highest priority task in the ready state.

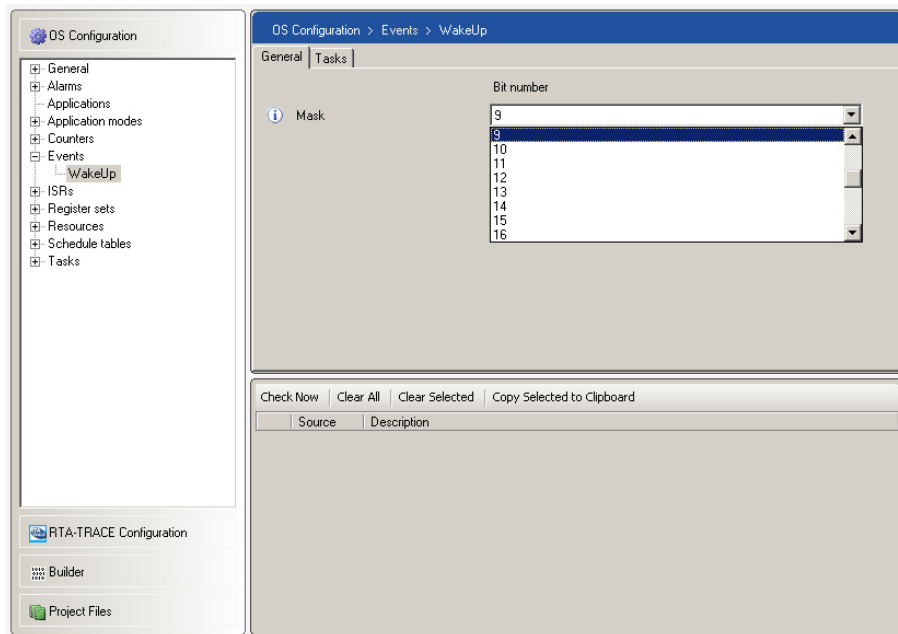


Figure 7.2: Configuring an Event mask in **rtaoscfg**

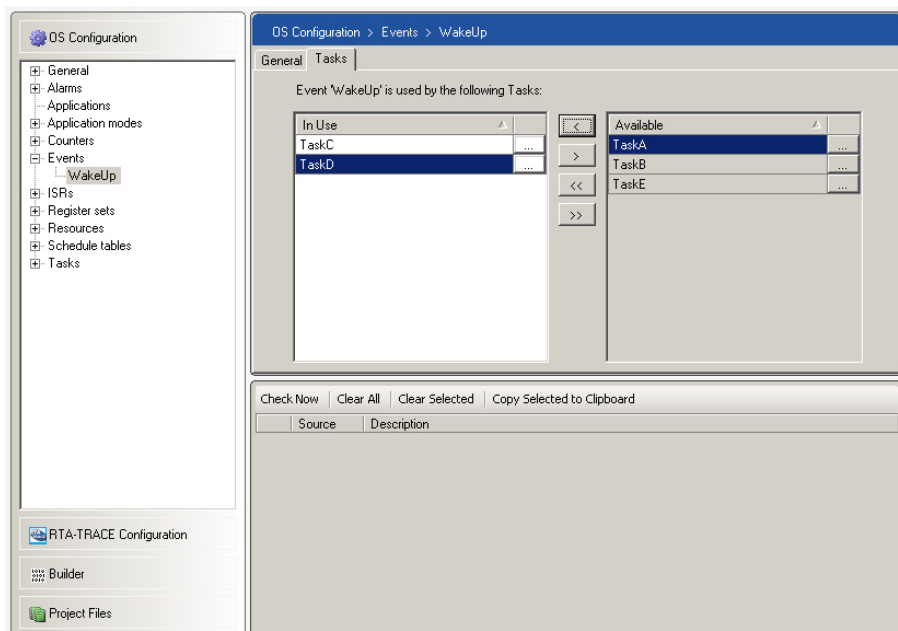


Figure 7.3: Selecting the Task to Wait on an Event

2. The event has occurred

In this case the task remains in the running state and will continue to execute at the statement immediately following the `WaitEvent()` call.

7.2.1 Single Events

To wait on a single event you simply pass in the event mask name to the API call. Code Example 7.1 shows how a task can wait for events.

```
#include <Os.h>
TASK(ExtendedTask) {
    ...
    WaitEvent(Event1); /* Task enters waiting state in API call if
                        Event1 has not happened */
    /* When Event1 is set, ExtendedTask resumes here */
    ...
}
```

Code Example 7.1: Waiting on an Event

In AUTOSAR OS it is illegal to set events for a task that is in the suspended state. In practice this means that the structure of a task that waits on events is typically an infinite loop that waits on events as shown in Code Example 7.2.

```
#include <Os.h>
TASK(ExtendedTask){
    /* Entry state */
    while(true){
        WaitEvent(Event1);
        /* State 1 */
        WaitEvent(Event2);
        /* State 2 */
        WaitEvent(Event3);
        /* State 3 */
    }
    /* Task never terminates */
}
```

Code Example 7.2: Simple 3-state State Machine with Events

7.2.2 Multiple Events

Because an AUTOSAR OS event is just a bit mask, you can wait on multiple events at the same time by bit-wise ORing a set of bit masks.

When your task waits on multiple events it will be resumed when any one of the events upon which you are waiting occurs. When you resume from waiting on multiple events, then you will need to work out which event (or events) has occurred.

OSEK provides the `GetEvent()` API call so that allows you to get the current set of events that are set for the task.

Code Example 7.3 shows how a task can wait on multiple events simultaneously and then identify which of the events has been set when it resumes.

```
#include <Os.h>
TASK(ExtendedTask){
    EventMaskType WhatHappened;
    while(true){
        WaitEvent(Event1|Event2|Event3);
        GetEvent(Task1, &WhatHappened);
        if( WhatHappened & Event1 ) {
            /* Take action on Event1 */
            ...
        } else if( WhatHappened & Event2 ) {
            /* Take action on Event2 */
            ...
        } else if( WhatHappened & Event3 ) {
            /* Take action on Event3 */
            ...
        }
    }
}
```

Code Example 7.3: Waiting on Multiple Events

7.2.3 Deadlock with Extended Tasks

While AUTOSAR OS provides freedom from deadlock in mutual exclusion over a critical section (see Chapter 6) you are not protected from building systems with events that can deadlock. If you have extended tasks that mutually set and wait on events sets, then it is possible that two (or more) tasks will be waiting on events that are only set by other tasks that are waiting. It is, of course, impossible for basic tasks in the system to deadlock, even if there are deadlocking extended tasks present.

Code Example 7.4 shows two tasks that will deadlock if there no other task set either `Ev1` or `Ev2`.

```
#include <Os.h>
TASK(Task1) {
    while (1) {
        WaitEvent(Ev1);
        /* Never reach here - DEADLOCKED with Task2! */
        SetEvent(Task2,Ev2)
    }
}
TASK(Task2) {
    while (1) {
```

```

    WaitEvent(Ev2);
    /* Never reach here - DEADLOCKED with Task1! */
    SetEvent(Task1,Ev1)
  }
}

```

Code Example 7.4: Deadlock with Extended Tasks

OS configuration does not capture which tasks/ISRs set events, only which tasks can wait on events. It is therefore impossible for RTA-OS3.x to statically determine whether your extended tasks will deadlock or not. However, the following design approaches may help:

- use basic tasks only;
- analyze your code to show that there is no circular waiting of events on the transitive closure of all SetEvent()/WaitEvent() pairs.

7.3 Setting Events

Events are set using the SetEvent() API call.

The SetEvent() call has two parameters, a task and an event mask. For the specified task, the SetEvent() call sets the events that are specified in the event mask. The call does not set the events for any other tasks that share the events.

You can bit-wise OR multiple event masks in a call to SetEvent() to set multiple events for a task at the same time

Events cannot be set for tasks that are in the suspended state. So, before setting the event, you must be sure that the task is not suspended. You can do this using the GetTaskState() API call, but note that there is a potential race-condition when this is called for tasks of higher priority than the caller. The caller may be preempted between the call to the API and the evaluation of the result and the state of the task that was requested may have changed in the intervening time.

An extended task is moved from the waiting state into the ready state when any one of the events that it is waiting on is set.

Code Example 7.5 shows you how a task can set events.

```

#include <Os.h>
TASK(Task1) {
    TaskStateType TaskState;

    /* Set a single event */

```

```

SetEvent(Task2, Event1);

/* Set multiple events */
SetEvent(Task3, Event1 | Event2 | Event3);
...
/* Checking for the suspended state */
GetTaskState(Task2,&TaskState);
if (TaskState != SUSPENDED) {
    SetEvent(Task2, Event1);
}
...
TerminateTask();
}

```

Code Example 7.5: Setting Events

A number of tasks can wait on a single event. However, you can see from Code Example 7.5 that there is no broadcast mechanism for events. In other words, you cannot signal the occurrence of an event to all tasks waiting on the event with a single API call.

Events can also be set by alarms and schedule tables.

7.3.1 Setting Events with an Alarm

Alarms can be used to periodically activate extended tasks that don't terminate. Each time the alarm expires, the event is set. The task waiting on the event is then made ready to run.

7.3.2 Setting Events with a Schedule Table Expiry Point

Expiry points on schedule tables can be used to program (a)periodic activations of extended tasks that do not terminate. Each time the expiry point is processed, the event is set. The task waiting on the event is then made ready to run.

7.4 Clearing Events

An event can be set by any task or ISR, but it can only be cleared by the owner of the event.

When a task waits on an event, and the event occurs, then a subsequent call to `WaitEvent()` for the same event will return immediately because the event is still set.

Before waiting for the event occurring again the last event occurrence of the event must be cleared.

Events are cleared using the `ClearEvent(EventMask)` API call. The EventMask must correspond to the one that is declared.

Code Example 7.6 shows how a task typically uses `ClearEvent()`.

```
#include <Os.h>
TASK(ExtendedTask){
    EventMaskType WhatHappened;
    ...
    while( WaitEvent(Event1|Event2|Event3)==E_OK ) {
        GetEvent(Task1, & WhatHappened);
        if(WhatHappened & Event1 ) {
            ClearEvent(Event1);
            /* Take action on Event1 */
            ...
        } else if( WhatHappened & (Event2 | Event3 ) {
            ClearEvent(Event2 | Event3);
            /* Take action on Event2 or Event3*/
            ...
        }
    }
}
```

Code Example 7.6: Clearing Events

When a task terminates all the events that it owns are cleared automatically.

7.5 Simulating Extended Tasks with Basic Tasks

Basic tasks can only synchronize at the start or end of task execution.

If other synchronization points are required then the event mechanism provides one way to do this. However, extended tasks typically have greater overheads than basic tasks. On resource-constrained systems, synchronization can be built using basic tasks only.

For example, if a task is built as a state machine (using a C switch statement, for instance) then you can set a state variable, issue a `TerminateTask()` call and wait for re-activation. Code Example 7.7 shows how this can be achieved.

```
#include <Os.h>
/* Create a "State" variable that remains in scope between task
   activations */
uint8 State;
TASK(Task1) {
    switch (State) {
        case 0:
            /* Synchronization point 0. */
            State = 1;
            break;
```

```
    case 1:
        /* Synchronization point 1. */
        State = 2;
        break;
    case 2:
        /* Synchronization point 2. */
        State = 0;
        break;
}
TerminateTask();
}
```

Code Example 7.7: Multiple Synchronization Points in a Basic Task

7.6 Summary

- Events are synchronization objects that can be waited on by extended tasks.
- An event can be used by multiple tasks.
- Setting an event is not a broadcast mechanism to signal all tasks that are waiting.
- Tasks, ISRs, alarms and schedule tables can set events.

8 Counters

Counters register how many “things” have happened in the OS in terms of ticks. A tick is an abstract unit. It is up to you to decide what you want a tick to mean and, therefore, what are the “things” the counter is counting.

You might define a tick to be:

- Time, for example a millisecond, microsecond, minute etc and the counter then tells you how much time has elapsed.
- Rotation, for example in degrees or minutes, in which case the counter would tell you by how much something has rotated.
- Button Presses, in which case the counter would tell you how many times the button has been pressed.
- Errors, in which case the counter is counting how often an error has occurred.

An ISR (or sometimes a task) is used to drive a counter. The driver is responsible for making the correct RTA-OS3.x API call to “tick” the counter or to tell RTA-OS3.x that the counter has “ticked” to a required value.

8.1 Configuring Counters

Each counter has 4 mandatory attributes:

Name is the name of the counter. RTA-OS3.x creates a handle for each counter using an identifier of the same name as the counter.

Type defines the counter model. AUTOSAR provides two models

Software counters are those where the count value is maintained internally by the OS. You will need to provide a counter driver that tells the RTA-OS3.x to increment the counter by one tick. Further details are provided in Section [8.2.1](#).

Hardware counters are those where a peripheral maintains the count value. You will need to provide a counter driver that tells the OS when a requested number of ticks have elapsed. The OS will also require your driver to provide implementations of callback routines that RTA-OS3.x uses to manage the peripheral at runtime. Further details are provided in Section [8.2.2](#).

A software counter is sufficient when you need a relatively low resolution, for example one millisecond or greater. You should use a hardware counter when you need very high resolution for example in the microsecond range, or where you need to accurately synchronize scheduling of tasks in RTA-OS3.x to an external source, for example a TPU or a global (network) time source.

Maximum Value defines the maximum count value for the counter. All counters wrap around to zero on the tick after the maximum allowed value has been reached¹. In many cases, you will simply use a full modulus wrap for the counter, so this will be 65535 ($2^{16} - 1$) for a 16-bit counter and 4294967295 ($2^{32} - 1$) for a 32-bit counter. The maximum counter value for your port can be found in your *Target/Compiler Port Guide*. This corresponds to the AUTOSAR OS counter attribute MAXALLOWEDVALUE.



For hardware counter you must ensure that MAXALLOWEDVALUE+1 is equal to the modulus of the peripheral.

Minimum Cycle defines the shortest number of ticks allowed when setting a cycle value for an alarm or a schedule table offset. In most cases, you will want this to be 1 tick. However, if you want to build systems where you enforce a minimum separation between alarms on the counter, then you may choose a larger value. This corresponds to the AUTOSAR OS counter attribute MINCYCLE.

Ticks per base is a legacy attribute from AUTOSAR OS that defined the number of underlying counter driver ticks required for each tick on the counter. You can assign any value to this attribute because it is not used by RTA-OS3.x. This corresponds to the AUTOSAR OS attribute TICKSPERBASE.

There is an additional optional attribute:

Seconds Per Tick defines the duration of a tick of the counter in seconds. This should be defined if you want to use the tick/time conversion features provided by AUTOSAR OS. Further details are given in Section 8.5.

Figure 8.1 shows how a counter called MillisecondCounter is declared.

8.2 Counter Drivers

RTA-OS3.x does not take control of any of your hardware to provide counter drivers. This makes RTA-OS3.x very easy to integrate with any tick source

¹This means that the maximum allowed value is equal to the modulus-1 of the counter.

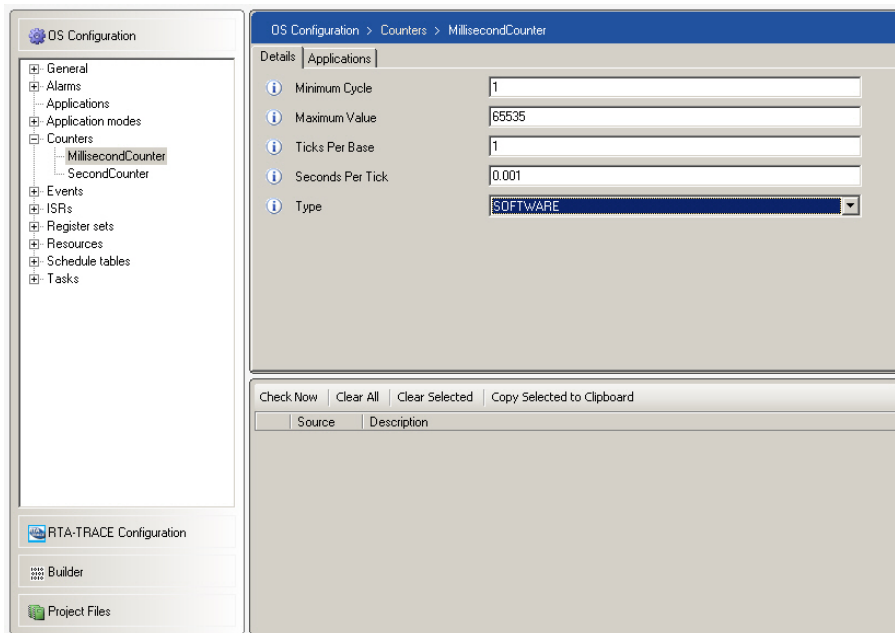


Figure 8.1: Declaring a Counter

for example timer ticks, error counts, button presses, TPU peripherals, etc. This means that you need to provide a driver for every counter you declare in RTA-OS3.x and interface this to the OS.

The interface between the driver and the counter depends on the type of the counter:

Software Counters are incremented by an API call.

Hardware Counters The count value is held in an external hardware peripheral. Your application must provide a more complex driver which tells RTA-OS3.x when a requested number of ticks have elapsed. RTA-OS3.x uses special callbacks to set a requested number of ticks, cancel a request, get the current count value and get the status of the counter.

8.2.1 Software Counter Drivers

For each of your software counters, you need to provide the driver that provides the tick. All software counters are initialized to zero by RTA-OS3.x during `StartOS()` and count upwards.

The software counter driver model is standardized in AUTOSAR OS and is shown in Figure 8.2.

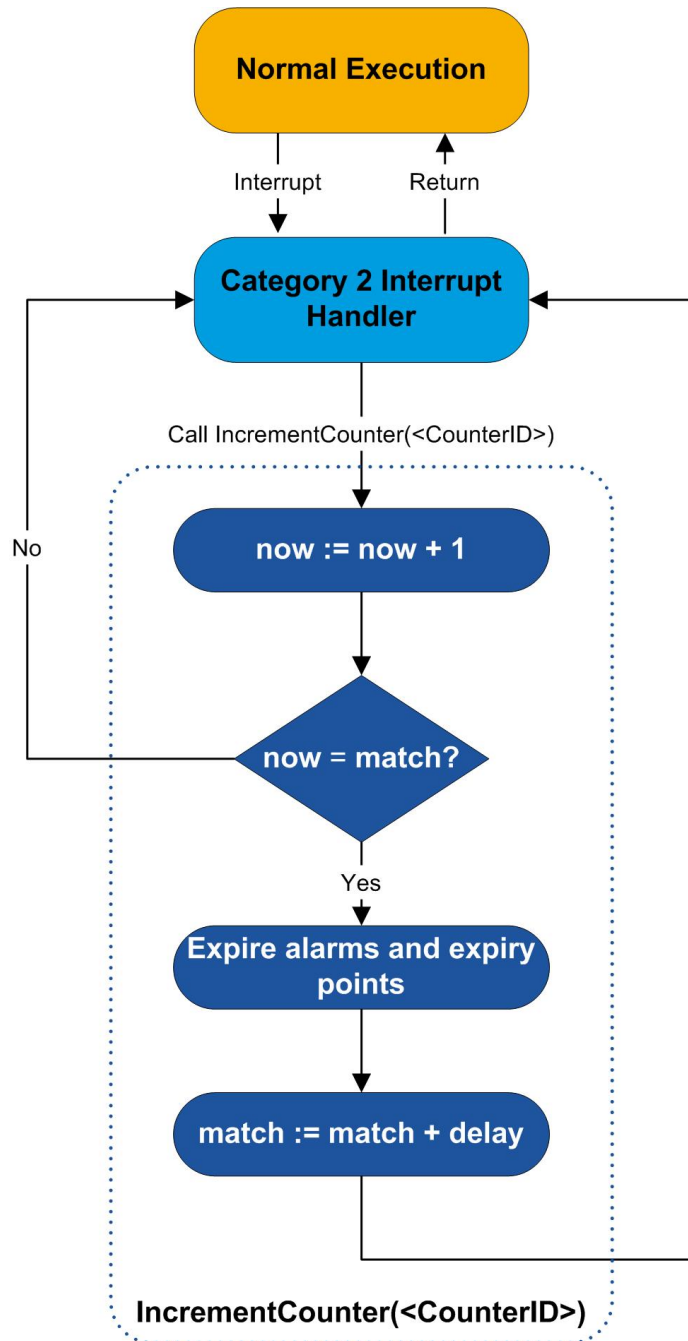


Figure 8.2: Ticked Counter Driver Model

Incrementing Software Counters

You use the API call `IncrementCounter(CounterID)` to increment the counter value held in RTA-OS3.x. The software counter wraps around to zero when one is added to `MAXALLOWEDVALUE`.

You can make the call to `IncrementCounter(CounterID)` from most places in your application code. One of the most common uses of a counter is to provide a time-base to RTA-OS3.x for activating tasks based on alarms (see Chapter 9) or Schedule Tables (see Chapter 10). In this case, you will need to provide a periodic timer interrupt that calls `IncrementCounter(CounterID)` on each expiry.

Code Example 8.1 shows how a millisecond interrupt would driver a counter called `TimeCounter`.

```
#include <Os.h>
ISR(HandleTimerInterrupt) {
    DismissTimerInterrupt();
    IncrementCounter(TimeCounter);
}
```

Code Example 8.1: Using a periodic interrupt to tick a software counter

Another common use of software counters is as part of a fault-tolerant system where some action needs to be taken when an error threshold is exceeded. A software counter can be used to register the number of errors and you can then use an alarm to trigger a recovery action (for example, activate an error recovery Task).

Code Example 8.2 shows how a task called `Critical` might log errors on a counter called `ErrorCounter`.

```
#include <Os.h>
TASK(Critical){
    ...
    if (Error) {
        IncrementCounter(ErrorCounter);
    }
    ...
    TerminateTask();
}
```

Code Example 8.2: Using a periodic Task to tick a software counter

Static Counter Interface

As the AUTOSAR API call takes the name of a counter as a parameter, this means that RTA-OS3.x must internally de-reference the parameter before up-

dating the OS data structures. It also means that the compiler needs to push a parameter on the stack on entry.

Typically however, you know at build time which counter you will be ticking from where. You will also probably be driving the counter from an interrupt handler - the last place where you need to waste time unnecessarily.

RTA-OS3.x recognizes this and generates a dedicated API call called `Os_IncrementCounter_<CounterID>()` for each counter that has been declared in the configuration file (where CounterID is the name of the counter).



The API call `Os_IncrementCounter_<CounterID>()` is not necessarily portable to other AUTOSAR OS implementations.

As an example, consider an application containing two Counters: one called TimeCounter and one called AngularCounter. **rtaosgen** will generate the two API calls shown in Code Example 8.3.

```
Os_IncrementCounter_TimeCounter();
Os_IncrementCounter_AngularCounter();
```

Code Example 8.3: Static Software Counter Interface

The interrupt handlers that you supply to service the timer and angular interrupts must call these API calls.

Code Example 8.4 shows how these interrupt handlers might look.

```
#include <Os.h>
ISR(HandleTimerInterrupt) {
    ServiceTimerInterrupt();
    Os_IncrementCounter_TimeCounter();
}
ISR(HandleAngularInterrupt) {
    ServiceAngularInterrupt();
    Os_IncrementCounter_AngularCounter();
}
```

Code Example 8.4: Interrupt Handlers for Code Example 8.3

If you have multiple software counters that you need to tick at the same rate, then you can make multiple `Os_IncrementCounter_<CounterID>()` calls within your handler as shown in Code Example 8.5

```
#include <Os.h>
ISR(MillisecondInterrupt) {
    ServiceTimerInterrupt();
    Os_IncrementCounter_Counter1();
    Os_IncrementCounter_Counter2();
    ...
}
```

```
    Os_IncrementCounter_CounterN();  
}
```

Code Example 8.5: Making multiple calls to the static software counter interface



There is an `Os_IncrementCounter_<CounterID>()` API call available for each counter you declare. These static API calls are faster and use less RAM than the AUTOSAR `IncrementCounter(<CounterID>)` API call because the calls do not require a parameter and do not need to work out which counter is being ticked. You should decide which version is appropriate for your application and choose accordingly.

8.2.2 Hardware Counter Drivers

For each of your hardware counters, you need to provide the hardware counter driver that calls RTA-OS3.x and a set of callbacks that are used by RTA-OS3.x. As with software counters, RTA-OS3.x provides a well-defined interface for connecting the advanced counter driver to the OS.



The AUTOSAR OS standard does not specify a standard API call for dealing with hardware counters. If you are porting your application from another OS to RTA-OS3.x, then you may need to change the hardware counter driver API calls.

For each hardware counter, RTA-OS3.x knows what the next action driven by the counter is, whether that is to expire an alarm or process an expiry point on a schedule table or both. RTA-OS3.x also knows how many ticks need to elapse before this happens. This is called the **match** value.

When you use a software counter, *the driver tells RTA-OS3.x each time a tick has elapsed*. RTA-OS3.x counts ticks internally and, when the match value is reached, the action is taken. RTA-OS3.x then calculates the next match value and the process repeats.

By contrast, when you use an hardware counter, *RTA-OS3.x tells the driver*, through a callback function, when the next action is needed. Your peripheral counts the requested number of ticks and generates an interrupt when the correct number have elapsed. In the interrupt handler you make the `Os_AdvanceCounter_CounterID()` API call to tell RTA-OS3.x to process the next action due on CounterID. RTA-OS3.x does this and the process repeats.

The driver model is shown in Figure 8.3.

Normally, you will use an interrupt to drive both software and hardware counters. With a software counter, an interrupt occurs for each counter tick, *whether or not there is anything for RTA-OS3.x to do*. With a hardware counter, an interrupt occurs only when RTA-OS3.x needs to do something.

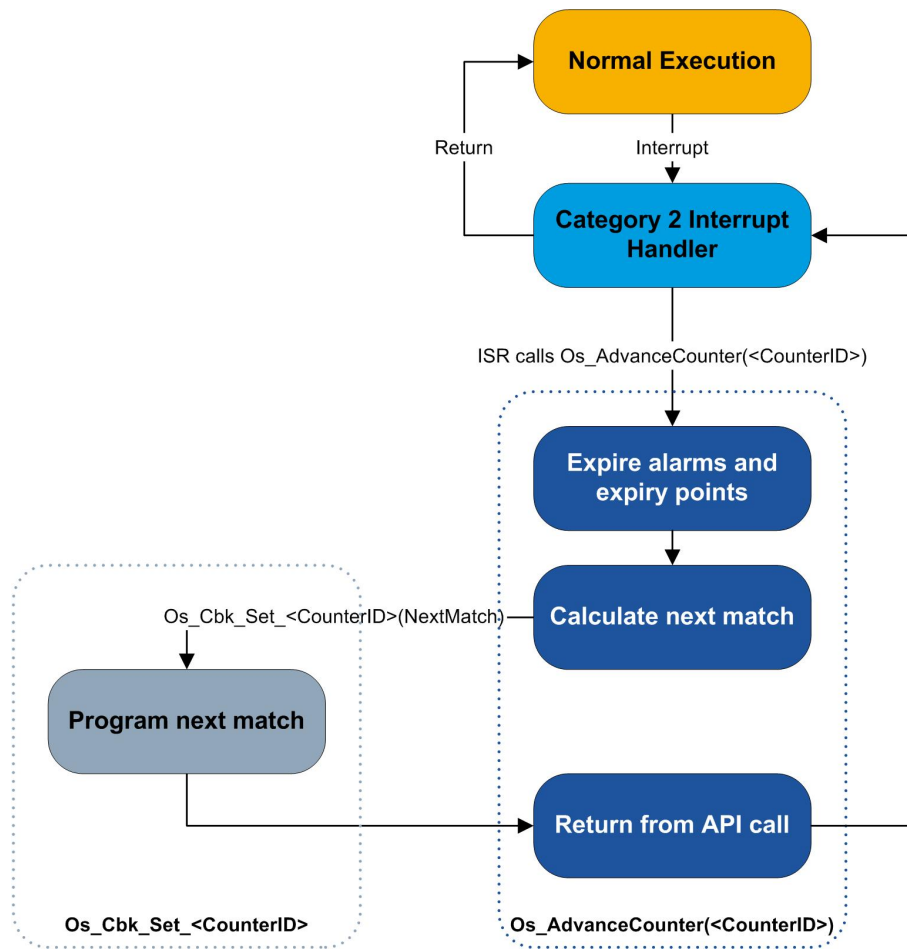


Figure 8.3: Advanced Counter Driver Model

This means that hardware counters reduce interrupt interference to the absolute minimum required.

Advancing Hardware Counters

You use the API call `Os_AdvanceCounter_<CounterID>()` to tell RTA-OS3.x that the match value has been reached.



You are responsible for writing the driver that calls `Os_AdvanceCounter_<CounterID>()` and ensuring that the next action is taken at the correct time.

The `Os_AdvanceCounter_<CounterID>()` API call cause the next alarm and/or expiry point to be processed and will set up the next match value by calling a callback you provide or, if there are no actions left to do (i.e. there are no active alarms or schedule tables on the counter), cancel interrupts from the driver. More detailed information about writing hardware counter drivers can be found in Chapter 11.

Callback Functions

For a software counter communication is one way - the driver tells RTA-OS3.x when a *single* tick has happened. For hardware counters the driver has to tell RTA-OS3.x when *multiple* ticks have happened. However, RTA-OS3.x also needs to tell the driver to driver do things. This is done using a set of callback functions that provide an abstraction between RTA-OS3.x and any type of peripheral you want to use as the driver. The exact functionality of the callbacks depends on the peripheral you are using as your hardware counter driver. Further information on writing callbacks can be found in Chapter 11.

However, by way of a short overview, four callbacks are required:

`Os_Cbk_Set_<CounterID>()`

This callback sets up the state for an interrupt to occur when the next action is due. The callback is passed the absolute value of the counter at which an action should take place. For counters, this callback is used in two distinct cases:

1. Starting
 - Setting the initial interrupt source when a schedule table or an alarm is started on the counter.
2. Resetting
 - Shortening the time to the next counter expiry.

The second case is needed because you can, for example, make a `SetRelAlarm(WakeUp, 100)` call when the next interrupt is due in more than 100 ticks.

`Os_Cbk_State_<CounterID>()`

This callback returns whether the next action on the counter is pending or not and, if the action is not pending, the number of ticks remaining until the match value is reached.

`Os_Cbk_Now_<CounterID>()`

This callback needs to return the current value of the external counter. This is used for the `GetCounterValue()` API call. See Section 8.4.

`Os_Cbk_Cancel_<CounterID>()`

This callback must clear any pending interrupt for your counter and ensure that the interrupt cannot become pending until a `Os_Cbk_Set_<CounterID>()` call is made. If you do not cancel all the alarms on the counter and/or stop schedule tables driven by the counter, then this call is not needed.

8.3 Accessing Counter Attributes at Runtime

The RTA-OS3.x API call `GetAlarmBase()` always returns the configured counter values. The structure of `GetAlarmBase()` is shown in Code Example 8.6.

```
AlarmBaseType Info;
GetAlarmBase(Alarm2, &Info);
MaxValue = Info.maxallowedvalue;
BaseTicks = Info.ticksperbase;
MinCycle = Info.mincycle;
```

Code Example 8.6: Using `GetAlarmBase()` to read static counter attributes

The configured values are can also be accessed as symbolic constants in the form shown below.

- `OSMAXALLOWEDVALUE_<CounterID>`
- `OSTICKSPERBASE_<CounterID>`
- `OSMINCYCLE_<CounterID>`

So Code Example 8.6 above could also have been written as shown in Code Example 8.7:

```

MaxValue = OSMAXALLOWEDVALUE_Alarm2;
BaseTicks = OSTICKSPERBASE_Alarm2;
MinCycle = OSMINCYCLE_Alarm2;

```

Code Example 8.7: Using macros to read static counter attributes

8.3.1 Special Counter Names

If a counter with the name `SystemCounter` is created, then it is possible in AUTOSAR OS R3.x to access the associated counter attributes with a short form of the macros by omitting the trailing `_CounterID`:

```

OSMAXALLOWEDVALUE_SystemCounter → OSMAXALLOWEDVALUE
OSTICKSPERBASE_SystemCounter    → OSTICKSPERBASE
OSMINCYCLE_SystemCounter        → OSMINCYCLE

```

RTA-OS3.x generates both forms of the macros for `SystemCounter` and you can use either version.

The `SystemCounter` also provides an additional macro to get the duration of a tick of the counter in nanoseconds called `OSTICKDURATION`.



The generation of a meaningful `OSTICKDURATION` macro requires the counter attribute “Seconds Per Tick” to be configured.

8.4 Reading Counter Values

You may find that your application has the need to be able to read the current value of a counter at runtime. For example, you might want to know how many errors an error counter has logged, how many times a button has been pressed or how much time has elapsed.

The current value of a counter can be read at runtime by calling the `GetCounterValue()` API as show in Code Example 8.8.

```

TickType HowMany;
GetCounterValue(ButtonPresses,&HowMany);

```

Code Example 8.8: Using `GetCounterValue()`

When you use `GetCounterValue()` you should be aware that:

- counters wrap around from `MAXALLOWEDVALUE` to zero, so the calculation needs to compensate for the wrap
- preemption can occur at the point the call returns meaning that when you resume the value of ‘Now’ will be old.

- when using a hardware counter, the counter driver will still be incrementing when the call returns. Even when preemption does not occur, the calculation performed immediately will be based on old data.

If you need to perform a simple calculation to work out how many ticks of the counter have elapsed since a previously read value, then you can avoid this potential race-condition by using the `GetElapsedCounterValue()` API call. The call takes a previously read counter value as input and calculates the ticks that have elapsed, including compensation for the counter wrapping. The calculation occurs at OS level (i.e. with interrupts disabled) so does not suffer from preemption effects.

Code Example 8.9 shows how you might use this feature to measure the end-to-end (response) time of a task.

```
#include <Os.h>
TickType Start;
ISR(CaptureTrigger){
    /* Dismiss interrupt */
    GetCounterValue(TimeCounter,&Start);
    ...
    ActivateTask(GenerateResponse);
}
TASK(GenerateResponse){
    TickType Finish;
    CalculateValue();
    WriteToDevice();
    GetElapsedCounterValue(TimeCounter,&Start,&Finish);
    ...
    TerminateTask();
}
```

Code Example 8.9: Using `GetElapsedCounterValue()`

If your counter is counting time ticks (as in Code Example 8.9), then this is referred to in AUTOSAR OS as a “free running timer”. There is nothing special about this type of counter - it is identical to any other type of counter - the only distinction is that the counter is one which is driven by a timer tick source.

The intended use of the free running timer functionality is to measure short, high accuracy, durations at runtime. If you need to do this, then you will probably need to use a hardware counter to get the required counter resolution.

8.5 Tick to Time Conversions

It is common for counters to be used as a time-base reference for the OS. For most of the applications that you write, the relative timing of events will

be the real-time values determined by your system requirements. You will most likely think about system configuration in terms of real-time values, nanoseconds, milliseconds etc, rather than in the more abstract notion of ticks.

If a counter configuration parameter ‘Seconds Per Tick’ has been configured, then RTA-OS3.x generates macros for you to use to convert between ticks and real-time units.

AUTOSAR OS states that tick to time conversion is for hardware counters only. However, the feature is generally useful for both software and hardware counters and the AUTOSAR XML configuration language supports configuration for both types of counter. In RTA-OS3.x this anomaly is resolved by providing tick to time conversion for both software and hardware counters. However, you should note that the provision of these macros for software counters is not necessarily supported by other AUTOSAR OS implementations.

The following macros are provided:

- OS_TICKS2NS_CounterID(ticks) converts ticks to nanoseconds
- OS_TICKS2US_CounterID(ticks) converts ticks to microseconds
- OS_TICKS2MS_CounterID(ticks) converts ticks to milliseconds
- OS_TICKS2SEC_CounterID(ticks) converts ticks to seconds

The values returned by these macros are of `PhysicalTimeType` rather than `TickType` that are used by the API calls that you might use the macros with, so you will need to cast them to an appropriate type.

Code Example 8.10 shows how these macros might be used in your application code to program a timeout using a statically defined “timeout” value.

```
#define TIMEOUT_MS 100 /* Set a timeout to be 100ms */
TickType TimeoutInTicks;
TimeoutInTicks =
    (TickType)((PhysicalTimeType)TIMEOUT_MS/OS_TICKS2MS_TimeCounter(1));
SetRelAlarm(TimeoutAlarm, TimeoutInTicks, 0);
```

Code Example 8.10: Programming an alarm with time rather than ticks (1)

In addition to these macros RTA-OS3.x, generates a macro called `OSTICKDURATION_<CounterID>` that returns the duration of a counter tick in nanoseconds so this makes it extremely useful if you want to program alarms of a fixed time, even if you change the underlying counter tick rate. Code Example 8.11 shows how Code Example 8.10 can be reworked using the

OSTICKDURATION_<CounterID> macro. This version offers slightly better performance because the duration of a single tick does not need to be calculated at runtime.

```
#define TIMEOUT_NS 100000000 /* Set a timeout to be 100ms */
TickType TimeoutInTicks;
TimeoutInTicks =
    (TickType)(TIMEOUT_NS/OSTICKDURATION_TimeCounter);
SetRelAlarm(TimeoutAlarm, TimeoutInTicks, 0);
```

Code Example 8.11: Programming an alarm with time rather than ticks (2)



The OSTICKDURATION_<CounterID> macros are provided by RTA-OS3.x and are not part of the AUTOSAR OS standard. Use of the macros is not portable to other implementations.

8.6 Summary

- Counters are used to register a count of some tick source.
- Counters are either software or hardware counters. You need to provide the appropriate driver for the type of the counter you configure.

9 Alarms

It is possible to construct systems that activate tasks at different rates using ISRs. However, for complex systems, this can become inefficient and impractical. Alarms provide a more convenient, and more portable, way of scheduling systems.

The alarm mechanism consists of two parts:

1. A counter.
These were covered in [Chapter 8](#).
2. One or more alarms attached to the counter.

The alarm part specifies an action to perform when a particular counter value is reached. Each counter in your system can have any number of alarms attached.

An alarm is said to have expired when the value of a counter equals the value of an alarm attached to the counter. On expiry, RTA-OS3.x will perform the action associated with the alarm. The action could be to activate a task, to execute an alarm callback routine, set an event or tick a software counter.

The alarm expiry value can be defined relative to the actual counter value or as an absolute value. If the alarm expiry is defined as relative to the actual counter, it is known as a relative alarm. If it is defined as an absolute value, it is known as an absolute alarm.

Alarms can be configured to expire once. An alarm that expires once is called a single-shot alarm.

An alarm can also be specified to expire on a periodic basis. This type of alarm is called a cyclic alarm. You can find out more about cyclic alarms in [Section 9.2](#).

9.1 Configuring Alarms

There are three parts to alarm configuration:

1. Naming - Each alarm in your system needs to be allocated a unique name. As for other OS objects, this is the name that you will use in your code to refer to the alarm at runtime.
2. Association of a counter - An alarm is statically bound to a counter at configuration time. Any setting of the alarm is done in terms of ticks of the associated counter.

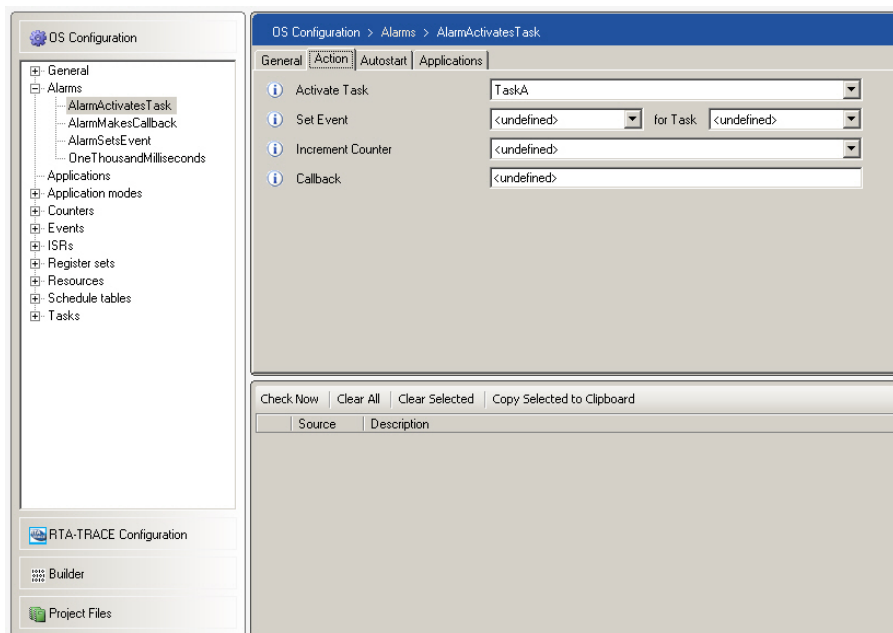


Figure 9.1: Activating a Task with an Alarm

3. Specification of the alarm's action.

Each alarm that you create is associated with up to 4 actions:

1. Activate a task.
2. Raise an event.
3. Execute a callback function.
4. Increment a (software) counter.

If you need to activate multiple tasks, set multiple events, make multiple callbacks or increment multiple counters on expiry, you will need multiple alarms with the same expiry value. (Schedule Tables provide an alternative mechanism that allows you to activate multiple tasks and/or set multiple events simultaneously. You can read about Schedule Tables in Chapter 10).

9.1.1.1 Activating a Task

The most common action for an alarm is to activate a task. This is the basis for building systems with periodically activated tasks - you create an alarm for each task and then program the alarm to occur at the required period. Figure 9.1 shows how to configure an alarm to activate a task.

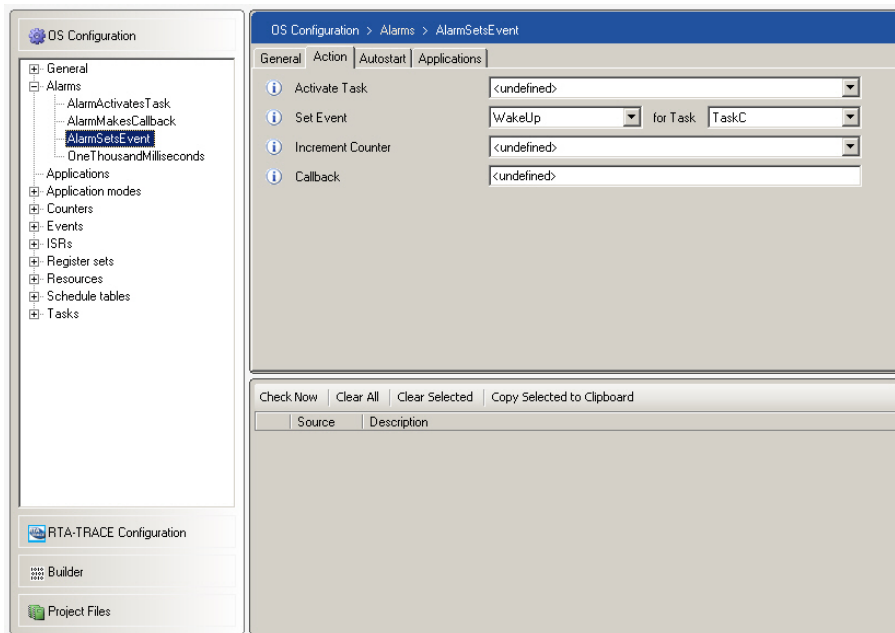


Figure 9.2: Setting an Event for a Task with an Alarm

In AUTOSAR OS, you may only activate a single task for each alarm. If you need multiple tasks to run when an alarm expires, then you can do this either by creating multiple alarms or by using task activation chains (see Section 4.9.1).

9.1.2 Setting an Event

An alarm can set an event for a specified task. When an event is set with an alarm, it has the same properties as it would if it were set using the `SetEvent()` API call. This means you need to specify both the event and the task for which the event is to be set. Figure 9.2 shows you how to set an event action for an alarm.

9.1.3 Alarm Callbacks

Each alarm can have an associated callback function. The callback is simply a C function that is called when the alarm expires.



Alarm callbacks are only permitted in Scalability Class 1 according to the AUTOSAR OS R3.x standard. This is because these callbacks run at OS level and can therefore interfere with timing protection and also present a hole in the memory protection scheme.



*RTA-OS3.x allows you to relax this restriction by allowing you to allow alarm callbacks in all Scalability Classes using a configuration option in **General → Optimizations → RTA-OS**.*

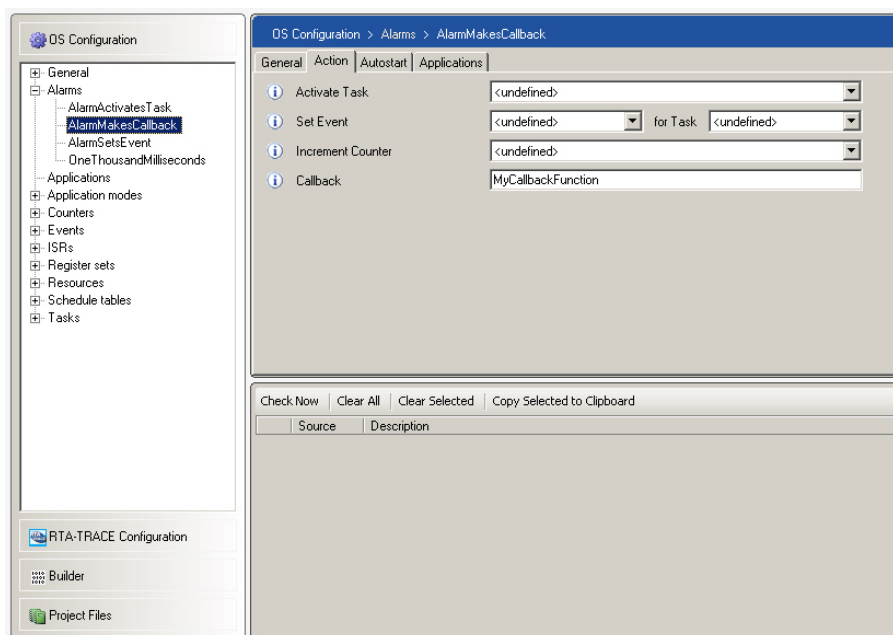


Figure 9.3: Configuring a Callback Routine for an Alarm

Figure 9.3 shows how to configure a callback routine for an alarm.

Each callback routine must be written using the `ALARMCALLBACK()` macro, shown in Code Example 9.1.

```
ALARMCALLBACK(UserProvidedAlarmCallback) {

    /* Callback code. */

}
```

Code Example 9.1: An Alarm Callback



Callback routines run at OS level, which means Category 2 interrupts are disabled. You should therefore aim to keep your callback routines as short as possible to minimize the amount of blocking that your tasks and ISRs suffer at runtime.

The only RTA-OS3.x API calls that you can make inside the callback are the `SuspendAllInterrupts()` and `ResumeAllInterrupts()` calls.

9.1.4 Incrementing a Counter

Incrementing a software counter from an alarm allows you to cascade multiple counters from a single ISR. A counter ticked from an alarm inherits the period of the alarm. So, if you have an alarm that occurs every 5 milliseconds, you can use the alarm to drive a second ticked counter that ticks every 5 milliseconds. Figure 9.4 shows you how this is configured in RTA-OS3.x.

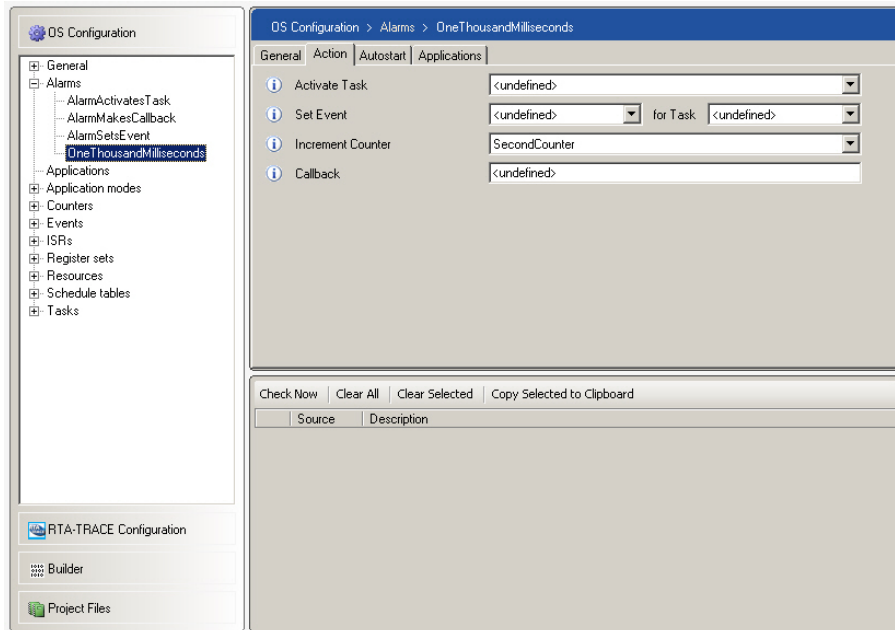


Figure 9.4: Cascading counter increments from an alarm

Code Example 9.2 shows how you would drive Counter1ms from an interrupt. Every fifth interrupt registered on Counter1ms would cause the alarm to expire and increment the cascaded Counter5s :

```
#include <Os.h>
ISR(MillisecondInterrupt){
    CLEAR_PENDING_INTERRUPT();
    Os_IncrementCounter(Counter1ms);
    /* Every 5th call internally performs
       Os_IncrementCounter(Counter5ms) */
}
```

Code Example 9.2: Cascading Counters

Cascaded counters must have a tick rate that is an integer multiple of the counter driving the alarm. You can configure systems with multiple levels of cascading. However, RTA-OS3.x will generate an error if you try and configure a system with a cycle in the cascade or you try and increment a hardware counter.



The timing properties of a cascaded counter are defined relative to timing properties of the first counter in the cascade. The earliest counter in the cascade therefore determines the base tick rate from which all other counters are defined. If you change the tick rate of the earliest counter, then the entire timing behavior of the application will be scaled accordingly.

9.2 Setting Alarms

Two API calls are provided for setting alarms:

- `SetAbsAlarm(AlarmID, start, cycle);`
Sets the alarm to expire when the counter value next reaches the value `start`. You should be aware that if the underlying counter already has value `start` when the call is made, then the alarm will not occur until the counter has 'wrapped around'.
- `SetRelAlarm(AlarmID, increment, cycle);`
Sets the alarm to expire `increment` ticks from the current count value when you make the call. This means that `increment` is a tick offset from the current counter tick value.

In these two API calls, a `cycle` value of zero ticks indicates that the alarm is a single-shot alarm, which means that it will expire only once before being canceled. A `cycle` value greater than zero defines a cyclic alarm. This means that it will continue expiring every `cycle` ticks after the first expiry has occurred. Setting a non-zero `cycle` value gives you an easy way to configure periodic alarms that occur with a periodicity of `cycle` ticks.

Selecting Parameters

If the activated task is BCC1 or ECC1/2 there will be no queued activation. This means that if the `start` or `increment` value is very short, or the `start` value is very close to the current counter value, then this may cause undesired side effects. The alarm will try to activate the task while a previously activated instance is still executing. The activation would be lost and an `E_OS_LIMIT` error would be raised (see Chapter 13 for more information about error codes and how to debug use of RTA-OS3.x at runtime). You must make sure that enough time is allowed for the task to complete before the next alarm which results in a re-trigger of the task occurs.

9.2.1 Absolute Alarms

Single Shot

An absolute alarm specifies the absolute value of the underlying counter at which the alarm expires. Single shot absolute alarms are useful for monitoring things against a pre-defined threshold value - the alarm can be configured to expire when the threshold is exceeded. You might want to count the number of errors that occur in data samples taken at runtime and then trigger a recovery action when the number of errors reaches a dangerous level. This is shown in Code Example 9.3.

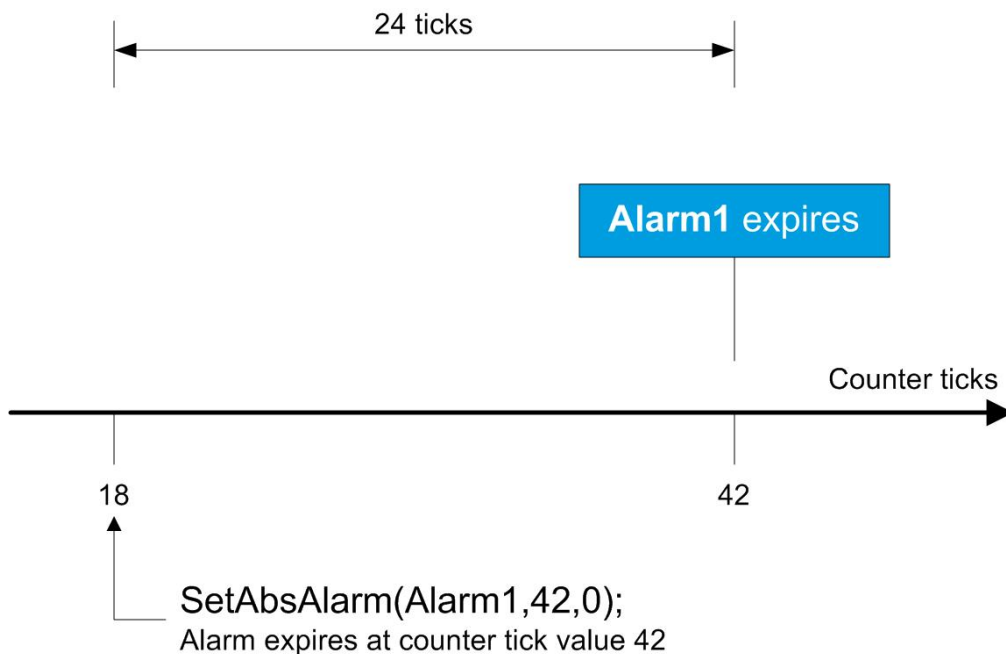


Figure 9.5: Illustration of an Absolute Single Shot Alarm

```
/* Expire when counter value reaches 42. */
SetAbsAlarm(DangerLevelReached, 42, 0);
```

Code Example 9.3: Absolute single shot alarm

Code Example 9.3 is illustrated in Figure 9.5.

A single shot alarm is useful when you need to program a timeout that waits for a fixed amount of time and then takes an action if the timeout occurs.

Cyclic

If an absolute alarm specifies a non-zero cycle value then it will first expire at the specified start tick and then every cycle ticks thereafter. This is shown in Code Example 9.4.

```
/* Expire when counter value reaches 10 and then every 20 ticks
   thereafter */
SetAbsAlarm(Alarm1, 10, 20);
```

Code Example 9.4: Absolute cyclic alarm

The behavior of the code example is illustrated in Figure 9.6.

For absolute alarms, an absolute start value of zero ticks is treated in the same way as any other value - it means expire the alarm when the counter reaches the value zero.

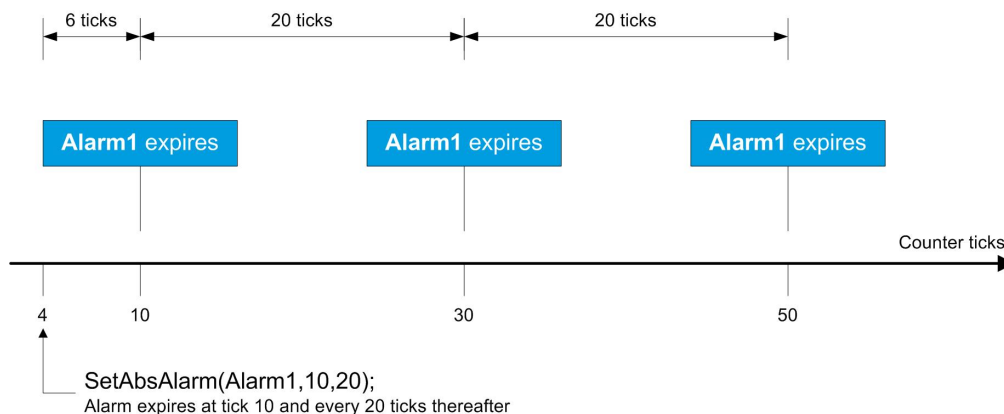


Figure 9.6: Illustration of the Absolute Cyclic Alarm

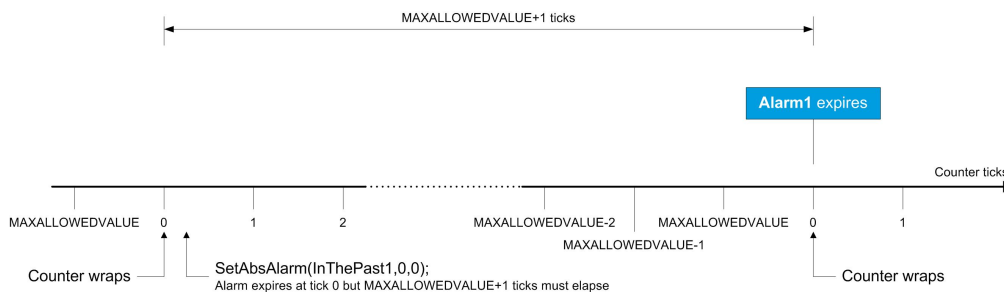


Figure 9.7: Setting an alarm in the past

For example, if the current counter value was zero then you would not see your alarm expire until the $\text{MAXALLOWEDVALUE}+1$ number of counter value ticks had happened. On the other hand, if the counter value was already at MAXALLOWEDVALUE , then you would see the alarm expire on the next tick of the counter.

Setting Alarms in the past

With an absolute alarm it is possible to set the start time to be a value that is already in the past. This does not mean that the alarm will not happen. Recall that counters wrap around when they reach MAXALLOWEDVALUE . So, when you set an alarm in the past you might have to wait up to $\text{MAXALLOWEDVALUE}+1$ (i.e. the counter modulus) ticks until the alarm occurs.



If you set an alarm to start at tick T and the value of the counter is already T then the alarm will not expire immediately. This is because T is already in the past when the alarm is set.

A common error is to set an absolute alarm to occur at zero when the OS starts and then wonder why it does not occur when expected. This is because zero is already in the past! The effect is shown in Figure 9.7.

Synchronizing Absolute Cyclic Alarms to a Counter Wrap

Setting an alarm to occur periodically at a known synchronization point is extremely important for real-time systems. However, in AUTOSAR OS, it is not possible to set an absolute alarm to occur periodically each time the underlying counter wraps around.

For example, assume you have a counter that counts in degrees with a resolution of one degree and you want to activate a task at “top dead center”, i.e. on each revolution of the crankshaft.

For example, assume that the counter has a modulus of 360 ticks. What you need to say is `SetAbsAlarm(Alarm1, 0, 360)`. This is forbidden by the AUTOSAR OS standard because the `cycle` parameter cannot be greater than `MAXALLOWEDVALUE`, which is always the modulus-1 (in this case 359).

If you need this type of functionality, you must provide code that resets an absolute single-shot alarm each time the alarm expires.

For example, if `Task1` is attached to `Alarm1`, then the body of `Task1` will need to reset the alarm when the task is activated as shown in Code Example 9.5.

```
TASK(Task1) {
    /* Single-shot alarm reset at top dead center = 0 = 360
       degrees. */
    SetAbsAlarm(Alarm1, 0, 0);
    /* User code. */
    TerminateTask();
}
```

Code Example 9.5: Resetting an Alarm when a Task is Activated

9.2.2 Relative Alarms

Single-Shot

A relative alarm specifies the absolute value of the underlying counter at which the alarm expires. Single shot relative alarms are useful when you want to timeout some activity at runtime. For example, you might want to wait for an external event and then activate a task if the event does not occur.

Code Example 9.6 shows how an absolute single shot alarm can be set.

```
/* Timeout 42 ticks from now */
SetRelAlarm(Timeout, 42, 0);
```

Code Example 9.6: Relative single shot alarm

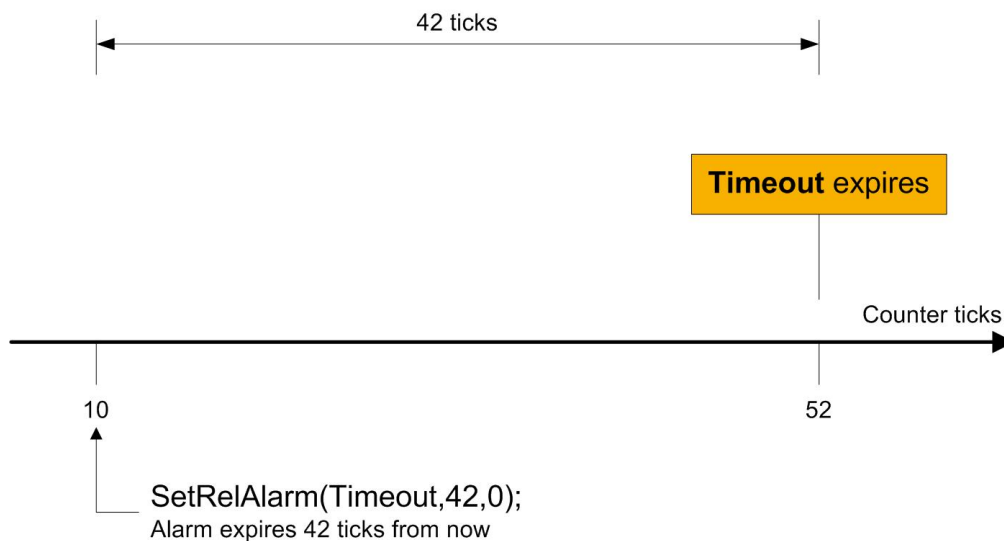


Figure 9.8: Illustration of a Relative Single Shot Alarm

Code Example 9.6 is illustrated in Figure 9.8

A single shot alarm is useful when you need to program a timeout that waits for a fixed amount of time and then takes an action if the timeout occurs.

In AUTOSAR OS, the use of zero for increment in `SetRelAlarm()` is forbidden. If you use zero for increment, then an `E_OS_VALUE` error will be returned.

Cyclic

Code Example 9.7 shows a relative alarm that expires after 10 ticks and then every 20 ticks thereafter.

```
/* Expire after 10 ticks, then every 20 ticks. */
SetRelAlarm(Alarm1, 10, 20);
```

Code Example 9.7: Relative cyclic alarm

In Figure 9.9, you can see how this alarm can be visualized.

9.3 Auto-starting Alarms

It is possible to start alarms by calling `SetRelAlarm()` or `SetAbsAlarm()` in the main program. However, the easiest way to set cyclic alarms is to make them auto-started. Auto-started alarms are started during `StartOS()`.

Auto-started alarms can be set on a per application mode basis so you can choose in which application modes the alarm is auto-started. Each auto-started alarm must also specify whether it is started at an absolute or a rela-

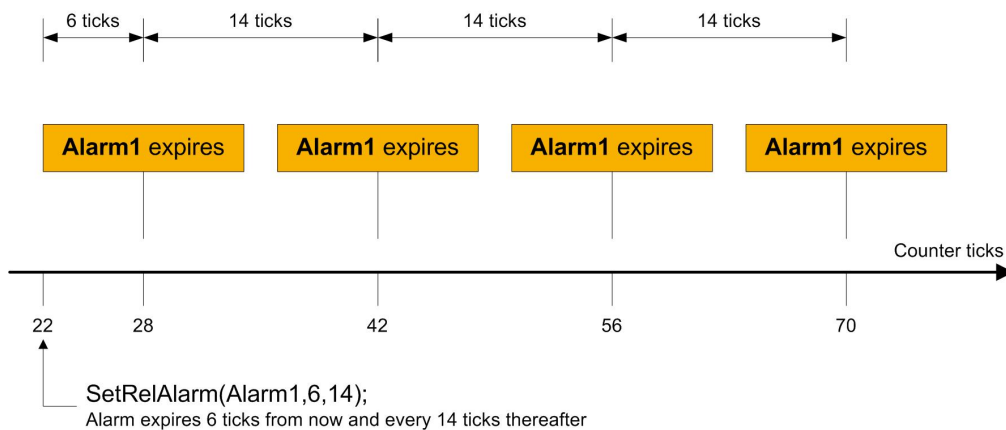


Figure 9.9: Illustration of a Relative Cyclic Alarm

tive counter value and the associated increment/start and cycle parameters must be configured.



Even though alarms may be started in different application modes it is not possible to assign different auto-start parameters for each mode.

Figure 9.10 shows how alarms can be set to auto-start from the Startup Modes pane.

RTA-OS3.x ensures that software counters are initialized to zero during `StartOS()` (hardware counters will be set to the value configured by your own application initialization code). As a result of this, you must take care if you use the a start time of zero ticks for an absolute alarm because the zeroth tick will have already happened when the alarm is started. Although the alarm will be started, it will not expire until the associated counter has wrapped around to zero. On a 16-bit counter ticked every millisecond you would need to wait just over 65 seconds for this to happen, and on a 32-bit counter just under 48 days. Specifying that the alarm starts on the first (or later) tick means that the initial expiry will occur on the next tick of the counter.

Auto-started absolute alarms are useful if you require alarms to be synchronized with each other (i.e. the relative expiries between alarms have to occur a pre-configured number of ticks apart).

9.4 Canceling Alarms

You can cancel an alarm using the `CancelAlarm()` API call.

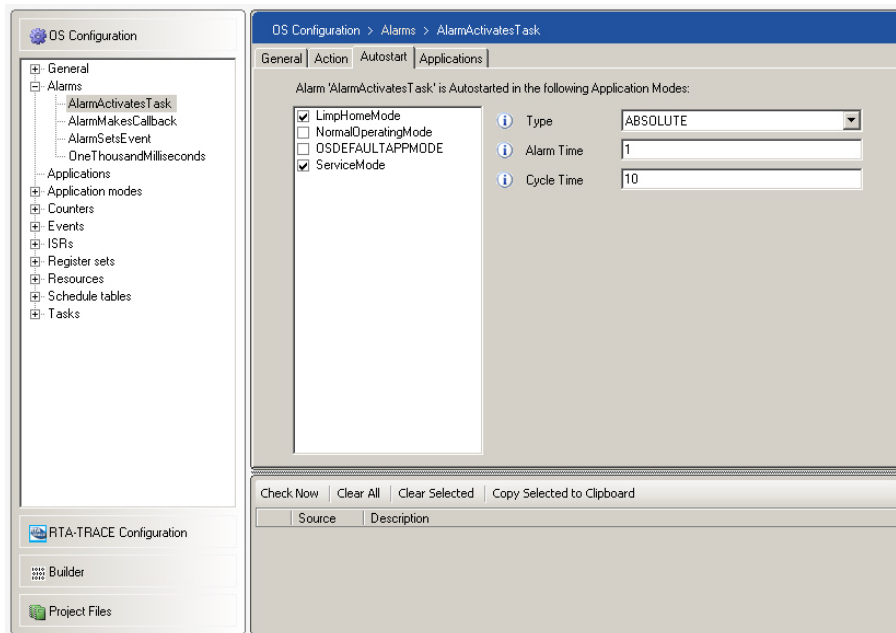


Figure 9.10: Auto-starting Alarms

An alarm may, for example, need to be canceled to stop a particular task being executed. An alarm can be restarted using the `SetAbsAlarm()` or the `SetRelAlarm()` API call.

9.5 Working out when an Alarm will occur

If you need to work out when an alarm will occur, for example, to avoid setting an absolute alarm when the absolute value has already been reached, then you can use the `GetAlarm()` API call.

The call returns the number of ticks remaining before the specified alarm expires. If the alarm is not set, then the API call returns the value `E_OS_NOFUNC` and the number of ticks to expiry is undefined. It is recommended that the return value of the call is checked before using the result. Code Example 9.8 shows the use of the API call.

```

TickType    TimeToExpiry;
TickType    SafetyMargin = 100;
StatusType  IsValid;
IsValid = GetAlarm(Alarm1, &TimeToExpiry);
if (IsValid != E_OS_NOFUNC) {
    if (TimeToExpiry <= SafetyMargin) {
        Log(InsideSafetyMargin);
    }
}

```

Code Example 9.8: Getting the time to expiry

You should exercise caution when making runtime decisions based on the number of ticks returned by the call, especially if the underlying counter has a high resolution. As with reading counter values with `GetCounterValue()`, preemption can occur between getting the value and using it for calculation. This means that you may read a (long) time to expiry but then be preempted to resume shortly before the alarm expires (or even after it has expired).

9.6 Non-cyclic (aperiodic) Alarms

Cyclic alarms are only useful for programming cyclic behavior. In many systems, for example those that need to execute tasks periodically to poll data sources, this is ideal. However, you may need to program systems where the time between successive expiries of an alarm needs to change at runtime. For example, you might be calculating an engine shaft speed and using this to program the duration of spark or injection timing.

Aperiodic behavior with alarms need to be programmed using single-shot alarms that are set to the next expiry value by the activated task.

In Code Example 9.9, a task runs every millisecond and polls a counter that registers degrees of rotation of a crankshaft. The task calculates the position and speed of the crank. The speed is used to determine the duration of the spark timing. The spark is started and an alarm is set to expire after `SparkTiming` ticks.

```
TASK(MillisecondTask) {
    ...
    GetElapsedCounterValue(ShaftEncoder,&Position,&DegreesRotation);
    RevsPerMinute = (DegreesRotation/360) * 1000 * 60;
    SparkTiming = Lookup(RevsPerMinute);
    if (Position = 90) {
        StartSpark();
        SetRelAlarm(TimeCounter, SparkTiming, 0); /* Activates
            SparkOff on expiry */ }
    }
    ...
    TerminateTask()
}
TASK(SparkOff){
    StopSpark();
    TerminateTask();
}
```

Code Example 9.9: Aperiodic Alarm Example

9.7 Summary

- Alarms are set on an underlying counter.

- You can set multiple alarms on each counter.
- Each alarm specifies an action, either:
 - activation of a task,
 - setting an event,
 - execution of a callback, or
 - ticking a ticked counter.
- Alarms can be set to expire at an absolute or relative (to now) counter value.
- Alarms can be auto-started.

10 Schedule Tables

In Chapter 9 you saw that you can build systems requiring periodic and aperiodic behavior relatively easily. However, one of the limitations of alarms is that you can only perform one action per alarm. If you need to build a system where you have a phased sequence of task activations and guarantee some separation in time (temporal separation) then you need to be quite careful how you start and stop the alarms.

While it is possible to build such a system with alarms, there is nothing, other than code review, that prevents the timing properties of the application being accidentally modified at runtime. Furthermore, you saw that if you wanted to define multiple task activations at a single point in time, you were forced to create multiple alarms when what you really want to do is to activate multiple tasks from a single alarm.

AUTOSAR OS addresses the limitations of alarms by providing an OS object called a schedule table.

A schedule table comprises a set of expiry points that occur on statically configured offsets from a notional zero. The offsets are specified in ticks of a statically bound counter - just like the expiry of alarms. The key difference between schedule tables and alarms is that the expiry points on a schedule table maintain their relative separation (to each other). In most cases, this separation is fixed at build time. However, for explicitly synchronized schedule tables this separation can vary within a pre-configured range of values. The schedule table can be started and stopped as a composite unit and, whenever it is restarted, the expiry points always have the same relative execution behavior: any given point A will always be followed by the next point B.

Schedule tables adopt the following terminology:

Initial Offset is the offset to the first expiry point on the schedule table. It is therefore the smallest offset configured.

Duration is the number of ticks from zero before the schedule table stops.

Final Delay is the difference between the offset to the final expiry point and the duration. It is therefore equal to the value of duration minus the longest offset.

Delay is the number of ticks between adjacent expiry points and is equal to the longer offset minus the shorter offset. If the schedule table repeats, then the delay between the last and the first expiry point is equal to the Final Delay plus the Initial Offset.

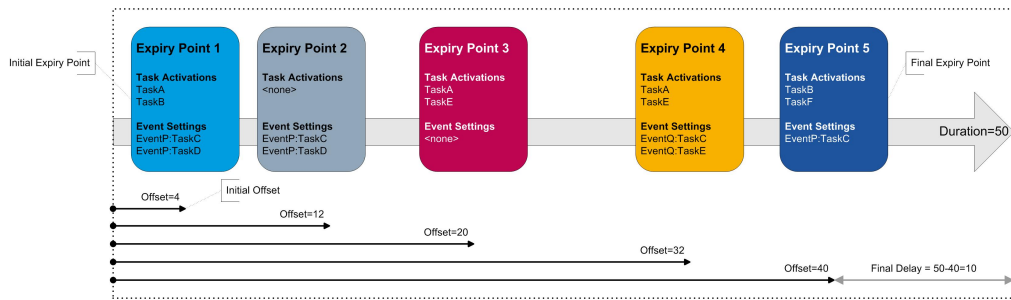


Figure 10.1: Visualizing a Schedule Table

Repeating defines whether the schedule table runs once each time it is started and then stops automatically, or whether it repeats until it is stopped using an OS API call.

An expiry point is similar to an alarm in that it indicates a number of ticks at which RTA-OS3.x needs to take some action. However, expiry points differ from alarms in what actions can be taken. The following table shows the differences:

Action	Alarm	Expiry Point
ActivateTask()	✓- one task	✓- multiple tasks
SetEvent()	✓- one event	✓- multiple events
Callback	✓	✗
IncrementCounter()	✓	✗

Figure 10.1 shows the anatomy of a schedule tables with 5 expiry points and a duration of 50 counter ticks. When the schedule table was started¹, each expiry point would occur every 50 ticks with offset ticks from the notional zero point.



We use the term “notional zero” to mean the point from which offsets are measured. The notional zero is unrelated to values on the underlying counter.

When a schedule table is started (see Section 10.4) the notional zero will be mapped onto the appropriate “now” value of the underlying counter.

Configuration of a schedule table has two parts:

1. configuration of the attributes of the table itself (Section 10.1);
2. configuration of the expiry points on the table (Section 10.2).

¹And assuming it was configured as repeating

In addition (since configuration of a Schedule Table can be complex), **rtaoscfg** also provides a *Schedule Table Assistant* which is described in Section 10.3.

10.1 Configuring a Schedule Table

Each schedule table needs to be assigned a unique name so that it can be referred to in API calls.

A schedule table, like an alarm, is driven by an OS counter. The counter provides the schedule table with a tick source that provides the reference for the expiry point offsets. All offsets are specified in terms of ticks of the OS counter.

The counter used to drive the schedule table can be shared with other schedule tables and also with alarms. However, while multiple alarms can be running on the same counter concurrently, it is only possible to have one schedule table running per counter at any point in time. The following table shows the constraints between the schedule tables and alarms attached to the counter.

Object	Configuration	Active Objects at Runtime
Alarm	0..n	0..n
Schedule Table	0..n	0..1

Each schedule table has a statically configured length called the *duration*. The duration is specified in ticks of the counter and must be in the range $\text{MINCYCLE} \leq \text{Duration} \leq \text{MAXALLOWEDVALUE} + 1$.

A schedule table has two modes of operation:

Single shot: the schedule table runs once and stops automatically at the end. Each expiry point is processed once. Single-shot schedule tables are useful when you want to start a phased sequence of actions in response to some triggering event.

Repeating: the schedule table runs multiple times. When the end of the table is reached, it wraps around to zero starts again. Expiry points are therefore processed periodically with period equal to the duration.

Figure 10.2 shows the configuration of a schedule table called *MasterPlan*, driven by Millisecond counter, which repeats with a duration of 50 ticks.

Figure 10.3 shows how the schedule table in Figure 10.1 with the configuration from Figure 10.2 would run when started at a counter value of 0 ticks and again at a counter value of 42 ticks.

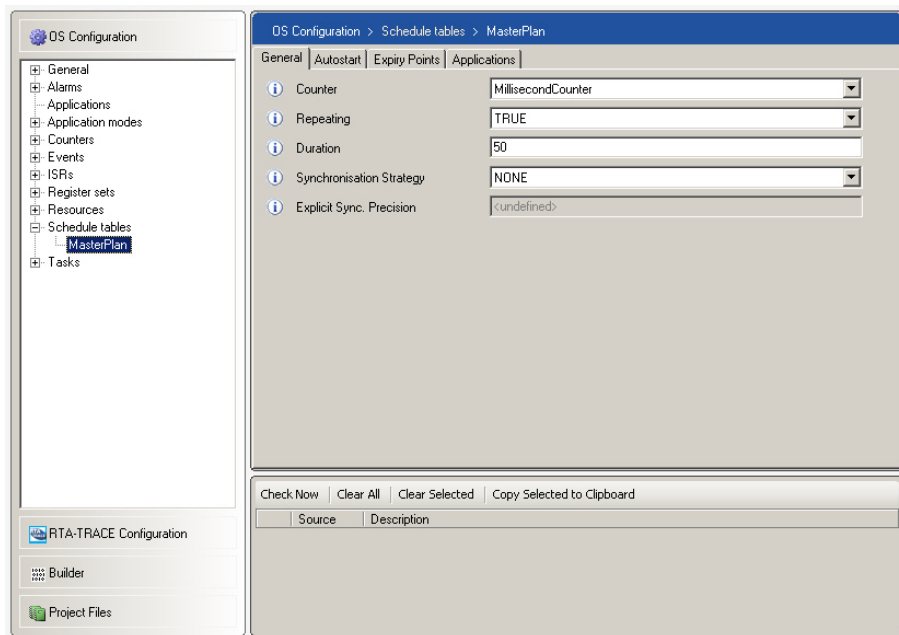


Figure 10.2: Schedule Table Configuration

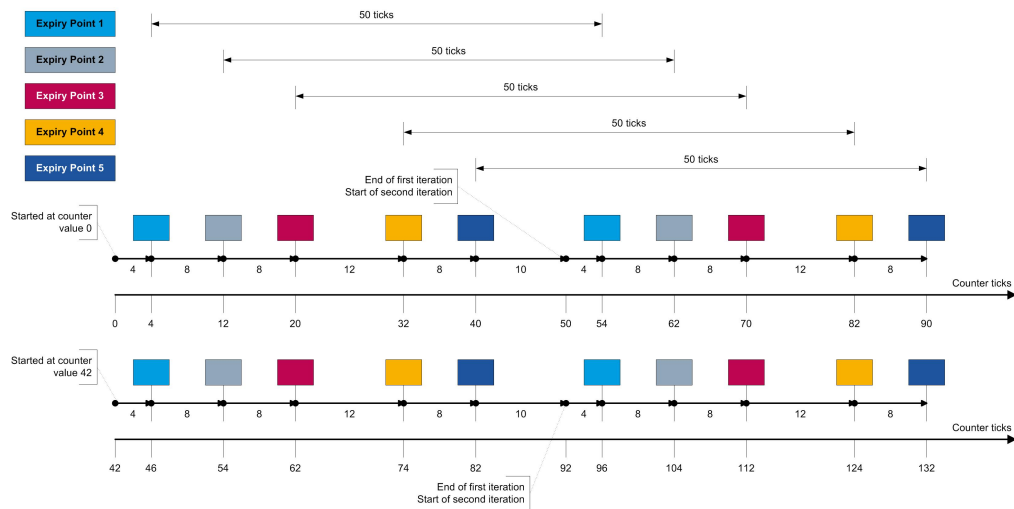


Figure 10.3: Visualizing a Schedule Table

10.1.1 Synchronization

In Figure 10.3, expiry points occur at the same relative offsets from the notional zero each time the table repeats. However, the counter tick value at which expiry points occur depends on when the schedule table was started.

It may be the case that your application requires expiry points to occur at a specific counter value. For example, the counter may be counting degrees of rotation on a crankshaft and it is essential that expiry points occur at 0°, 90°etc. Alternatively, you may want expiry points to be processed synchronously with a global time signal.

Schedule tables support these use-cases using synchronization. Three synchronization strategies are supported:

None no strategy is used. RTA-OS3.x does not do anything to ensure synchronization. This is the default behavior if no strategy is configured.

Implicit assumes that the counter used to drive the schedule table is the counter with which synchronization is required.. Section 10.9 provides additional information about implicit synchronization.

Explicit assumes that there are two counters:

- a *drive counter* which is the OS counter used to process expiry points on the schedule table.
- a *synchronization counter* which is external to the OS (i.e. it is not an OS counter object as described in Chapter 8) with which synchronization is required.

RTA-OS3.x is told the value of the synchronization counter and uses this to synchronize the schedule table.

Section 10.10 describes explicit synchronization in more detail.

10.2 Configuring Expiry Points

Each schedule table contains one or more expiry points. An expiry point marks the offset on the table where an action is required. Each expiry point has the following attributes:

- The offset from the start (the notional zero) of the schedule table.
- Zero or more tasks to activate
- Zero or more events to set for a specified task

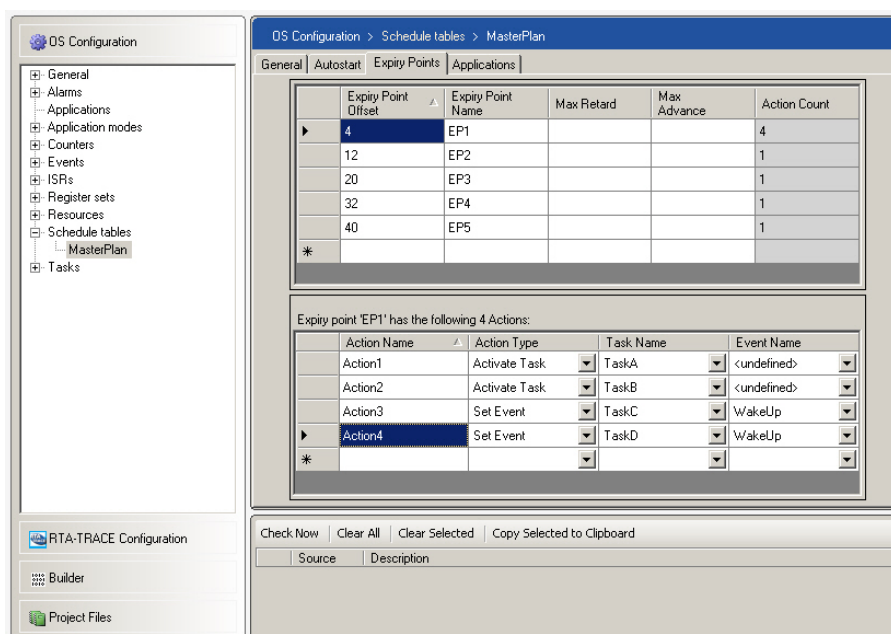


Figure 10.4: Specifying Expiry Points

Each expiry point must activate at least one task or set at least one event. It is not possible to have an expiry point with no action.

Recall that offsets are defined in ticks of the OS counter which drives the schedule table. This means that the properties of the counter place limitations on configuration.

The offset can be zero or in the range MINCYCLE to the duration of the schedule table and there must be at least MINCYCLE ticks difference between any pair of offsets².

The limitation to MINCYCLE occurs because delays of less than MINCYCLE cannot be programmed on the counter.

For example, if the counter has a MINCYCLE of 10 then:

- offsets between 1 and 9 inclusive are invalid
- offsets of 10 and 20 are valid
- offsets of 10 and 19 are invalid

Figure 10.4 shows how to specify expiry points.

²More formally, $\forall x, y \in \text{Offsets} \bullet \text{abs}(x - y) \geq \text{MINCYCLE}$

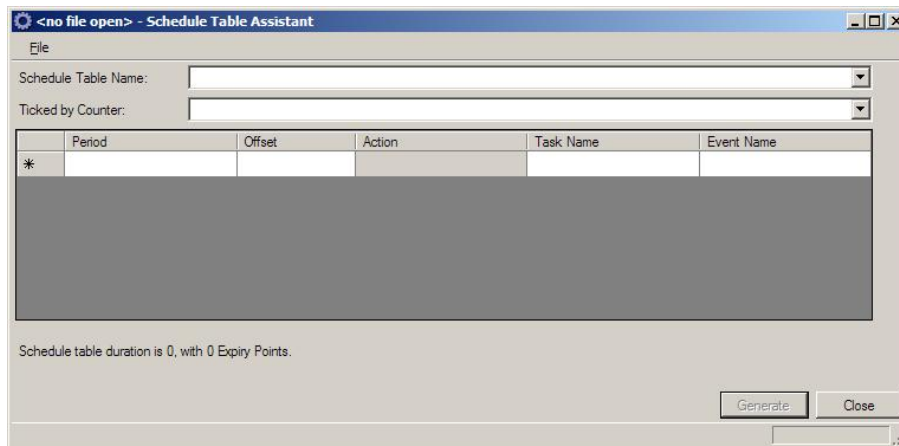


Figure 10.5: The Schedule Table Assistant

The upper part of the workspace shows the expiry points and their associated offsets. The lower part of the workspace shows the actions for the selected expiry point and the control for adding and removing expiry points.

10.3 The Schedule Table Assistant

Many applications are based around a number of periodic tasks. Since setting up a single Schedule Table with multiple (en)harmonic periods can be time consuming and error prone, **rtaoscfg** provides an assistant plug-in which can automatically generate the framework for such a system.

While the *Schedule Table Assistant* creates the necessary Schedule Table, Counter, Task and Event containers as required, it is still necessary for the end-user to complete their definitions (with Counter frequency, Task priority etc.).

The configuration of a periodic system can also be saved to disk in case future editing is required.

Figure 10.5 shows the initial view of the Schedule Table Assistant.

The data-entry grid requires that the period and offset are entered in terms of “ticks” (as specified by AUTOSAR) - the mapping of ticks to real time units is a function of the Counter driving the Schedule Table and is therefore carried out within the counter’s configuration page.

If, in addition to a period and offset, a grid entry specifies a Task name only, then the action is configured as *Activate Task*. If an Event name is specified as well, then the action becomes *Set Event*.

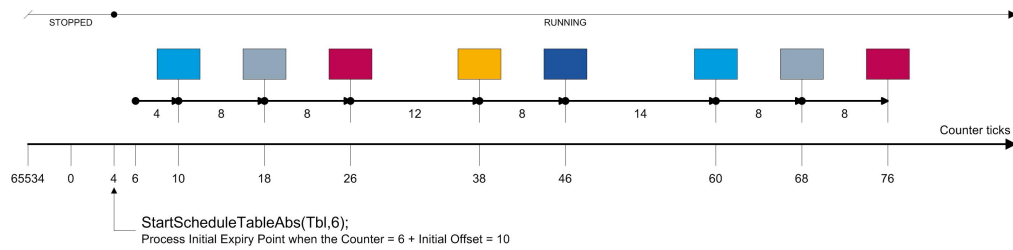


Figure 10.6: Starting a schedule table at an absolute count value

While the configuration is being entered, the message below the data-entry grid displays the length and size of the Schedule Table which to be generated.

The Assistant can overwrite a Schedule Table if one with the same name already exists, allowing a Schedule Table to be updated with new periodic information. After saving the file **File → Save** the periodic configuration can be reloaded and regenerated in the future.

The Schedule Table and supporting containers are generated when the **Generate** button is pressed - if a project is not already open, the “New Project” dialog will be presented as described in Section 3.1.1.

10.4 Starting Schedule Tables

10.4.1 Absolute Start

The `StartScheduleTableAbs(ScheduleTableID, Start)` API call is used to start a schedule table at the absolute counter value `Start` as shown in Code Example 10.2:

```
/* Start Schedule Table Tbl when the counter reaches tick 6 */
StartScheduleTableAbs(Tbl, 6);
```

Code Example 10.1: Using `StartScheduleTableAbs()`

The schedule table is in the `SCHEDULETABLE_RUNNING` state when the call returns, unless it is an implicitly synchronized schedule table, in which case it has state `SCHEDULETABLE_RUNNING_AND_SYNCHRONOUS`.

The first expiry point will be processed when `Start + Initial Offset` ticks have elapsed. Figure 10.6 shows the schedule table from Figure 10.1 when started according to Code Example 10.4.

Absolute start is extremely useful for building schedule tables that are synchronized to specific values of an external (hardware) counter. In fact, if a schedule table is configured with implicit synchronization then this is the only way to start a schedule table.

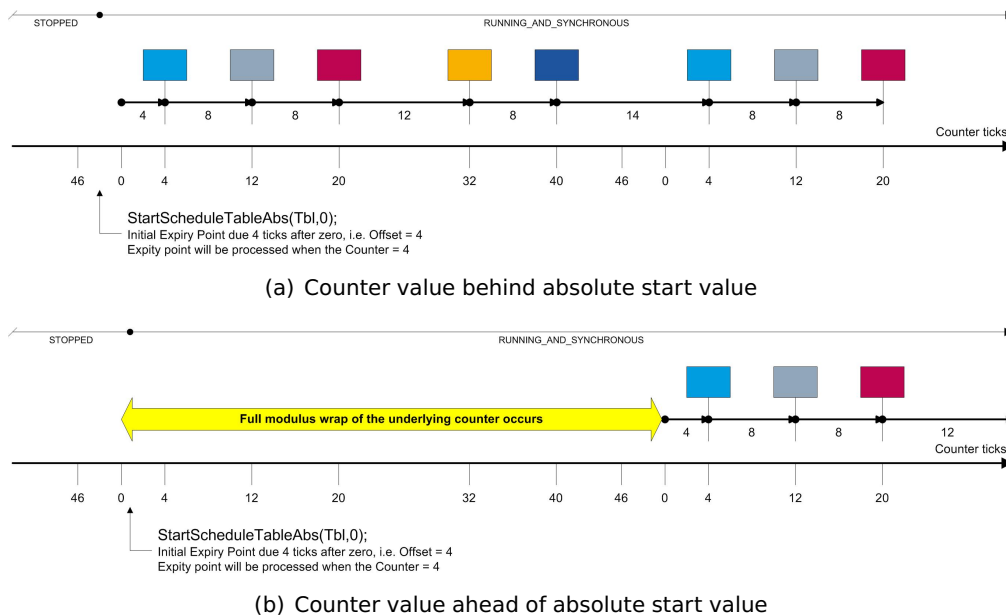


Figure 10.7: Synchronizing a schedule table with an absolute count value

Assume we have the same schedule table layout as before, but driven by a counter that has a modulus equal to the duration of the schedule table. We want the expiry point offsets to represent the *actual count value* at which the schedule table is processed. This can be achieved trivially by making an absolute start at zero as shown by Code Example 10.2:

```
/* Start ImplicitlySynchronizedTable when the counter wraps to
   zero */
StartScheduleTableAbs(ImplicitlySynchronizedTable, 0);
```

Code Example 10.2: Starting an implicitly synchronized schedule table

Figure 10.7(a) shows that the schedule table would wait until the counter reaches zero before starting and that each expiry point occurs at a counter value equal to its offset.



When starting a schedule table at an absolute count value note that the counter may have already passed the value. The schedule table will not start until the counter reaches the value again (i.e. after it has wrapped around).

This issue is most often encountered when a schedule table is started using `StartScheduleTableAbs(MyTable, 0)` immediately that RTA-OS3.x starts (or even auto-starting the schedule table at zero) because the counter value

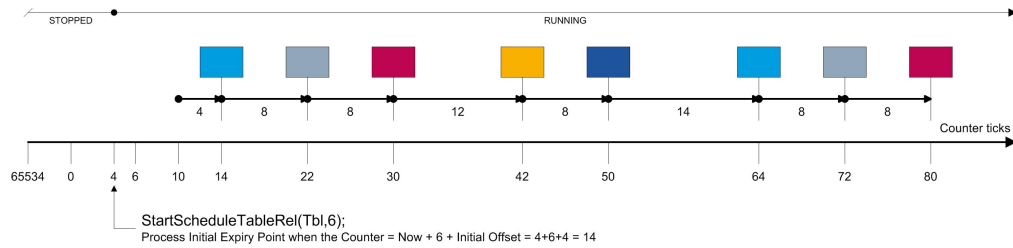


Figure 10.8: Starting a schedule table at a relative count value

of zero is already in the past. A full modulus wrap of the counter (plus the initial offset) must elapse before the first expiry point is processed³

Figure 10.7(b) shows what would happen if the schedule table shown in Figure 10.7(a) was started just after the counter reached zero.

10.4.2 Relative Start

The `StartScheduleTableRel(ScheduleTableID, Offset)` API call is used to start a schedule table at a relative number of ticks from now. This is the same concept as relative start of an alarm using `SetRelAlarm()` (see Section 9.2.2).

```
/* Start Schedule Table Tbl 6 ticks from now */
StartScheduleTableRel(Tbl, 6);
```

Code Example 10.3: Using `StartScheduleTableRel()`

The `Offset` parameter of the `StartScheduleTableRel()` call specifies the relative number of ticks from now at which RTA-OS3.x will start the schedule table.

The schedule table is in the `SCHEDULETABLE_RUNNING` state when the call returns. The first expiry point will be processed when `Offset + Initial Offset` ticks have elapsed on the counter. Figure 10.8 shows the schedule table from Figure 10.1 when started according to Code Example 10.4.



If the underlying counter is a hardware counter you must ensure that the `Offset` passed to `StartScheduleTableRel()` has not already expired before the call returns.

For schedule tables that are driven by a software counter, the counter cannot be incremented while the `StartScheduleTableRel()` is executing because both API calls execute at OS level and are thus are serialized.

³On a 32-bit counter running at 1ms you would need to wait approximately 50 days for this to happen.

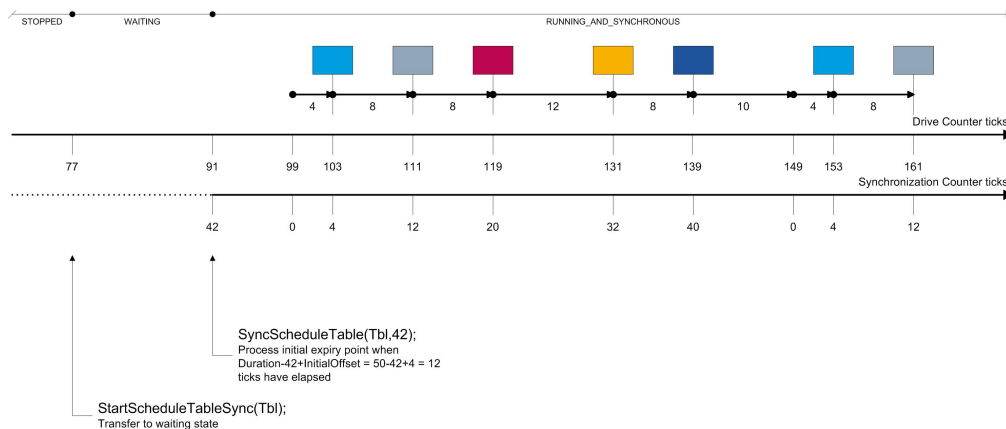


Figure 10.9: Starting a schedule table synchronously

Relative start is prohibited for schedule tables that use implicit synchronization, so the following code will generate an error:

```
StartScheduleTableRel(ImplicitlySynchronizedTable, 6); /* returns
E_OS_ID */
```

Code Example 10.4: Forbidden use of StartScheduleTableRel()

10.4.3 Synchronized Start

The StartScheduleTableSynchron(ScheduleTableID) API call is used to start a schedule table synchronously after RTA-OS3.x has been told the value of the synchronization counter. It is only possible to use this API for schedule tables with explicit synchronization.

```
/* Start Schedule Table when the synchronization count is
   provided */
StartScheduleTableSynchron(ExplicitlySynchronizedTable);
/* Table now waits */
...
SyncScheduleTable(ExplicitlySynchronizedTable,42);
/* Table processes first expiry point Duration-42+InitialOffset
   ticks from now */
```

Code Example 10.5: Using StartScheduleTableSynchron()

Figure 10.9 shows what happens when a schedule table is started according to Code Example 10.4.

The schedule table is in the SCHEDULETABLE_WAITING state when the call returns.

StartScheduleTableSynchron() assumes that the notional zero of the schedule table needs to be synchronized with the absolute value zero on the synchronization counter.

When a schedule table is started by calling `StartScheduleTableSynchron()` it will not run until `SyncScheduleTable(ScheduleTableID, Value)` is subsequently called. The first expiry point will then be processed when `Duration - Value + Initial Offset` ticks have elapsed on the drive counter.



If a `SyncScheduleTable()` is never executed then the table will not leave the `SCHEDULETABLE_WAITING` state unless it is stopped using `StopScheduleTable()`.

10.5 Expiry Point Processing

Expiry points are processed in response to an `IncrementCounter()` or `Os_AdvanceCounter()` API call being made on the counter driving the schedule table.

When the counter reaches the value at which the next expiry point is due, RTA-OS3.x will:

- activate all configured tasks, in descending priority order; then
- set all events for tasks, again in descending priority order; then
- set up the match value on the underlying counter at which the next expiry point is due.

10.6 Stopping Schedule Tables

A schedule table can be stopped at any point by calling the `StopScheduleTable(ScheduleID)` API. Schedule table processing stops immediately.

A repeating schedule table will run until it is stopped by calling `StopScheduleTable()` or the table is switched by calling `NextScheduleTable()` (see Section 10.7).

A schedule table that is configured as single shot (the repeating attribute is false) will stop automatically when `Final Delay` ticks have elapsed after RTA-OS3.x has processed the final expiry point as shown in Figure 10.10.

10.6.1 Restarting a stopped schedule table

You can re-start a schedule table that has been stopped by calling `StartScheduleTable[Abs|Rel|Synchron]()`. The schedule table will re-start at its notional zero as shown in Figure 10.11. There is no mechanism for starting a schedule table part-way through.

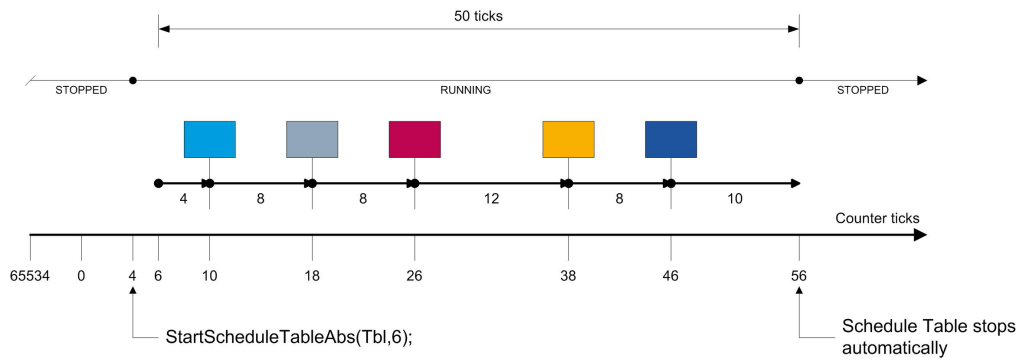


Figure 10.10: Non-repeating schedule tables stop automatically

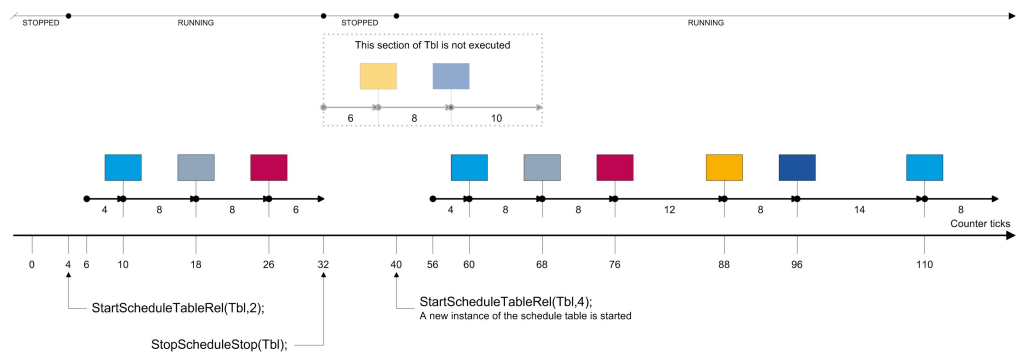


Figure 10.11: Schedule tables always start at their notional zero

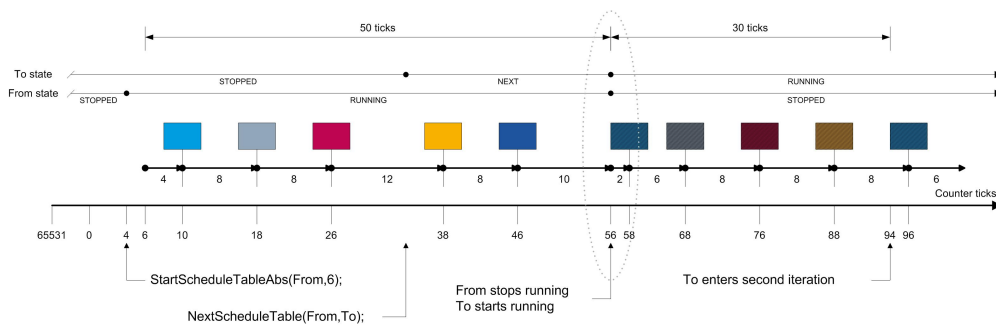


Figure 10.12: Switching between schedule tables

10.7 Switching Schedule Tables

You can switch from one schedule table to another at runtime using the `NextScheduleTable()` API call. The switch between schedule tables always occurs at the end of the table - i.e. final delay ticks after the final expiry point is processed. Code Example 10.6 shows how the API call is made.

```
/* Start To after From has finished */
NextScheduleTable(From, To);
```

Code Example 10.6: Switching a schedule table

It is only valid to switch between schedule tables that are driven by the same counter. The schedule table in the switch do not need to have the same number of expiry points, the same duration, the same initial offset, etc.

When a call to `NextScheduleTable()` is made, the delay between the final expiry point on `From` and the first expiry point on `To` is given by:

$$\text{Delay} = \text{From.FinalDelay} + \text{To.InitialOffset}$$



If the Current schedule table has a final delay of zero ticks and the Next schedule table has an initial offset zero, then the delay between expiry points will be zero ticks.

Figure 10.12 shows the process of switching from one schedule table `From` with duration 50 ticks to schedule table `To` with duration 30 ticks.

If you make multiple calls to `NextScheduleTable()` while `From` is running then the next table that runs will be the one you specified in your most recent call.



If you stop the `From` schedule table then the `To` schedule table will remain in the state `SCHEDULETABLE_NEXT` indefinitely.

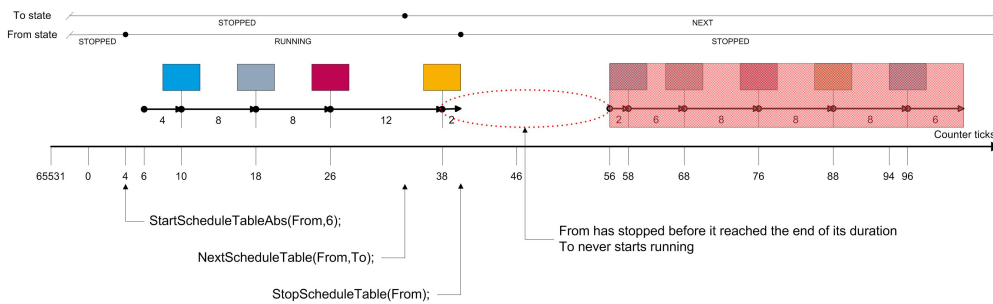


Figure 10.13: Leaving a schedule table in the SCHEDULETABLE_NEXT state

Figure 10.13 shows the effect, where schedule table From is stopped before the transition to schedule table To occurs, leaving To in the SCHEDULETABLE_NEXT state.

A side-effect of this problem is that the application will not be able to run schedule table To because it is not possible to start a schedule table that is not in SCHEDULETABLE_STOPPED. The problem can be fixed at runtime by checking if the table is waiting and then calling StopScheduleTable(To) to reset its state to SCHEDULETABLE_STOPPED. Code Example 10.7 shows how this might be achieved.

```
ScheduleTableStatusType FromState,ToState;
GetScheduleTableStatus(From, &FromState);
GetScheduleTableStatus(To, &ToState);
if ( (ToState == SCHEDULETABLE_NEXT) &&
      (FromState == SCHEDULETABLE_STOPPED) {
    /* Reset state of To */
    StopScheduleTable(To);
}
StartScheduleTableAbs(To,99);
```

Code Example 10.7: Resetting the state of a schedule table

Note that to avoid race conditions between reading and evaluating the state conditions, it is essential that any code like this executes at the same priority as the highest priority OS object that manipulates either of the two schedule tables.

10.8 Choosing an Synchronization Strategy



If you need to synchronize a schedule table then you are strongly advised to use implicit synchronization as this model correctly decouples the OS from the counter driver.

This is especially true when synchronizing with a global time provided by a time-triggered networking technology like FlexRay or TTP. The technologies use hardware to perform clock synchronization across a distributed network

and provide the current view of global time to the CPU. Even if the hardware loses synchronization on the network, it will still provide the CPU with a best-effort global time reference from which a schedule table can be driven and re-synchronization will be attempted automatically. These hardware solutions fix the clock synchronization problem in a far more efficient and accurate way that you can hope to achieve in software. You should therefore not re-introduce these problems in your application by using AUTOSAR OS's explicit synchronization model.

The only legitimate use-case for explicit synchronization is when you need to synchronize processing in the OS with a sporadically generated and inaccurate (i.e. subject to large amounts of drift) clock source, for example a global time signal broadcast over a CAN network.

10.9 Implicit Synchronization

Implicit synchronization requires that:

1. The schedule table and the OS counter must wrap at the same value. This guarantees that an offset of X ticks from zero on the table can be aligned exactly with tick X of the counter.
This is configured by setting the duration equal to MAXALLOWEDVALUE+1 of the counter.
2. The schedule table is only started at a known counter value. This means that the table can only be started using StartScheduleTableAbs().

RTA-OS3.x does not do anything to maintain synchronization as this is the responsibility of the counter driver.

The following table shows the differences between the 'implicit' and 'none' synchronization strategies:

	None	Implicit
Minimum Duration	Mincycle	MaxAllowedValue+1
Maximum Duration	MaxAllowedValue+1	MaxAllowedValue+1
Start...Abs()	✓	✓
Start...Rel()	✓	✗
Start...Synchron()	✗	✗

Note that a schedule table that uses the 'none' synchronization strategy, has a duration equal to MAXALLOWEDVALUE+1 and is always started using StartScheduleTableAbs() will behave identically at runtime to a schedule table using the 'implicit' synchronization strategy.

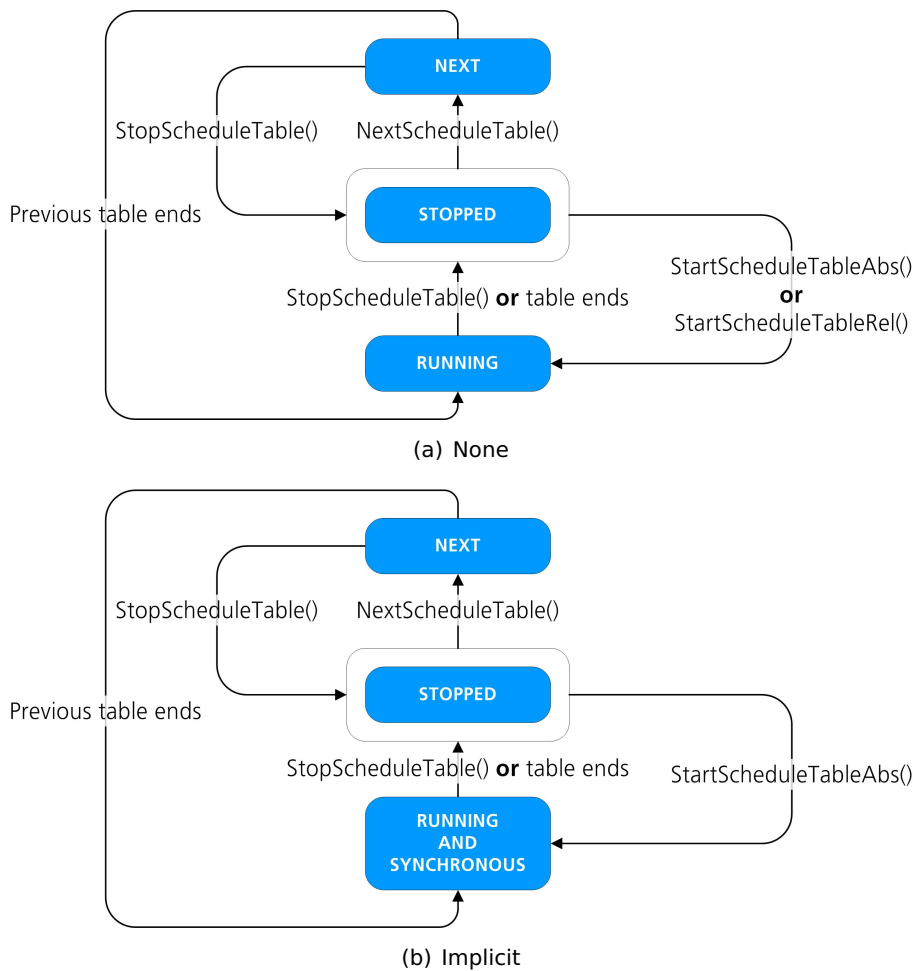


Figure 10.14: State transitions for none and implicitly synchronized schedule tables

The 'implicit' synchronization strategy differs from the 'none' synchronization strategy only in that 'implicit' forces RTA-OS3.x to check that the constraints are not violated using additional configuration-time and run-time checks.

Figure 10.14 shows how similar the state models are, with the exceptions that `StartScheduleTableRel()` cannot be used and the `SCHEDULETABLE_RUNNING` state used for the 'none' synchronization strategy is simply renamed `SCHEDULETABLE_RUNNING_AND_SYNCHRONOUS` in the 'implicit' synchronization strategy.

Implicit synchronization is very efficient when integrated with a hardware counter. When each expiry point is processed, RTA-OS3.x uses the hardware counter callback `Os_Cbk_Set_<CounterID>` to program the counter to interrupt when the delay to the next expiry point elapses. Any re-synchronization

occurs in the scope of the driver itself and the schedule table always remains synchronized with the counter.

Additional information about writing hardware counter drivers is provided in Chapter 11.

10.10 Explicit Synchronization

Recall from Section 10.1.1 that explicit synchronization requires two counters:

- a *drive counter* that is used to process expiry points on the schedule table.
- a *synchronization counter* that is external to the OS (i.e. it is not an OS counter object) with which synchronization is required.

With none and explicit synchronization, the delays between expiry points are static.

With explicit synchronization RTA-OS3.x adjusts the delay between adjacent expiry points dynamically (within configurable bounds) in order to keep the schedule table synchronized with the synchronization counter. The delay is extended if the schedule table is ahead of the synchronization counter and the delay is reduced if the schedule table is behind of the synchronization counter.

10.10.1 Counter Constraints

Explicit synchronization in AUTOSAR OS assumes that the following constraints between the schedule table, the drive counter and the synchronization counter:

1. the schedule table duration is no greater than the modulus of the drive counter.
2. the schedule table duration is equal to the modulus of the synchronization counter.
3. the drive counter and the synchronization counter have the same resolution - one tick of the drive counter must take the same time as one tick of the synchronization counter.

The first constraint can be checked by RTA-OS3.x because the drive counter is an OS object. The remaining constraints cannot be checked by RTA-OS3.x because the synchronization counter is not visible as a configurable object.

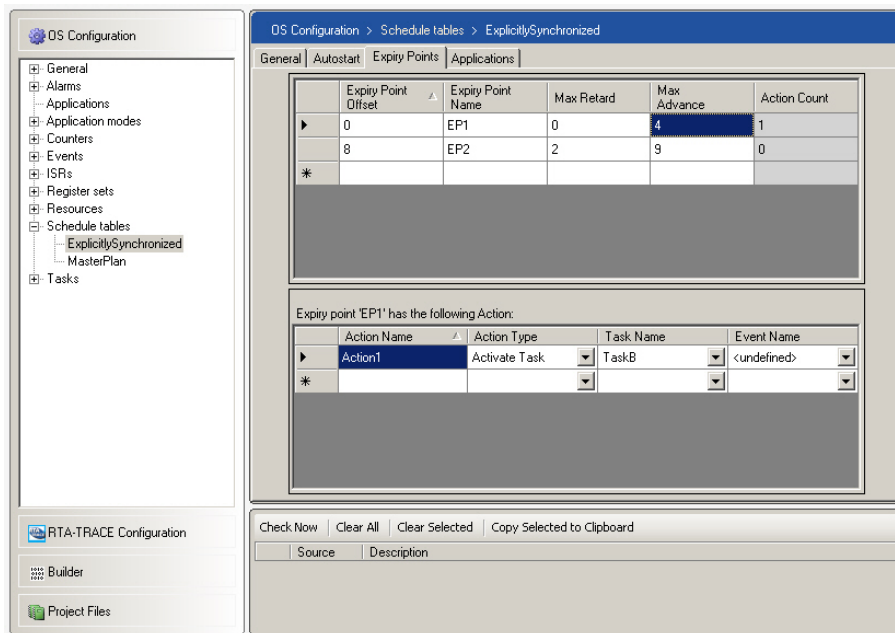


Figure 10.15: Specifying Max Advance and Max Retard for Expiry Points



It is your responsibility to check that your synchronization counter satisfies the started constraints. Explicit synchronization will not work if the constraints are not satisfied.

10.10.2 Specifying Synchronization Bounds

Adjustment Range

The amount by which RTA-OS3.x can adjust the delay between expiry points is statically configured using two parameters:

Max Retard defines the maximum number of ticks that can be *subtracted* from the expiry point's offset.

This is equivalent to shortening the delay from the previous expiry point by the specified number of ticks.

Max Advance defines the maximum number of ticks that can be *added* to the expiry point's offset.

This is equivalent to lengthening the delay from the previous expiry point by the specified number of ticks.

Figure 10.15 shows how the synchronization bounds are specified for an expiry point.

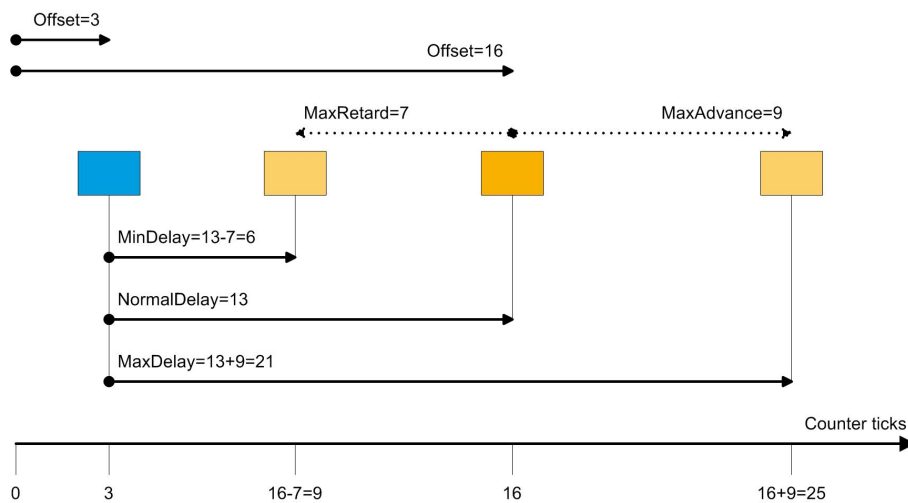


Figure 10.16: Impact of Max Retard and Max Advance on delays

Figure 10.16 shows the impact of these parameters on the offsets to an expiry point and the delays from the previous expiry point.

Each expiry point can specify its own adjustment values. If the adjustment values for an expiry point are not specified, then no adjustment will be made at runtime.

In common with other schedule table configuration parameters, the permitted values are limited by the attributes of the drive counter. The minimum delay between expiry points must be greater than MINCYCLE. This is guaranteed when the following condition is true for every pair of expiry points (i,j):

$$\text{Offset}_i - (\text{Offset}_j - \text{MaxRetard}_j) \geq \text{MINCYCLE}$$

Note that all adjustment calculations are evaluated modulo the duration of the schedule table. The initial expiry point can therefore specify a Max Retard value greater its offset, providing that the result of the modulus subtraction is greater than the value of final expiry point's offset plus MINCYCLE.

Figure 10.17 shows how the delay between two adjacent expiry points can be reduced from 3 ticks to 1 tick using Max Retard of 2 ticks.

The maximum delay between two adjacent expiry points must be less than the schedule table's duration:

$$\text{Offset}_i + \text{MaxAdvance}_i \leq \text{DURATION}$$

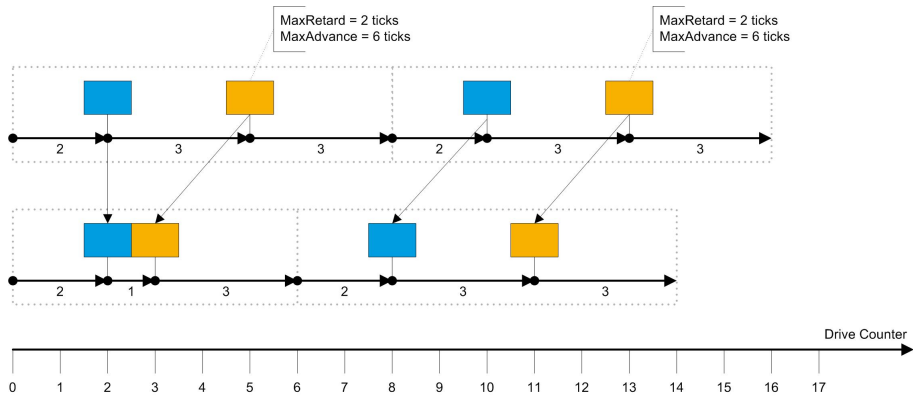


Figure 10.17: Minimizing the delay between expiry points using max retard

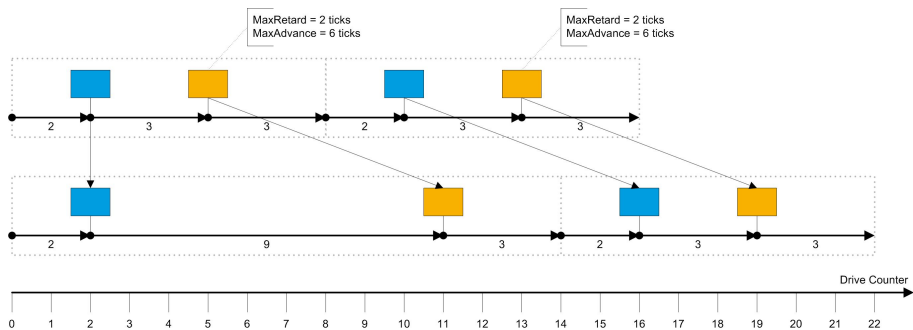


Figure 10.18: Maximizing the delay between expiry points using max advance

Figure 10.18 shows how the delay between two adjacent expiry points can be increased from 3 ticks to 9 ticks using max advance equal to 6 ticks.

Deviation

The schedule table is driven by a different counter to the one with which synchronization is required so it is possible for the schedule table to drift with respect to the synchronization counter. In AUTOSAR OS the drift is called the *deviation*.

However, explicitly synchronized schedule tables are constrained so that the synchronization counter and the schedule table have the same modulus. RTA-OS3.x can therefore use the value of the synchronization counter to calculate the deviation between the current point on the schedule table and the synchronization counter.

The deviation is the smallest number of ticks that must be added to (or subtracted from) the current value 'now' value of the schedule table to equal the value of the synchronization counter, modulo the duration of the schedule table.

Deviation is defined to be:

positive if ticks must be *added* to the 'now' value of the schedule table. This means that synchronization counter is running *ahead* the schedule table and delays will need to be reduced. Shortened delays make expiry points happen earlier - in effect "speeding up" the schedule table.

negative if ticks must be *subtracted* from the 'now' value of the schedule table. This means synchronization counter is running *behind* the schedule table and the delays will need to be extended. Extending delays make expiry points happen later - in effect "slowing down" the schedule table.

Figure 10.19 shows the deviations that are possible for each synchronization counter value with a counter modulus of 8 where the schedule table has a 'now' value of 7 ticks.

The red triangle shows the number of ticks that must be subtracted from 'now' to equal the value of the synchronization counter. The green triangle shows the number of ticks that must be added to 'now' to equal the synchronization counter. The deviation is given by the smallest triangle. Note that if adding and subtracting give the same deviation, then it doesn't matter whether we choose to add or subtract.

RTA-OS3.x assumes the deviation is positive if the value is less than or equal to half the duration of the schedule table.

The following table shows how the deviation values for any schedule table value and synchronization counter value for a modulus (duration) of 8 ticks.

		Schedule Table Now							
		0	1	2	3	4	5	6	7
Sync. Count	0	0	-1	-2	-3	+4	+3	+2	+1
	1	+1	0	-1	-2	-3	+4	+3	+2
	2	+2	+1	0	-1	-2	+3	+4	+3
	3	+3	+2	+1	0	-1	-2	-3	+4
	4	-4	+3	+2	+1	0	-1	-2	-3
	5	-3	-4	+3	+2	+1	0	-1	-2
	6	-2	-3	-4	+3	+2	+1	0	-1
	7	-1	-2	-3	-4	+3	+2	+1	0

10.10.3 Startup

It is common to start an explicitly synchronized schedule table with the `StartScheduleTableSynchron()` API call described in Section 10.4.3. The schedule table will wait for the synchronization counter to be provided (using

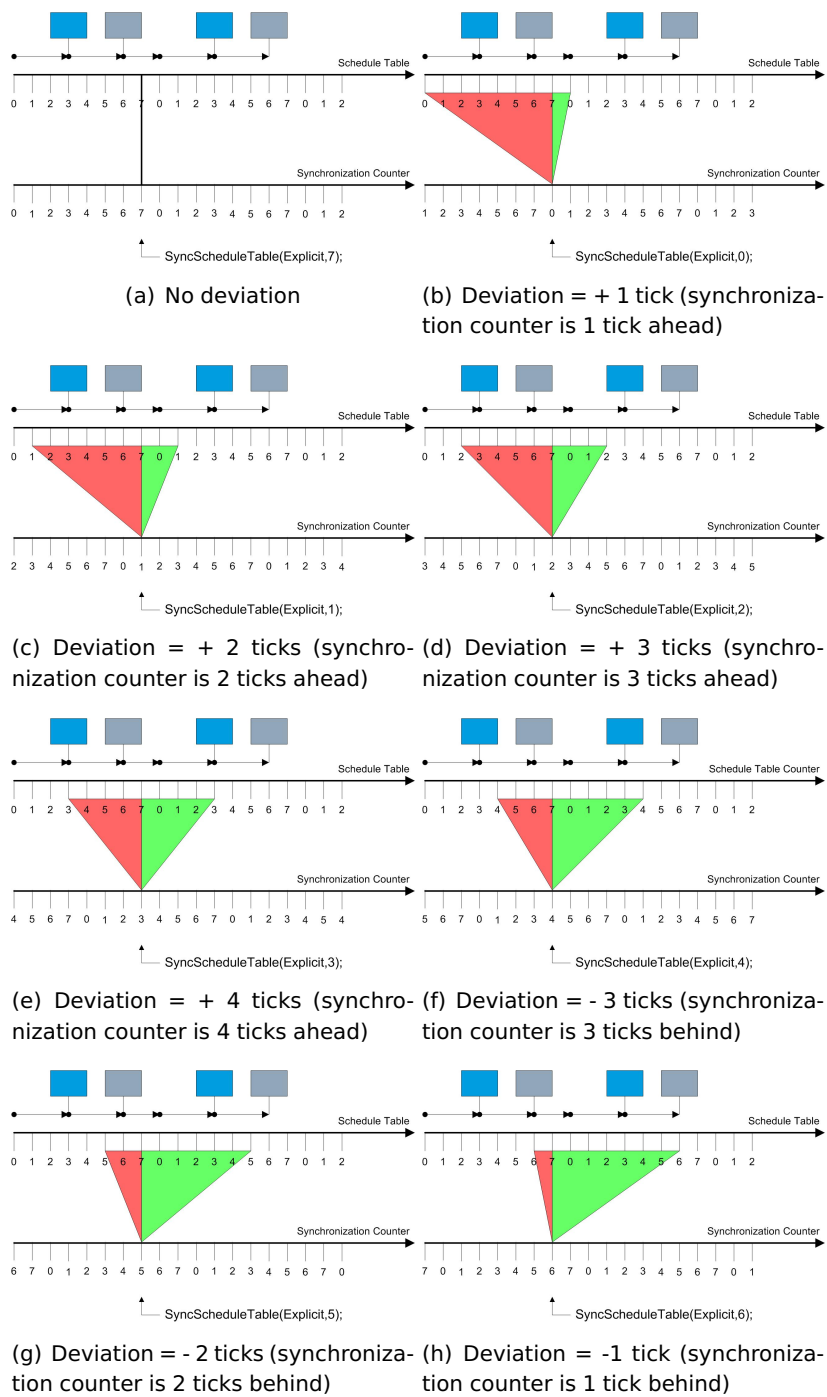


Figure 10.19: Schedule table deviations

the `SyncScheduleTable()` call described in Section 10.10.4). This provides what is known as *hard* startup. This model means that the schedule table only starts once a synchronization is established.

However, it may be the case that a schedule table needs to run and process expiry points before a synchronization count is provided i.e. it needs to run asynchronously. When a synchronization count is provided, then RTA-OS3.x can synchronize the schedule table.

Asynchronous start is trivially done using a `StartScheduleTableAbs()` or `StartScheduleTableRel()` API call.



The schedule table runs asynchronously with the (as yet unknown) synchronization count for an indefinite amount of time. During this time, there is no difference between the 'explicit' and 'none' synchronization strategies. If your system can run safely for an indefinite time without synchronization, then it follows that it must be able to run safely all the time without synchronization. Using explicit synchronization in this case is unnecessary as a synchronization strategy of 'none' would suffice and would have a lower memory footprint and higher performance.

10.10.4 Synchronizing the Schedule Table

Whenever an expiry point is processed, RTA-OS3.x calculates the delay to the next expiry point at runtime.

The `SyncScheduleTable(Tbl, Value)` API call is used to tell RTA-OS3.x the current value of the synchronization counter and additionally that synchronization should start (if not already being performed).

There are two parts to synchronization:

1. The application calls `SyncScheduleTable(Tbl, Value)` and RTA-OS3.x calculates the current deviation between the schedule table and the synchronization counter.
2. Every time an expiry point is processed, RTA-OS3.x uses the calculated deviation and the adjustment limits (Max Retard and Max Advance) to calculate the delay to the next expiry point required to maintain (or gain) synchronization.

If `SyncScheduleTable(Tbl, Value)` is called multiple times before the next expiry point is processed then only the most recently calculated values will be used.

Expiry Point Adjustment

If a synchronization count has not yet been provided, expiry points are processed as described in Section 10.5.

When a synchronization count is provided, each time an expiry point is processed, RTA-OS3.x uses the currently stored deviation to adjust the delay to the next expiry point within the limits specified by max retard of the next expiry point and max advance of the current expiry point.

The adjustment made depends on the sign of the deviation:

Negative deviations occur when the schedule table is running behind the synchronization counter and needs to 'catch up'.

Catching up is done by reducing the time to next expiry point. The maximum value of the reduction is the minimum of the deviation and the max retard of the next expiry point, i.e. the reduction is limited by the max retard value.

Positive deviations occur when the schedule table is running ahead of the synchronization counter and needs to 'slow down'.

Slowing down is done by increasing the time to next expiry point. The maximum value of the increase is the minimum of the deviation and the max advance of the current expiry point, i.e. the increase is limited by the max retard value.

The amount of adjustment made is subtracted from the deviation and the delay to the expiry point is processed.

If the deviation is non-zero when the next expiry point is processed then a further adjustment is made.

Figure 10.20(a) shows a schedule table and the associated max retard and max advance parameters for each expiry point. The effect of synchronization requests triggered by calls to `SyncScheduleTable()` on the delays between expiry points is shown in Figure 10.20(b)

10.10.5 Synchronization

An explicitly synchronized schedule table is said to be synchronous (with the synchronization counter) if the deviation is not more than a specified *precision*. If you configure a schedule table to use the 'explicit' synchronization strategy then it is mandatory to specify a precision. The precision can be set to any value in the range 0 to the duration.

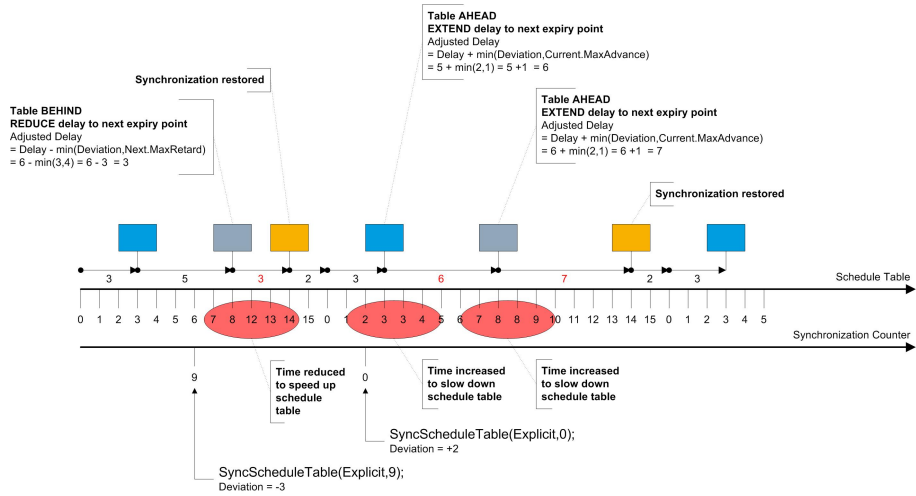
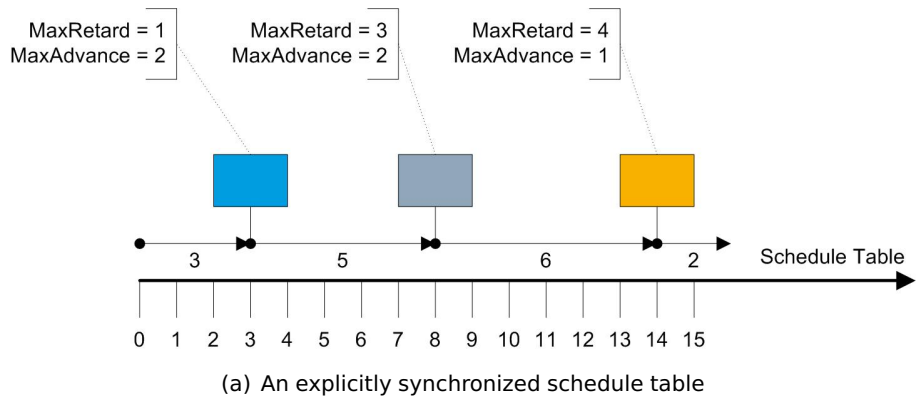


Figure 10.20: Synchronizing a Schedule Table

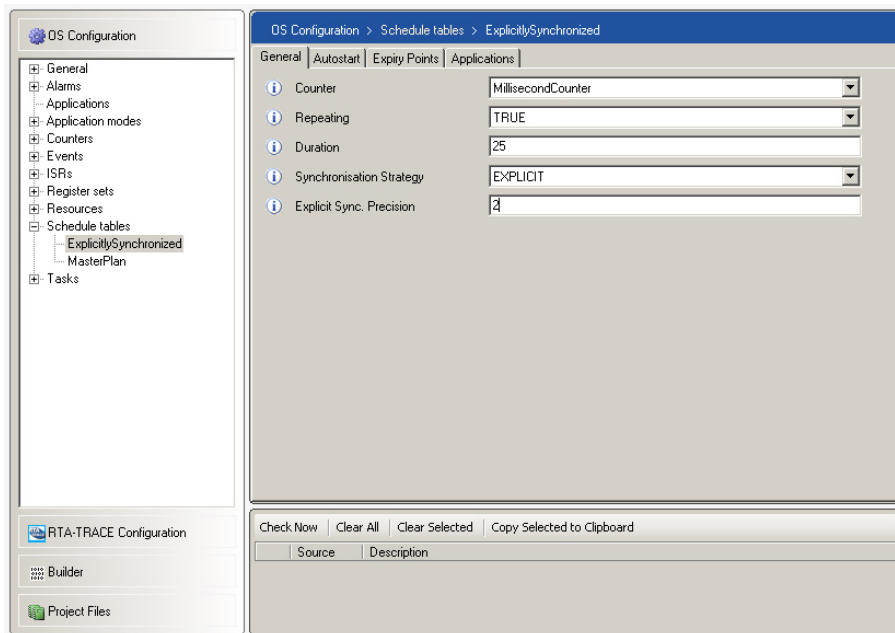


Figure 10.21: Specifying the precision

Figure 10.21 shows the configuration of an explicitly synchronized schedule table with a precision of 2 ticks on the underlying counter (in this case two milliseconds).

The following table shows the deviation values you saw in Section 10.10.2. Assuming a precision of 2, the shaded cells show under which deviations the schedule table be considered synchronous.

		Schedule Table Now								
		0	1	2	3	4	5	6	7	
Sync. Count	0	0	+1	+2	+3	+4	-3	-2	-1	
	1	+1	0	+1	+2	+3	+4	-3	-2	
	2	+2	+1	0	+1	+2	+3	+4	-3	
	3	+3	+2	+1	0	+1	+2	+3	+4	
	4	+4	+3	+2	+1	0	+1	+2	+3	
	5	-3	+4	+3	+2	+1	0	+1	+2	
	6	-2	-3	+4	+3	+2	+1	0	+1	
	7	-1	-2	-3	+4	+3	+2	+1	0	

A precision of zero means that the deviation must be zero for the schedule table to be synchronous. A precision of duration/2 or greater will mean that the table is always synchronous. This is because the deviation is calculated modulo the duration, so the schedule table and the synchronization counter can never be more than half the duration out of phase.

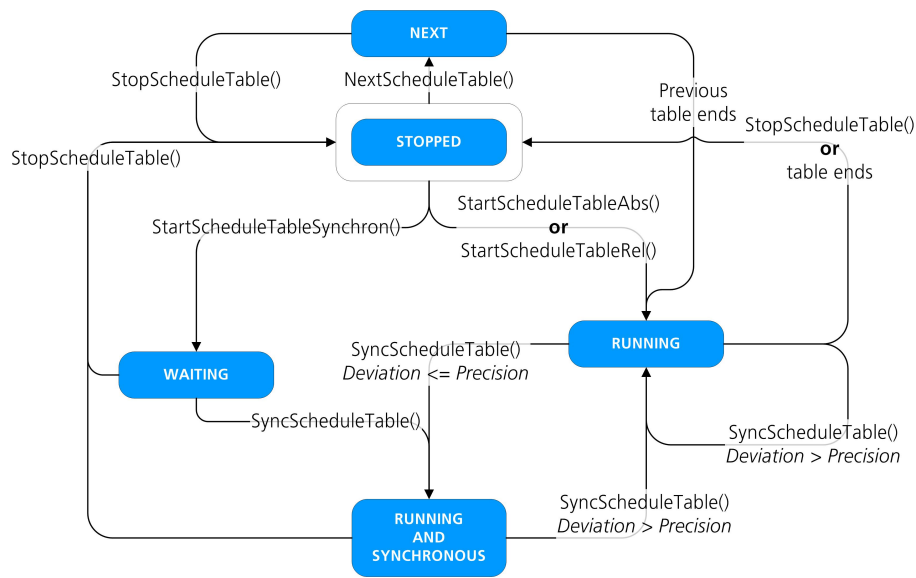


Figure 10.22: Explicit Synchronization States

The state of a running explicitly synchronized schedule table is either:

SCHEDULETABLE_RUNNING_AND_SYNCHRONOUS if the schedule table is running and deviation \leq precision

SCHEDULETABLE_RUNNING if the schedule table is running and deviation $>$ precision

An explicitly synchronized schedule table will alternate between these two states at runtime.

Figure 10.22 shows the state transition model for an explicitly synchronized schedule table.

10.10.6 Disabling Synchronization

Synchronization of an explicitly synchronized schedule table can be disabled using `SetScheduleTableAsync()`. When the call is made, the schedule table continues to run and expiry points are processed, but RTA-OS3.x stops any expiry point adjustment.

Synchronization can be re-started by calling `SyncScheduleTable()`.

Code Example 10.8 shows examples of these calls.

```

SetScheduleTableAsync(Explicit);
...

```

```
SyncScheduleTable(Explicit,0);
```

Code Example 10.8: Stopping and restarting synchronization

10.11 Schedule Table Status

You can query the state of a schedule table using the `GetScheduleTableStatus()` API call. The call returns the status through an out parameter. Code example 10.9 shows how to get the status.

```
ScheduleTableStatusType State;  
GetScheduleTableStatus(Table, &State);
```

Code Example 10.9: Getting the status of a schedule table

The status will be either:

SCHEDULETABLE_STOPPED if the table is not started.

SCHEDULETABLE_RUNNING if the schedule table is started.

SCHEDULETABLE_RUNNING_AND_SYNCHRONOUS if the schedule table is started and synchronized.

SCHEDULETABLE_NEXT if the schedule table has been started by a call to `ScheduleTableNext()` but is not yet running (because another schedule table on the same counter has not yet finished).

SCHEDULETABLE_WAITING if the schedule table has been started by a call to `StartScheduleTableSynchron()` but a call to `SyncScheduleTable()` has not been made.

10.12 Summary

- Schedule tables provide a way of planning a series of actions statically at configuration time.
- A schedule table is associated with exactly one OSEK counter, may specify a duration, and contains one or more expiry points.
- Expiry points in RTA-OS3.x are created implicitly by specifying offsets for stimuli implemented on a schedule table.
- You can switch between schedule tables, but only at the notional end of the table.
- Schedule tables can be synchronized with a global tick source using either:

- implicit synchronization, where counter driving the schedule table is the global time source
- explicit synchronization, where counter driving the schedule table is ticked by a local counter and the table is synchronized manually by telling the OS the current value of the global tick source.

11 Writing Hardware Counter Drivers

You have seen that RTA-OS3.x provides a simple, elegant and powerful interface for driving counters. The hardware counter driver mechanism provides great flexibility by placing the software/hardware interaction in the domain of user-supplied code. This allows easy integration of drivers for novel hardware and application requirements, and the ability to “piggyback” driver operation on hardware that is also used for other functions.

As the owner of your hardware you know how you want to use it in your application and therefore you are responsible for providing the hardware counter driver functions.

This chapter offers some guidelines to help you in the construction of hardware counter drivers. Much of this knowledge has been gained while constructing drivers for assorted peripheral timers, but it should be applicable to other peripherals which increment in response to some external event (e.g. interrupts generated by the rotation of a toothed wheel).

The example code is structured for ease of explanation and understanding. Different control structures may result in small improvements in the quality of generated code on some targets (e.g. replacing a **while(1)** loop using **if . break** exits with a **do . while** loop with appropriately modified conditions). If you choose to make this type of optimization, then you should take care to ensure that the required semantics and orderings of operations are maintained (e.g. note that the “&&” logical and operator in C imposes both ordering and lazy evaluation).

11.1 The Hardware Counter Driver Model

The hardware driver concept assumes an underlying free-running peripheral counter. The counter has an initial value established by the user, counts up from zero and wraps back to zero as it reaches its modulus.



These are the assumptions of the model. In later sections of this chapter you will see how to implement this model with hardware which does not require these constraints.

A hardware counter driver uses the `0s_AdvanceCounter_<CounterID>()` API call to tell RTA-OS3.x to expire an alarm and/or the schedule table expiry points associated with a counter as soon as possible after it/they become due and to program the next alarm or expiry point.

In this chapter we use the terms:

now is the counter’s current (continuously increasing) value.

old is the previously programmed compare value.

match is the (absolute value of the) count at which the next alarm or schedule table expiry point is due.

- is a binary subtraction modulo the counter's modulus.

The code examples in this chapter make use of the functions:

```
clear_pending_interrupt()  
set_pending_interrupt()  
enable_interrupt_source()  
disable_interrupt_source()
```

These functions refer to operations performed on the status/control registers of the counter peripheral used to provide the hardware counter functionality. You are responsible for providing these functions (or equivalent code) in your hardware drivers.

11.1.1.1 Interrupt Service Routine

Typically you will call the hardware driver interface of RTA-OS3.x from a user-supplied Category 2 ISR.

The ISR is triggered by each **match** and will call `Os_AdvanceCounter_<CounterID>()` to tell RTA-OS3.x that a **match** has occurred. RTA-OS3.x will then setup the delay until the next **match**. In general, there are three classes of behavior the ISR. These are described here, along with their implications for system behavior and schedulability analysis, in order that appropriate choices can be made when implementing the ISR component of hardware counter drivers.

Simple handlers can deal with a single **match** value being processed per ISR. This class of handler must complete before the next interrupt becomes due.

Re-triggering handlers can deal with one or more matches becoming due before it completes handling of the interrupt by which it was first triggered. Such a handler processes one **match** per invocation, and exits with the invoking interrupt still pending if another **match** is already due.

Looping handlers can deal with one or more matches becoming due before it completes handling of the interrupt which first triggers it. Such a handler is able to process multiple expiries in turn, and only exits when either no **match** is due or when an interrupt is pending. Any interrupt handler which is capable of looping is a looping handler.

When it can be guaranteed the handler can complete before the next **match** becomes due then a simple handler is the best choice because they typically have a smaller worst-case execution time than re-triggering or looping handlers. The choice between re-triggering and looping is influenced by the following factors:

1. Some hardware will not support re-triggering behavior, so a looping handler must be used.
2. When the interrupt that invokes the handler is at the same level as another interrupt in the system, and that other interrupt has a higher arbitration precedence (i.e. will be handled first if both are pending) then a re-triggering handler is preferred because it reduces latency for the other interrupt. In practice, this is of particular concern for architectures with a single interrupt priority level.
3. A re-triggering handler typically has smaller execution time than a looping handler when a single **match** is processed. Note that it is not normally relevant that a looping handler may be “more efficient” when several expiries are handled in one invocation. Worst case behavior occurs when each **match** is handled by a separate invocation.

A simple handler is recommended if the handler’s worst case response time (i.e. the time between the interrupt becoming ready and the handler terminating) is known to be smaller than the minimum interval between interrupts. If this cannot be guaranteed, then a re-triggering handler should be used unless the hardware characteristics prohibit it.

11.1.2 Callbacks

Recall from 8.2.2 that four callbacks are also required as part of the hardware counter driver:

1. `Os_Cbk_Now_<CounterID>()` which must return the **now** value of the peripheral counter.
2. `Os_Cbk_Cancel_<CounterID>()` which clears any pending interrupt for the counter and ensures that the interrupt will not become pending until after a `Os_Cbk_Set_<CounterID>()` call has been made. This behavior is required if any of the following conditions apply to your application:
 - (a) the alarms driven by the counter are never stopped directly by the application calling `CancelAlarm();`
 - (b) the schedule tables driven by the counter are stopped directly by the application calling `StopScheduleTable();`

(c) you have a schedule table that does not repeat (in this case RTA-OS3.x may need to cancel the interrupt when the schedule table stops).

If none of these conditions apply, then you can simply provide a 'stub' call to implement `Os_Cbk_Cancel_<CounterID>`.

3. `Os_Cbk_State_<CounterID>()` is called by RTA-OS3.x when `GetAlarm()` or `GetScheduleTableStatus()` is called by your application code and the relevant alarm or schedule table is running. The call returns an `Os_CounterStatusType` which is a C structure of the form:

```
struct {
    TickType Delay;
    boolean Pending;
    boolean Running;
}
```

The `Delay` field, when defined, gives the number of ticks from the previous **match** at which the next **match** is due, i.e. `Delay` is the relative time between matches. The `Pending` field is set to true if the next **match** is already pending. When the `Pending` field is false, then the `Delay` holds the relative number of ticks from **now** that remain until the next **match** becomes due. This behavior is required if the application interrogates the status.

4. `Os_Cbk_Set_<CounterID>()` establishes a state in which an interrupt will become due the next time the counter matches the supplied value. The callback is passed the absolute **match** value at which the next **match** is due. The callback is used to start the counter and also to shorten the time to the next **match**. This secondary behavior is needed because you can set alarms (or start schedule tables) that need to begin at an **match** closer to now than the currently programmed **match** value.

All of the hardware driver callbacks run at OS level. This means that they will not be preempted by Category 2 ISRs and do not, therefore, need to be reentrant.

11.2 Using Output Compare Hardware

This section considers the construction of drivers for output compare (sometimes known as compare/match) counter hardware. Such hardware has the property that an interrupt is raised when a counter value (advanced by some outside process such as a clock frequency or events detected by some sensor) matches a compare value set by software. It is assumed that both the

counter value and the current compare value can be read by software. In this section, it is assumed that the registers of the counter hardware are mapped to the variables `OUTPUT_COMPARE` and `COUNTER`.

The section outlines appropriate call back functions, followed by several interrupt handlers making different assumptions about required behavior and hardware facilities.

Initially, a counter with the same modulus as `TickType` is considered. `TickType` usually has a modulus of 2^{16} on 16-bit targets and 2^{32} on 32-bit targets.

With full modulus arithmetic, the number of ticks in a delay can be determined by subtracting the start value from the end value. When the current counter value (`COUNTER`) is subtracted from the next compare value (`OUTPUT_COMPARE`), the result is the number of ticks before the **match** is reached. If this value is read after the next **match** is set, and found to be greater than the currently required delay, then the counter has passed the next **match** and there will be an extra modulus wrap (i.e. `TickType` ticks) before the compare occurs. This can happen if the delay before the next **match** is very short (for instance, one tick), in which case there is a race condition between the counter passing the intended **match** and the setting of the **match**.

11.2.1 Callbacks

Cancel

The `Os_Cbk_Cancel_<CounterID>()` call must ensure that no further interrupts will be taken. This is a hardware dependent operation that would typically be achieved by disabling interrupt generation by the counter device.

```
FUNC(void, OS_APPL_CODE) Os_Cbk_Cancel_<CounterID>(void){
    clear_pending_periodic();
    disable_interrupt_source();
}
```

Now

The `Os_Cbk_Now_<CounterID>()` call reads the free-running counter to provide the current **now** value.

```
FUNC(TickType, OS_APPL_CODE) Os_Cbk_Now_<CounterID>(void){
    return (TickType)COUNTER;
}
```



Special care may be required when reading the counter on 8-bit devices to ensure that a consistent value is obtained: in some cases, the high and low bytes must be read in a particular order in order to latch then release the counter. Similar considerations may apply when writing compare values.

Set

The `Os_Cbk_Set_<CounterID>()` call causes the interrupt to become pending when the counter value next matches the supplied parameter value. This is achieved by disabling compare matching, clearing any pending interrupt, setting the compare value, and ensuring that the interrupt is enabled. If the hardware does not provide the ability to disable compare matching, this can be simulated by setting the compare value to one less than the current counter value (thus ensuring that a **match** will not occur before the next time that the compare value is set).

Note that it may not be necessary to disable compare matching. If it can be guaranteed that a **match** will not occur between system start up and the **match** at which the hardware counter is started, then disabling compare matching is not necessary. In the example below, this is achieved by setting the compare register to the previous value of the counter, thus ensuring that a “match” interrupt will not be generated until ticks equal to the modulus of the counter have occurred. This will be long enough to perform the rest of the `Os_Cbk_Set_<CounterID>()` function. (Note that this approach can only be used if the compare register is not shared with anything else).

```
FUNC(void, OS_APPL_CODE) Os_Cbk_Set_<CounterID>(TickType Match){
    /* prevent match interrupts for modulus ticks*/
    OUTPUT_COMPARE = COUNTER - 1;
    clear_pending_interrupt();
    OUTPUT_COMPARE = Match;
    enable_interrupt_source();
}
```

The code is carefully structured to avoid two potential race conditions that can arise from dismissing the interrupt in a way that can result in unexpected interrupts being generated or expected interrupts being lost. These race conditions are as follows:

1. Pre-existing values of the compare and counter values may lead to an interrupt being raised before the compare register is set, which results in a situation where the interrupt appears to have been caused by the action of `Os_Cbk_Set_<CounterID>()` (rather than previous compare/counter values).

2. Using the `clear_pending_interrupt()` call after the compare register is set avoids the first race condition (without the need to disable the match interrupt), but may result in the situation where a very short delay (for instance, one tick after the value of the counter register when `Os_Cbk_Set_<CounterID>()` is called) is ignored. In some cases, a full counter wrap will occur before the compare causes an interrupt. Depending on the hardware, this may result in no interrupt occurring (even after a counter wrap).

In all situations, careful consideration should be given to the use of very short delays, as the counter may reach the next **match** even before it has been set, particularly if the execution path between user code which reads the current value of **now**, calculates the next **match** and sets the **match** is long. If this occurs, a full counter wrap will need to occur before the **match** occurs.

In the above example, match interrupts are prevented by means of changing the output compare register. In subsequent examples, the way in which this is achieved is not specified. Rather, it is assumed that a function `disable_compare()` is provided to prevent the hardware from generating match interrupts.



If the counter is used for some other purpose (in addition to its function as the driver for the hardware counter), the `disable_compare()` function must not halt the counter as this will lead to counter drift for other users. The re-enabling of compare matching needs to be done atomically with the assignment of the compare register. If this is not done, another race condition may exist if a short delay is set into the output compare register.

The callback shown above only works for alarms/schedule tables that you do not adjust once they have been started. If you plan to make `Set[Abs|Rel]Alarm()` calls or to `NextScheduleTable()` calls, then you need a different `Os_Cbk_Set_<CounterID>()` callback. The callback needs to be able to reset a currently programmed **match** value for a new **match** that is due to occur means it is fewer ticks from **now** than the old **match** value.

We also assume that delays due to higher priority interrupts are relatively small compared with an entire wrap of the counter modulus.

A naïve implementation would (atomically) reprogram the compare value with **match**. This is wrong because a higher priority interrupt (e.g. Category 1) could delay the write to the hardware register, so that by the time you write **match** to the compare register, **now** is already greater than **match**. This would cause all processing of the whole schedule to cease for 2^{16} (or 2^{32} or modulus) ticks. In fact, it is perfectly possible that, by the time we are

ready to write **match** to the compare register, **now** is already greater than both **match** and **old**.

Your implementation of `Os_Cbk_Set_<CounterID>()` must distinguish between the starting case (where interrupts are stopped) and the resetting case (where the schedule is running and it is being used to shorten the delay to an existing OLD compare value).

In this second case, your implementation of `Os_Cbk_Set_<CounterID>()` must return with the compare register containing the new **match** value and either;

- **now** has not exceeded **match**; or
- the compare interrupt flag is already pending. Note that if the interrupt flag is pending, it does not matter if **match** or even **old** has been passed by **now** as the hardware counter driver code you write that deals with `Os_AdvanceCounter_<CounterID>()` will (eventually) catch up to the correct time.

First you must write the new **match** to the compare register:

- If **now** is between **match** and **old**, i.e. $\text{old} - \text{match} > \text{now} - \text{match}$, then **now** has already passed **match**. You must ensure that the interrupt flag is pending before returning.
- If **now** is not between **match** and **old** then either you can return with no flag pending or both **match** and **old** have been passed and you must ensure the pending flag is set before returning. You can test for both values having been passed using $\text{now} - \text{old} < \text{old} - \text{now}$.

```
FUNC(void, OS_APPL_CODE)Os_Cbk_Set_<CounterID>(TickType Match){
    TickType Old = (TickType)COMPARE;
    TickType Now = (TickType)COUNT;
    /* Update COMPARE with new Match */
    COMPARE = Match;
    if ((Old-Match > Now-Match) || (Now-Old < Old-Now)){
        set_pending_interrupt();
    }
}
```

State

The `Os_Cbk_State_<CounterID>()` call is only made when the alarm or schedule table is running and must first check whether the next **match** has already occurred (i.e. the interrupt is pending, this can occur because all of the callbacks are executed at OS level, which will prevent the resulting ISR from preempting the currently executing task). If this is not the case, then the remaining time to the next **match** is also required.

```
FUNC(void, OS_APPL_CODE)
    Os_Cbk_State_<CounterID>(Os_CounterStatusRefType State){
    State.Delay = OUTPUT_COMPARE - COUNTER;
    State.Running = True
    if (interrupt_pending()) {
        State.Pending = True;
    } else {
        State.Pending = False;
    }
}
```



The Delay value is calculated before checking whether the interrupt is pending. This is necessary to avoid a race condition in which the interrupt becomes pending after checking but before calculating Delay, which would result in an invalid value.

11.2.2 Interrupt Handlers

Simple

In the simplest case, it is only necessary to clear the interrupt and make the required `Os_AdvanceCounter()` call. `Os_AdvanceCounter()` calls the callback `Os_Cbk_Set_<CounterID>()` to program the next **match**. This assumes that the latency of the handler to the statement at which it has set the next **match** value - i.e. after the call to `Os_AdvanceCounter()` on the compare value) is known to be less than the shortest time between two matches driven by the counter, so the **match** will be ahead of **now**.

```
#include <Os.h>
ISR(Advanced_Driver)
{
    clear_pending_interrupt();
    Os_AdvanceCounter_<CounterID>();
}
```

It is essential that the **match** is always advanced to be ahead of **now**. If the **match-now** is shorter than the handler response time, then this will not be the case and an additional full wrap of the peripheral counter will be introduced before the next **match** occurs. In order to verify that a simple handler

may be used safely, you should use schedulability analysis to verify that the simple handler can complete before its next invocation.

Re-Triggering

When matches may be too close together for the handler to advance the compare value before the next **match** is due, the handler must account for the situation in which the next **match** is already due.

This example considers the use of an output compare timer with hardware interlocking to prevent the accidental clearing of an interrupt which is raised during the clearing sequence. It is assumed that for this type of interlock, clearing the interrupt is achieved by reading the status register, then writing the status register (with a bit pattern that clears the interrupt bit). In this example, the interlock consists of two functions:

1. `prepare_interrupt_clear()`
2. `commit_interrupt_clear()`

While the driver is still running, the **match** is advanced (in the case of a full wrap, advancing by 0 is correct) and the first part of the interrupt clearing sequence is performed (reading the status register). Then a check is made that the new **match** is ahead of **now**. If this check shows that an interrupt will not be raised when the counter advances to the compare value (i.e. the next **match** is not yet due), then the interrupt clearing sequence is completed (by writing to the status register with the flag bit clear). If the check fails (i.e. the new expire is already due) then the interrupt is left pending and the handler will be re-triggered to deal with the **match**.



*The two-stage interrupt clearing sequence is required to avoid a race condition in which the counter reaches the next **match** between being tested and the interrupt being cleared. This would otherwise result in the interrupt for the next **match** being cleared. The required hardware behavior is that if the interrupt is raised again after the first stage of the sequence, then the second stage will not clear the interrupt.*

A similar approach can be taken with devices where the interrupt can be re-asserted by software. In these case, the interrupt can be cleared on entry to the handler, then re-asserted if the next **match** is due. In this case no race condition can occur (assuming there is no problem associated with software asserting an interrupt which the hardware is already asserting).

```
ISR(OutputCompareInterrupt){
    Os_CounterStatusType State;
    TickType remaining_ticks;
    Uint16 clear_tmp;
```

```

Os_AdvanceCounter_<CounterID>();

Os_Cbk_State_<CounterID>(&State);

if (State.Running == True) {
    OUTPUT_COMPARE += State.Delay;
    clear_tmp      = prepare_interrupt_clear();
    remaining_ticks = OUTPUT_COMPARE - COUNTER;
    if ((State.Delay == 0)
        || ((remaining_ticks != 0)
            && (remaining_ticks <= State.Delay))) {
        commit_interrupt_clear(clear_tmp);
    }
}
}

```



Some output compare hardware requires that the compare register be written to arm each interrupt. In such cases it is necessary to structure the code (as is the case above) so that the compare register is written to its previous value in the case of a Delay value of zero.

Looping

This section considers a generic looping ISR structure TickType modulus counter with programmable output compare.

```

#include <Os.h>
ISR(Advanced_Driver){
    Os_CounterStatusType State;
    TickType remaining_ticks;

    clear_pending_interrupt();

    while(1) {
        Os_AdvanceCounter_<CounterID>();
        Os_Cbk_State<CounterID>(&State)

        if (State.Running == False) {
            /* Exit 1: all alarms/schedule tables stopped */
            return;
        }

        OUTPUT_COMPARE += State.Delay;

        if (State.Delay == 0u) {
            /* Exit 2: full wrap */
            return;
        }
    }
}

```

```

    remaining_ticks = OUTPUT_COMPARE - COUNTER;

    if ((remaining_ticks != 0u) &&
        (remaining_ticks <= State.Delay)) {
        /* Exit 3: match is in the future */
        return;
    }

    if (interrupt_pending()) {
        /* Exit 4: interrupt pending */
        return;
    }
}
}
}

```

This interrupt handler first dismisses the invoking interrupt, then enters a loop which processes the **match** and checks whether any further matches need to be processed by this invocation. This check has four exit conditions, which must be evaluated in the order shown.

Exit 1 is taken if the counter/schedule has stopped, so no further action is necessary. If the counter has not stopped, then the next **match** is set to the required number of ticks (which will be zero in the case of a full wrap). Checks must then be made to determine whether an interrupt will be raised when the next **match** is due.

Exit 2 is taken if the Delay value indicates that a full wrap of the timer is required before the next **match** is due. Therefore, no change to the **match** value is necessary. A Delay value of 0 ensures that the new **match** is ahead of **now** (and consequently that the interrupt will be asserted when it is reached). Exiting here ensures that the following checks will not immediately identify a **match** between **now** and the **match** when a full wrap has been requested and the counter has not yet moved on¹.

Exit 3 is taken if the current timer value has not yet reached the next **match**. This check is done by determining if the time until the next interrupt (i.e. OUTPUT_COMPARE - COUNTER) is less than the Delay until the next **match**. Note that the cast to TickType is necessary to ensure that the counter modulo behavior is accounted for. The counter modulus must be the same TickType for this to work correctly (Section 11.2.3 explains how to cope with a hardware modulus not equal to TickType). If the counter has moved on by less than Delay ticks, then an interrupt will be

¹It is assumed that the interrupt will not be re-asserted while the counter and **match** continue to **match**, only when the **match** first occurs. If this is not the case, it must be ensured that the handler never exits in that state, perhaps by avoiding Delay values of zero.

raised at the correct time and the handler can exit, otherwise, the new **match** may be missed.

Exit 4 accounts for a race-condition between setting the next **match** and checking that it is ahead of the counter, since the counter can advance before the Exit 3 check is made. If Exit 3 is not taken, the next **match** is due **now**. If the interrupt is pending, Delay has already been recognized by the hardware, so the handler can exit and be re-invoked by the pending interrupt (it would not be acceptable to exit with an interrupt pending with no **match** due). Note that this construction means that it does not matter whether the interrupt is pending or not when Exit 3 is not taken because the counter has advanced by exactly the Delay value: either the pending interrupt or looping results in the next **match** being processed.

If no exit is taken, then the next **match** is due (or overdue) and another call to `Os_AdvanceCounter()` is made. The next **match** is processed and the exit checking is repeated.

Note that the typical behavior of this handler is expected to be a single `Os_AdvanceCounter()` call, because the next **match** will be in the future (i.e. it behaves just like a simple handler). Consequently, the handler should be as fast as possible for that case (since the worst-case behavior occurs when each **match** is triggered by a separate interrupt).



*It is important that you understand the interrupt behavior of the counter/compare hardware in use. When the **match** is set equal to the counter, there are three possible behaviors: the interrupt becomes pending as the **match** is set, the interrupt becomes pending as the counter moves beyond the **match**, or the counter needs to completely wrap around before the interrupt becomes pending again.*

In the example above, the test for Exit 3 assumes the counter hardware exhibits the first or third behavior. With the second behavior, it is necessary to exit if `remaining_ticks` is zero, as the interrupt will be asserted after the counter and **match** value have been observed as equal.

11.2.3 Handling a Hardware modulus not equal to TickType

The driver outlines presented in Section 11.2 so far have assumed that the counters and compare registers are the same width as `TickType` and arithmetic is unsigned modulo `TickType`. Some hardware may not have this property.

There are two cases:

1. the modulus of the hardware is less than the TickType
2. the modulus of the hardware is greater than the TickType

Both cases can be handled by changing aspects of the driver. The following sections discuss the changes in more detail.

Modulus less than TickType

In this case, we assume that the counter itself wraps to zero after some value ($m - 1$) (i.e. the counter has modulus m , where m is smaller than TickType). This increases the complexity of the drivers, but might be imposed by hardware behavior or may be necessary to support some other system requirement. For example, a timer set up with a modulus of 50000 and tick of 1ms could provide a 50ms interrupt via overflow used to drive a software counter and output compare interrupts used to provide drive a hardware counter.

Such a modulus requires modification to calculations which derive new compare values and which check the relationship between compare and counter values. The following example assumes that TickType has modulus 2^{16} .

If m is 2^x . (where $x < 16$), then it is simple to apply explicit modulus adjustments to arithmetic results by ANDing with $2^x - 1$. For 8 bit modulus, this would allow a compare value to be advanced by:

```
new_match = (old_match + Status.Delay) & 0xFF;
```

A similar operation can be applied to the result of calculating the ticks remaining to the **match**.

The calculations become more complex if the modulus value is not a power of two. Possible techniques are presented below.

Calculating of the new compare value must account for four possible results when the new Delay value is added to the **old** compare value is calculated using the TickType modulus of 2^{16} :

1. The Delay is zero. A full modulus wrap leaves the compare value unchanged.
2. The result is greater than the **old** compare value, but less than m . The result is the desired result.
3. The result is greater than m . The result of the addition needs to be wrapped at m . This can be achieved by subtracting m , avoiding the (often costly) modulus operator.

4. The result is less than the **old** compare value. The result of the addition wrapped at 2^{16} , so the result must have $(2^{16}-m)$ added to it to give the result of wrapping at m .

Note that if m is less than or equal to half the arithmetic modulus (i.e. less than or equal to half of 2^{16}), then the fourth case can never occur.

When checking whether the new output compare value has been set ahead of the counter, we consider three circumstances. No subtraction underflows the 2^{16} arithmetic modulus.

1. The Delay is zero, so the next **match** is known to be in the future. The handler is required to complete in less than the counter modulus.
2. The next **match** is greater than or equal to the counter so we can subtract counter from compare to give the interval until the next **match** then check whether this is less than or equal to the required Delay (otherwise, the next **match** has already occurred).
3. The next **match** is less than the counter value. Subtracting the **match** from counter gives the interval that remains when the interval to next **match** is subtracted from the modulus. Thus, we can calculate the interval to next **match** as $m - (\text{COUNTER} - \text{OUTPUT_COMPARE})$ and then check this result against the required Delay.

The same approach can be applied to the calculation of remaining time to **match** in the `Os_Cbk_State_<CounterID>()` call back.

Adding the mechanisms described above to conditions to the `ISR(OutputCompareInterrupt)` driver gives the following:

```
#include <Os.h>
/* The next line should result in a constant being substituted.
   We assume that the expression will be evaluated at compile
   time, avoiding modulus overflow at run time. m is the timebase
   modulus */

#define CMP_ADJUST ((TickType)65536u - m)

ISR(OutputCompare_SmallModulus){
    Os_CounterStatusType State;
    TickType counter_cache, remaining_ticks, new_match;
    clear_pending_interrupt();

    while(1) {

        Os_AdvanceCounter_<CounterID>();
```



```

Os_Cbk_State_<CounterID>(&State);

if (State.Running == False) {
    /* Exit 1: alarms/schedule tables stopped */
    return;
}

if (State.Delay == 0u) {
    /* OUTPUT_COMPARE = OUTPUT_COMPARE if
     * needed to arm next interrupt */
    /* Exit 2: full wrap */
    return;
}

new_match = OUTPUT_COMPARE + State.Delay;

if (new_match > OUTPUT_COMPARE) {
    if (new_match >= m) {
        new_match -= m;
    }
} else {
    new_match += (CMP_ADJUST);
}

OUTPUT_COMPARE = new_match;
counter_cache = COUNTER;

if (new_match >= counter_cache) {
    remaining_ticks = new_match - counter_cache;
} else {
    remaining_ticks =
        m - (counter_cache - new_match);
}

if ((remaining_ticks != 0u)
    && (remaining_ticks <= State.Delay)) {
    /* Exit 3: match in the future */
    return;
}

if (interrupt_pending()) {
    /* Exit 4: interrupt pending */
    return;
}
}
}
}

```

Modulus greater than TickType

The alternative case is where a hardware counter has a modulus that exceeds `TickType`. With a little care, such counters can be used to provide the behavior required for a `TickType` with a modulus of 2^{16} . We restrict our consideration to modulus values that are a power of two (e.g. a 32 bit counter). In these cases, the low 16 bits of the counter have the desired behavior, but overflow effects must be taken into account.

When the compare value is advanced in the interrupt handler, overflow from the bottom 16 bits must be propagated through the rest of the compare register. In addition, a Delay of 0 indicates that 2^{16} must be added to the compare value. Since the **match** can never be advanced by more than this, checks for the timer having passed the **match** can be carried out using the low 16 bits of the counter and compare registers.

When the `Os_Cbk_Set_<CounterID>()` call back is used, the **match** must be set so that it matches the counter when the low 16 bits of the counter next have the same value as the parameter passed to `Os_Cbk_Set_<CounterID>()`. This can be achieved as follows (assuming that counter and compare are 32 bit unsigned values):

```
FUNC(void, OS_APPL_CODE) Os_Cbk_Set_<CounterID>(TickType Match){
    uint32 to_compare;

    disable_interrupt_source();
    disable_compare();
    clear_pending_interrupt();

    OUTPUT_COMPARE = (COUNTER & 0xFFFF0000ul) | Match;
    to_compare      = OUTPUT_COMPARE - COUNTER;

    if ((to_compare == 0ul) || (to_compare >= 0x10000ul) {
        if(!(interrupt_pending())) {
            OUTPUT_COMPARE += 0x10000ul;
            to_compare      = OUTPUT_COMPARE - COUNTER;
            if ((to_compare == 0ul) || (to_compare >= 0x10000ul)){
                if(!(interrupt_pending())) {
                    OUTPUT_COMPARE += 0x10000ul;
                }
            }
        }
    }
    enable_interrupt_source();
}
```

The operations are carried out with interrupts from the hardware device disabled, in order to make them atomic with respect to the handler. First any

pending interrupts are cleared. This must be done after disabling comparison (for instance, setting the **match** to ensure that a pending interrupt can only be due to a **match** with the new **match**). Then, the compare register is set to the counter value with its lower 16 bits replaced by the Match parameter.

If the **match** lies in the future by less than 2^{16} ticks, then it has been set correctly. If there is a pending interrupt then the **match** must have been reached so the interrupt should be handled. Otherwise, the **match** is advanced by 2^{16} . The check must then be repeated to account for a race in which the counter could overtake the next **match** before it has been set. Checking twice is sufficient, assuming that the `Os_Cbk_Set_<CounterID>()` call completes in less than 2^{16} timer ticks.

This code assumes that the interrupt may or may not be pending if the **match** is set equal to the counter. If the interrupt is known to become pending when (or after) the two match, then the check for `to_compare` being zero should be removed.

Note that this function can be much simplified based on knowledge of application behavior. For example, if the counter is zeroed at startup and is started only once less than Match ticks after startup, then it is sufficient to set the compare value to Match.



Modulus 2^{16} behavior is not exhibited by the low 16 bits of a counter which has a modulus that is not a power of two: the last interval before the timer wraps consist of (counter modulus MOD 2^{16}) ticks.

11.3 Free Running Counter and Interval Timer

The counter compare handlers described in Section 11.2 allow the implementation of drift-free hardware counter drivers. However, not all target platforms provide such counter facilities.

Drift can be avoided when using a down counter if a separate free running counter is also available. The free running counter is used to provide a drift-free time reference, and the down counter is set up to interrupt when the next **match** becomes due. Some jitter (delay) may be introduced to individual matches due to delays in setting the interval for the down counter, but these do not accumulate (such jitter can be accounted for in the same way as jitter introduced in the handling of the interrupt). In this section, the down counter is considered to provide registers `COUNTER` and `DOWN_COUNTER` that can be used as variables. As in the previous example, both registers are taken to be `TickType` wide registers, and the values they use are taken to be unsigned `TickType` size integers.

11.3.1 Callbacks

All of the callbacks in this section assumed that the next **match** value is maintained in software and used in calculation of the down count value to the next interrupt. This can be declared as follows:

```
TickType next_match;
```

Cancel

The `Os_Cbk_Cancel_<CounterID>()` callback function only needs to disable the interrupt so the implementation is the same as before.

```
FUNC(void, OS_APPL_CODE) Os_Cbk_Cancel_<CounterID>(void){
    clear_pending_periodic();
    disable_interrupt_source();
}
```

Now

The `Os_Cbk_Now_<CounterID>()` callback function needs to return the value of the free-running counter.

```
FUNC(TickType, OS_APPL_CODE) Os_Cbk_Now_<CounterID>(void){
    return (TickType)COUNTER;
}
```

Set

Things start to change with the `Os_Cbk_Set_<CounterID>()`. The callback needs to set the `DOWN_COUNTER` so that it reaches zero (and interrupts) at a relative number of ticks from **now**. This is done by subtracting the `COUNTER` value from the `Match` value.



This relies on all three counters having the same modulus.

The callback must also log the next **match** value from the absolute `Match` parameter value passed into the call by RTA-OS3.x (this will be used by the `Os_Cbk_State_<CounterID>()` callback later).

```
FUNC(void, OS_APPL_CODE) Os_Cbk_Set_<CounterID>(TickType Match){
    /* Record value at which expire is due */
    next_match = Match;
    disable_compare();
    clear_pending_interrupt();

    /* set up interrupt when counter reaches match value */
    DOWN_COUNTER = next_match - COUNTER;
    enable_interrupt_source();
}
```

State

Note that the `Os_Cbk_State_<CounterID>()` call, below, could return `DOWN_COUNTER` as the `Status.Delay` value. If there is any jitter introduced by setting the down counter, this will reflect in the time at which the next **match** will be signaled, rather than when it is due. However, particularly with a non-TickType modulus where more calculation is avoided, the following may be acceptable.

```
FUNC(void, OS_APPL_CODE)
    Os_Cbk_State_<CounterID>(Os_CounterStatusRefType State){
    State.Delay = next_match - COUNTER;
    State.Running = True;

    if (interrupt_pending()) {
        State.Pending = True;
    } else
        State.Pending = False;
    }
    return;
}
```

11.3.2 ISR

This demonstrates a looping form of ISR: it loops until no due matches remain, rather than handling one **match** per invocation of the routine, as in a re-triggering form of ISR.

```
#include <Os.h>
ISR(IntervalTimerInterrupt){
    Os_CounterStatusType State;
    TickType remaining_ticks;

    clear_pending_interrupt();

    while(1) {
        Os_AdvanceCounter_<CounterID>();
        Os_Cbk_State_<CounterID>(&State)

        if (State.Running == True) {
            /* Exit 1: all alarms/schedule tables stopped */
            return;
        }

        next_match += State.Delay;
        /* Subtract adjustment for delay before COUNTER is set */
        remaining_ticks = next_match - COUNTER;

        if (State.Delay == 0u) {
            DOWN_COUNTER = remaining_ticks;
        }
    }
}
```

```

        /* Exit 2: full wrap */
        return;
    }

    if ((remaining_ticks!= 0u) &&
        (remaining_ticks <= State.Delay)) {
        DOWN_COUNTER = remaining_ticks;
        /* Exit 3: counter set for next expire */
        return;
    }

    /* assume we only get an interrupt due to setting the counter
       and we only set the counter when we are going to exit so
       no need to test for pending interrupt */
}
}

```

Note that exit 2 assumes that setting the counter to zero will result in an interrupt after one full wrap of ticks.

11.4 Using Match on Zero Down Counters

Some hardware might not provide a free running counter (or you might not want to use this for your hardware driver).

In this case you will have to use just the interval timer. This example assumes a 16-bit decrementing counter that raises an interrupt on reaching 0, and continues to decrement. Because the counter continues to decrement, the start point for the new countdown can be determined by adding the Delay to the counter value (assuming modulo 2^{16} arithmetic). It is desirable to minimize drift during the counter update. Preventing interrupts during the update, and adding an adjustment for the known time taken for update (to both the counter and next_match), may be able to reduce this to one tick per counter adjust (assuming the counter is asynchronous to the update, there will always be some uncertainty). counter_adjust is introduced to allow calculation of a **now** value: subtracting the counter value from next_match gives this. Note that the counter update and counter_adjust update must be atomic with respect to any call to obtain **now** for this to give the correct result.

When the driver is not running, the down counter is assumed to free-run. From start-up it runs downwards from zero and the value of **now** is (0 - counter). counter_adjust is used to hold the actual tick value that a *free running* counter would have reached the next time the DOWN_COUNTER has the value 0. this means that counter_adjust can be used to synthesize a virtual free-running counter for the purposes of the hardware counter driver.

11.4.1 Callbacks

Cancel

Canceling the driver is achieved as before.

```
FUNC(void, OS_APPL_CODE) Os_Cbk_Cancel_<CounterID>(void){
    clear_pending_periodic();
    disable_interrupt_source();
}
```

Now

The `Os_Cbk_Now_<CounterID>` callback cannot just return the value of the `DOWN_COUNTER` because the counter is not free running or monotonically increasing. Instead, the **now** value is calculated by subtracting the `DOWN_COUNTER` value from the `counter_adjust` to give the virtual free-running value.

```
FUNC(void, OS_APPL_CODE) Os_Cbk_Now_<CounterID>(void){
    return (counter_adjust - DOWN_COUNTER);
    /* counter_adjust is still correct adjustment
     * as counter runs to and through 0 */
}
```

Set

The race conditions discussed in Section 11.2.1 are still present in this model. If the interrupt is dismissed before the `DOWN_COUNTER` is set, there is a risk that an interrupt may occur between dismissing the interrupt and setting the down counter. If the interrupt is set after the down counter is set, a small delay could result in the expected interrupt being discarded. In the absence of specialized hardware protection, this can be avoided by the `disable_compare()` function setting the counter to modulus - 1, then dismissing the interrupt between determining the `AdjustedMatch` value and setting the counter (as shown in the above example).

```
TickType counter_adjust = 0;
FUNC(void, OS_APPL_CODE) Os_Cbk_Set_<CounterID>(TickType Match){
    TickType AdjustedMatch;
    AdjustedMatch = Match - (counter_adjust - DOWN_COUNTER);

    /* dismiss interrupt in a way that avoids race conditions */
    disable_compare();
    clear_pending_interrupt();

    DOWN_COUNTER    = AdjustedMatch;
    counter_adjust += AdjustedMatch;
    enable_interrupt_source();
}
```

State

`Os_Cbk_State_<CounterID>()` needs to set the Delay and can simply read the value of the `DOWN_COUNTER` to get this. The rest of the callback is identical to the others you have seen in this chapter.

```
FUNC(void, OS_APPL_CODE)
    Os_Cbk_State_<CounterID>(Os_CounterStatusRefType State){
        State.Delay = DOWN_COUNTER;
        State.Running = True;
        if (interrupt_pending()) {
            State.Pending = True
        } else {
            State.Pending = False;
        }
    }
}
```

11.4.2 Interrupt Handler

The following example shows an appropriate interrupt handler.

```
#include <Os.h>
ISR(MatchOnZeroInterrupt){
    Os_CounterStatusType State;
    TickType counter_cache;

    clear_pending_interrupt();

    while(1) {
        Os_AdvanceCounter_<CounterID>();
        Os_Cbk_State_<CounterID>(&State);

        if (State.Running == True) {
            /* Exit 1: all alarms/schedule tables stopped */
            return;
        }

        if (State.Delay == 0) {
            /* Exit 2: full wrap */
            return;
        }

        counter_cache = COUNTER + State.Delay;
        COUNTER = counter_cache;
        counter_adjust += State.Delay;

        if ((counter_cache != 0u) &&
            (counter_cache <= State.Delay)) {
            /* Exit 3: next match not yet been reached */
            return;
        }
    }
}
```



```

        if (interrupt_pending()) {
            /* Exit 4: interrupt pending */
            return;
        }
    }
}

```

The condition on Exit 3 assumes that the interrupt becomes pending when (not after!) the counter reaches zero, but may not do so if it is set to zero (if the counter is zero then the **match** is due and will be dealt with either by looping or re-entering via the pending interrupt). The same counter value must be used for both parts of the test otherwise races can occur if the counter changes between the two comparisons (hence the use of `counter_cache`).

If the behavior of the interrupt when the counter is set to zero is known, then the code can be simplified by removing Exit 4 and the associated test (since the interrupt status when `counter_cache` is zero will be known). If setting the counter to zero never causes the interrupt to become pending then that is the only change required. If setting the counter to zero always causes the interrupt to become pending, then Exit 3 should only check for `counter_cache` less than or equal to `Delay`. If the counter is zero, the interrupt will be pending and will cause the next event to be handled.

In the case of a very fast running clock (where the clock speed is greater than or equal to the processor speed), it will be necessary to add a correction to the counter to offset the number of ticks that occur between reading the counter and setting its new value. In any case, a drift of up to one tick cannot be avoided whenever the down counter is set. On a multiple interrupt level platform, it is desirable to disable all interrupts while reading/writing `COUNTER` to avoid the possibility of interruption between these operations, resulting in a large amount of drift.

11.5 Software Counters Driven by an Interval Timer

Using a periodic interval timer (or any per-event interrupt source) with an interrupt on zero can be used to synthesize a free-running counter in software. However, a handler of this form is of limited practical interest because there is one interrupt per tick. This means it is identical to incrementing a software counter. It is recommended that you use the software counter driver model instead.

11.6 Summary

- You need to provide an hardware driver for every hardware counter and advanced schedule.

- The driver interface comprises:
 - A Category 2 interrupt handler that tells RTA-OS3.x to take action; and
 - Four callback functions used by RTA-OS3.x to control the counter/schedule.
- If possible, you should use a free running counter with associated compare hardware and a simple interrupt handler.
- More advanced models can be supported though the interface if required.
- It is essential that you understand how your hardware generated the counter tick source and what happens when an interrupt from the device occurs.

12 Startup and Shutdown

Some operating systems that you might have used before will take control of the hardware. RTA-OS3.x, however, is different.

Initially the operating system is not running, so you are free to use the hardware as if no real-time operating system is being used. Until you explicitly start the operating system with an API call, it is not running.

RTA-OS3.x can be started in different application modes. A mode is a set or subset of the complete application functionality that corresponds with a specific function of the application. You will learn more about application modes in Section [12.2.2](#).

12.1 From System Reset to StartOS()

This section looks at what has to be done between an embedded processor “coming into life” when power is applied and the StartOS() API call being made to start RTA-OS3.x and your application. The details of what goes on in this period are naturally dependent on the particular embedded processor in use - the underlying principles are however the same. You should read this section in conjunction with the reference manual for your target processor and apply the concepts we describe to your own platform.

12.1.1 Power-on or Reset

When power is applied to an embedded processor, or the processor is reset, the processor does one of two things (depending on the type of processor).

It may start executing code from a fixed location in memory, or it may read an address from a fixed location in memory and then start executing from this address. The fixed location in memory that contains the address of the first instruction to execute is often called the “reset vector” and is sometimes an entry in the interrupt vector table.

In a production environment, the reset vector and/or the first instruction to be executed is usually in non-volatile memory of some variety. In a development environment it is often in RAM to permit easy re-programming of the embedded processor. Some evaluation boards (EVBs) have switches or jumpers that permit the reset vector and/or the first instruction to be in EEPROM or RAM.

Going from power-on or reset to the first instruction being executed is often referred to as “coming out of reset”. After a processor has come out of reset it usually:

- has interrupts disabled,

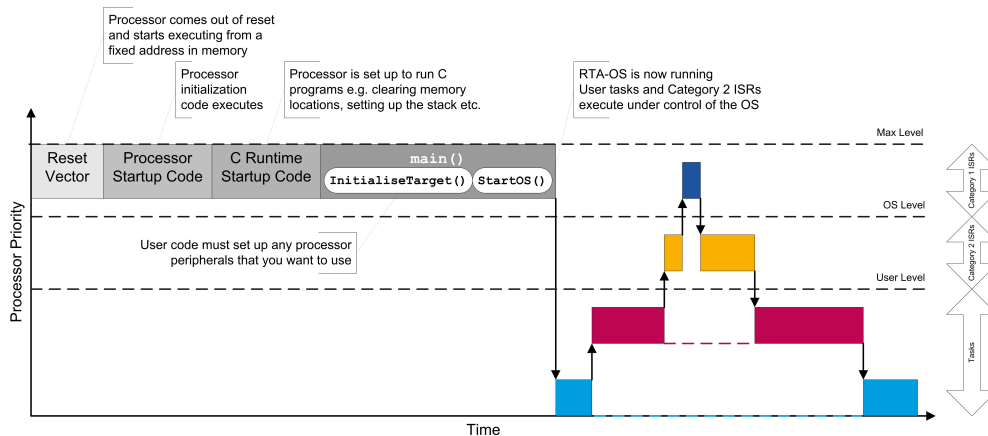


Figure 12.1: System Startup

- is in supervisor mode (if the processor supports it) - i.e. it can execute all instructions and access all addresses without causing an exception and has all forms of memory and I/O protection turned off.
- is in single-chip mode (if the processor supports it) - i.e. the chip is in a “self-contained mode” where external memory is not usable and external buses are disabled.

12.1.2 C Language Start-up Code

It is possible to have any code you would like executed when a processor comes out of reset but it is normal if using a high-level language such as C for this bootstrap code to be supplied with your compiler.

The compiler vendor supplies an object module or library that contains the bootstrap code. The bootstrap code usually does two key things:

1. it carries out basic processor configuration, for example bus configuration, enabling of access to internal RAM, etc.
2. it invokes the C language start-up code. Most of this is concerned with initializing data structures, clearing memory, setting up the stack pointer, etc.

Directives in the object module/library or in the linker configuration file are used to ensure that the bootstrap code (and reset vector value if needed) are placed in the correct location in memory.

The C language start-up code is usually supplied by the compiler vendor in an object module with a name like crt0 or startup and the code can usually

be identified in a map file by looking for a symbol with a name something like `_start` or `__main`. The source to this module is usually available to you.

For some target hardware, ETAS supplies a different version of the standard startup code that should be used with RTA-OS3.x applications. The *Target/Compiler Port Guide* and the example supplied with RTA-OS3.x will tell you how to use this.

The start-up code initializes the C language environment. For example, it sets up the stack pointer, the heap used for `malloc()` and it initializes global variables by copying their default values from ROM into RAM. Finally, the start-up code invokes the application start-up code.

12.1.3 Running `main()`

The application start-up code is typically in a function called from `main()`. The application start-up function has two things to do to work with RTA-OS3.x:

1. Initialize the target hardware into a state where RTA-OS3.x and the application can run
2. Call `StartOS()` to start RTA-OS3.x running.

For example the application start-up code for an RTA-OS3.x application may look like:

```
OS_MAIN(){
    InitializeTarget();
    StartOS(OSDEFAULTAPPMODE);
    /* Never reach here */
}
```

The macro `OS_MAIN()` is provided for your convenience by RTA-OS3.x to mark the main function of your application - you do not have to use this to work with RTA-OS3.x. The macro is used to handle the cases where using **`void main(void)`** is forbidden by the compilers.

The `InitializeTarget()` function in the above example need to be written by you to initialize the target hardware. The remainder of this section describes the types of things that you may have to do to initialize target hardware into a state where your application and RTA-OS3.x can run. This description is necessarily generic as every embedded processor is slightly different. It is probably wise to read this section in conjunction with the *Target/Compiler Port Guide* for your processor and the processor's reference guide.

Setting up Memory

In general, memory configuration is carried out by the bootstrap code that is run before the application start-up code is executed. In more complex embedded processors. However, the memory configuration set-up by the bootstrap code may not be what is required for the application. For example, if the processor has internal RAM and an external memory bus then it is most likely that the bootstrap code will have configured the processor to use the internal RAM. If your application needs to use RAM on the external memory bus, then you will need to configure the processor to use the external RAM. Configuring access to RAM typically involves programming bank select and mask registers - however the details depend on the embedded processor.

Setting up Peripherals

Most embedded applications make use of peripheral devices which may be part of the embedded processor or attached through I/O or memory buses. Examples are CAN controllers, Ethernet controllers and UARTs. It is generally a good idea to set-up peripheral devices before RTA-OS3.x is started since at this point the application code cannot be preempted and has complete control over interrupts.

Setting up Interrupts

Interrupt sources for Category 1 and 2 interrupts should be configured before `StartOS()` is called. Typically, you should ensure that the IPL is set to OS level and then both configure interrupt sources. You can also enable Category 1 interrupt sources here.



Do not enable Category 2 interrupts before calling `StartOS()` as this can result in a race condition where the interrupt needs to be handled before RTA-OS3.x has been initialized. You should use the `StartupHook()` to enable Category 2 interrupt sources. This model means that Category 2 interrupts will not be generated until `StartOS()` lowers the IPL just before it enters the idle mechanism.

On some microcontrollers it will be necessary to program priority registers in the hardware that configure interrupt priorities. The values you program must match the priority values that you told RTA-OS3.x at configuration time, otherwise your application will not work properly. On targets where this is the case, RTA-OS3.x will usually provide helper code so that you can do this job correctly. You should check the *Target/Compiler Port Guide* for any special instructions relating to target initialization.

Enabling Interrupts

Category 1 interrupts may also be enabled so that they generate interrupts immediately as the handling of Category 1 interrupts is completely outside the scope of RTA-OS3.x.

Category 2 interrupt sources must not actually generate interrupts until after `StartOS()` has completed initialization. You must not enable Category 2 interrupt sources before calling `StartOS()`. If you do this, then you can get a race condition where the interrupt occurs before RTA-OS3.x is correctly initialized.



Enabling Category 2 interrupt sources before `StartOS()` will result in undefined behavior.

RTA-OS3.x provides a safe way to enable Category 2 interrupt sources using the `StartupHook()` which is described in Section [12.2.1](#).

Setting up Timers

Most embedded applications use hardware timers. Timers are usually configured to “tick” and generate interrupts at a fixed frequency. The ISR associated with the timer interrupts then either activates a task directly or ticks an OSEK counter (i.e. calls `IncrementCounter(CounterID)`).

Setting up a hardware timer depends on the design of the timer but there are two common forms:

1. a count register is set to zero and a match register is set to the maximum value for the count register. The count register is incremented by the processor at a given frequency and, when it reaches the value in the match register, it generates an interrupt and resets the count register to 0.
2. a count register is loaded with the number of ticks to occur before an interrupt should be generated. The processor decrements the count register at a given frequency. When the register reaches zero, an interrupt is generated. Usually the ISR that handles the interrupt is responsible for reloading the count register.

The frequency at which timers must run will depend on your application. It is vital that all counters run at the frequency specified in their definition. If you have told RTA-OS3.x that a counter driven by a timer has a particular tick rate, i.e. you have specified the “Seconds Per Tick” attribute, then you must make sure that your timer hardware is configured to give a tick at the same rate.

12.2 Starting RTA-OS3.x

Once your hardware is initialized, you can start RTA-OS3.x

RTA-OS3.x is started only once a `StartOS()` call is made. This call is usually made from the main program. It is up to you to perform any hardware initialization that is necessary for the application. The initial state of RTA-OS3.x is described in the *Reference Guide*.

`StartOS()` takes a single application mode parameter. This parameter is either the default mode `OSDEFAULTAPPMODE` or another mode that has been configured in [rtaoscfg](#).

Have a look at the example main function in Code Example 12.1, which starts the operating system in the default application mode.

```
#include <Os.h>
OS_MAIN(){
    InitializeTarget();
    StartOS(OSDEFAULTAPPMODE);
    /* Never reach here */
}
```

Code Example 12.1: Example Main Function

The call to `StartOS()` does not return. Once the RTA-OS3.x is initialized, all interrupts are enabled and the `Os_Cbk_Idle()` runs until a higher priority task or ISR occurs.

Most RTA-OS3.x API calls can be made from the idle mechanism. However, you cannot use any calls that would require the idle mechanism to terminate (for example, it is not possible to call `TerminateTask()` from the idle mechanism).



You should not make RTA-OS3.x API calls that manipulate OS objects or enable Category 2 interrupts before calling `StartOS()`.

RTA-OS3.x can be suspended by disabling all Category 2 interrupts and ensuring that they will not be raised on some future event, such as an output compare match.

RTA-OS3.x will be suspended when no Category 2 interrupts are raised and the idle mechanism is running. You can resume RTA-OS3.x by re-enabling Category 2 interrupts and then resume making RTA-OS3.x calls.



Figure 12.2: Execution of the Startup Hook

12.2.1 Startup Hook

The Startup Hook is called by RTA-OS3.x during the StartOS() call after the kernel has been initialized, but before the scheduler is running.

StartOS() raises the interrupt priority level (IPL) to OS level as soon as it is called and lowers it to user level just before it returns. This means that the startup hook runs with Category 2 ISRs masked. That means you can safely enable interrupt generation in StartupHook() knowing that it will not actually result in an interrupt occurring until StartOS() has completed initialization and RTA-OS3.x is ready to run. At this point StartOS() un-masks Category 2 interrupts and the OS is running.

Figure 12.2 shows the execution of the Startup Hook relative to the initialization of RTA-OS3.x.

Code Example 12.2 shows how Startup Hook should appear in your code.

```
FUNC(void, OS_APPL_CODE) StartupHook(void) {
    /* Startup hook code. */
    EnableIOInterrupts();
    EnableTimerInterrupts();
    ...
}
```

Code Example 12.2: Using the Startup Hook

The Startup Hook is often used for the initialization of target hardware (for example the enabling of interrupts sources that have been configured in by the code you executed before the call to StartOS()).

12.2.2 Application Modes

Applications can be started in different modes, which might represents part of the complete functionality. These modes could correspond with specific functions of the application. You could have, for example, an end-of-line programming mode, a transport mode and a normal mode.

You can define as many application modes as you want. Figure 12.3 shows how to declare different application modes in `rtaoscfg`.



You must declare an application mode called OSDEFAULTAPPMODE.

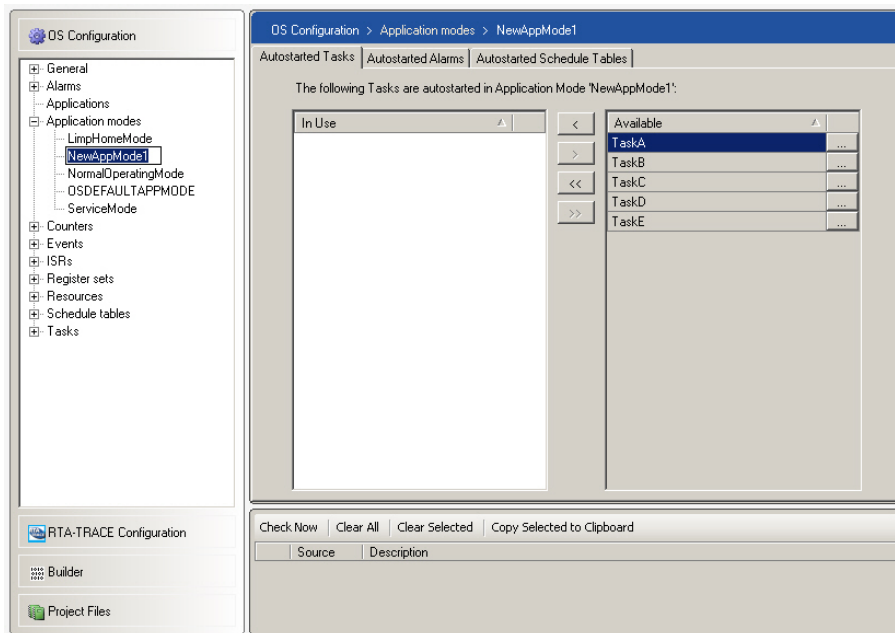


Figure 12.3: Configuring Application Modes

You can use the `GetApplicationMode()` API call to work out which mode you are in. This means that you can write application code that is mode-dependent. Code Example 12.3 shows how a task can be written so that it has different behavior in different modes.

```
TASK(Moded){
    AppModeType CurrentMode;
    GetApplicationMode(&CurrentMode);
    switch (CurrentMode) {
        case DiagnosticMode:
            DoExtendedFunctionality();
            break;
        case LimpHome
            DoBasicFunctionality();
            break;
        default:
            DoNormalFunctionality();
            break;
    }
    ...
}
```

Code Example 12.3: Adding modes to a task

Application modes can also be associated with a set of tasks, alarms and schedule tables that are started automatically when the operating system

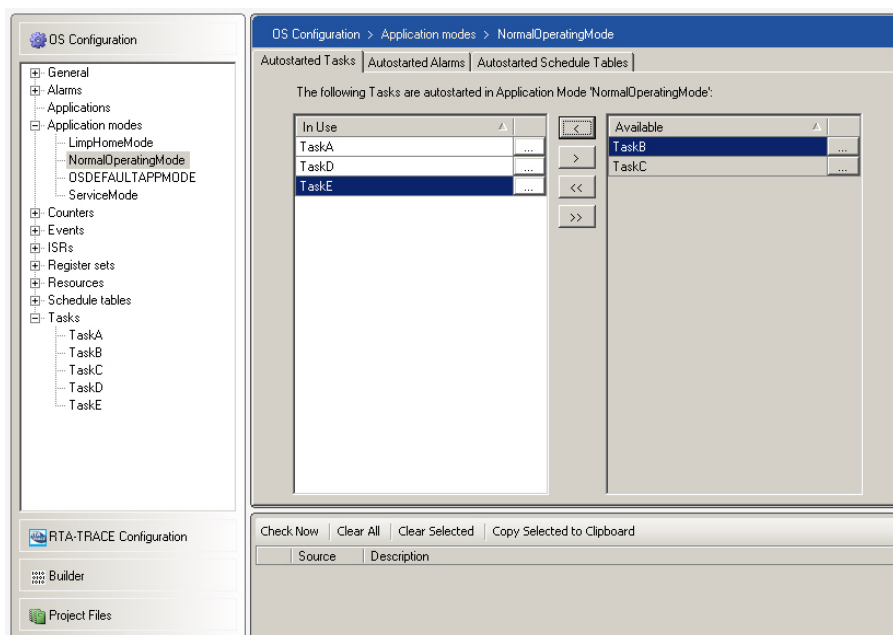


Figure 12.4: Declaring an auto-started tasks

starts. This means you can customize what happens during `StartOS()` for each of your declared modes.

Auto-starting Tasks

Any task can be auto-started in any application mode. When you auto-start a task the OS activates the task during the call to `StartOS()` i.e. an `ActivateTask()` API call is made internally. If you auto-start a basic task then it will have run and terminated before you reach the `Os_Cbk_Idle`. If you auto-start an extended task then it will run and either reach its first `WaitEvent()` API call for an event that has not yet been set or it will have terminated before you reach the `Os_Cbk_Idle`.



You do not need to auto-start tasks that you don't need to run immediately on startup. Tasks that are not auto-started can still be activated and run at a later stage through normal activation operations, expiry of alarms, processing of schedule table expiry points, etc.

Figure 12.4 shows that TaskA, TaskD and TaskE are auto-started in the `NormalOperatingMode` application mode.

Recall from Section 4.5.3 that the auto-start properties of a task can be configured for the task itself, so you have the option of specifying the tasks which are started in a chosen application mode or, if you prefer, the application modes in which a chosen task is auto-started. `rtaoscfg` will automatically keep these two views consistent if you make a change in the other view.

Auto-starting tasks is typically useful for two cases:

1. Running an initialization task before other tasks in the system start to execute.

If you need to do this, then you must ensure that the auto-started task has a higher priority than any of the tasks that need to run after the initialization task.

2. Starting extended tasks.

You will recall from Section 7.2 that you cannot set events for extended tasks in the suspended state and that the structure of the task is typically an infinite loop and a series of `WaitEvent()` calls. By auto-starting extended tasks you can avoid any potential errors that may occur by setting events on auto-starting extended tasks.



Auto-started tasks execute in priority order, from the highest to the lowest priority. If a higher priority task sets events for a lower priority task, then the events will be processed by the lower priority task when it executes.

Auto-starting Alarms

Alarms can also be auto-started in any application mode. When `StartOS()` returns, all auto-started alarms will have been set.

Figure 12.5 shows you how an alarm are configured for auto-starting.

Recall from Section 9.3 that you can also specify auto-starting from the alarm itself. As with auto-started task, `rtaoscfg` will automatically keep these two views consistent if you make a change in the other view.

Note however, that the configuration of an auto-started alarm's absolute or relative start tick value is configured for the alarm itself.



This means that the same auto-start characteristics apply across all application modes in which the alarm is auto-started.

The absolute and relative auto-start value have the same behavior as `SetAbsAlarm()` and `SetRelAlarm()` respectively and configuration uses the same types of parameters. If an alarm is auto-started, then you must specify an alarm time and a cycle time. The same restrictions apply for these parameters as for the `offset`, `start` and `cycle` parameters to the alarm API calls:

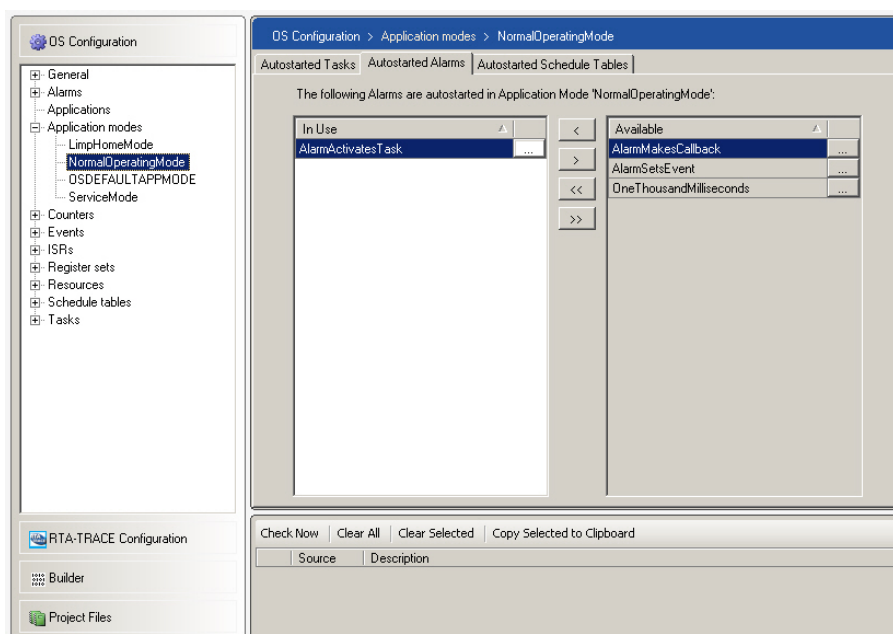


Figure 12.5: Auto-starting an Alarm

	Alarm Time		Cycle Time	
	Min	Max	Min	Max
Relative	1	MAXALLOWEDVALUE	MINCYCLE	MAXALLOWEDVALUE
Absolute	0	MAXALLOWEDVALUE	MINCYCLE	MAXALLOWEDVALUE



If you auto-start an alarm in absolute mode with alarm time zero, then the alarm will not expire until a full modulus wrap of the underlying counter has occurred (i.e. after MAXALLOWEDVALUE+1 ticks have elapsed) because 0 is already in the past. For example, if you have an alarm on a millisecond counter then it will not occur until 65536ms (65.5 seconds) have elapsed.

Auto-started alarms are useful when you want to start a set of cyclic (periodic) tasks when the OS starts. If you are using alarms to start multiple tasks and you need the tasks to run at specific cyclic rates *relative to each other*, then you must make sure that the alarms are auto-started. This is the only way to guarantee alarm synchronization.

Auto-starting Schedule Tables

Schedule tables can be auto-started in any application mode. When `StartOS()` returns, all auto-started schedule tables will be running.

Like alarms, schedule tables can be started at either an absolute or relative tick value. This has the same behavior as `StartScheduleTableAbs()` and `StartScheduleTableRel()` respectively and configuration uses the same

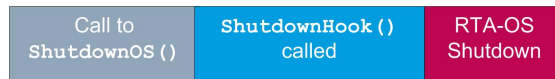


Figure 12.6: Execution of the Shutdown Hook

types of parameters. If a schedule table is auto-started, then you must specify an absolute start value or a relative offset depending on the mode in which you start the schedule table. The same restrictions apply for these parameters as for the offset, start and cycle parameters to the schedule table start API calls:

	Relative Offset		Absolute Value	
	Min	Max	Min	Max
Relative	1	MAXALLOWEDVALUE	-	-
Absolute	-	-	0	MAX



Schedule tables that are started with an absolute value zero will not expire until a full modulus wrap of the underlying counter has occurred (i.e. after MAXALLOWEDVALUE+1 ticks have elapsed) because the tick value of zero is already in the past when the schedule table is started.

12.3 Shutting Down RTA-OS3.x

The operating system can be shutdown at any point by making the ShutdownOS() API call. When this happens, RTA-OS3.x will immediately disable interrupts and then enter an infinite loop. If you have configured the ShutdownHook() it is called before the infinite loop is entered.

The ShutdownHook() is always passed a parameter that can be used to determine the reason for shutdown and then take any necessary action.

12.3.1 Shutdown Hook

The Shutdown Hook is called during the execution of the ShutdownOS() API call. Figure 12.6 shows the execution of the Shutdown Hook with respect to a ShutdownOS() API call.

Code Example 12.4 shows how Shutdown Hook might appear in your code.

```
FUNC(void, OS_APPL_CODE) ShutdownHook(StatusType Error) {
    /* Shutdown hook code. */
    switch (Error) {
        case E_OK:
            /* Normal shutdown */
            break;
        default:
            /* Abnormal shutdown */
            LogError();
    }
}
```

```

        break;
    }
    for(;;); /* Wait for reset */
}

```

Code Example 12.4: Using the Shutdown Hook

You should not normally return from the `ShutdownHook()`. If you do then RTA-OS3.x will disable all interrupts and enter an infinite loop running at OS level.

12.4 Restarting RTA-OS3.x

AUTOSAR OS does not provide any mechanism for restarting the OS at runtime other than through a watchdog reset. This is an unfortunate side-effect of `StartOS()` not returning when a `ShutdownOS()` call is made. This is a significant shortcoming in the AUTOSAR standard because it is an extremely common requirement to be able to restart the OS in different modes during runtime. For example, an ECU may have a power-saving mode or a “limp-home” mode.

RTA-OS3.x removes this limitation by providing two API calls that are used in combination to restart the OS.



Restarting of the OS is unique to RTA-OS3.x and is not part of the OSEK or AUTOSAR standards. Use of the features described in this section are therefore not portable to other implementations.

The API call `Os_SetRestartPoint()` places a marker in your code to which the API call `Os_Restart()` jumps when the call is made. `Os_SetRestartPoint()` cannot be made once `StartOS()` has been called and therefore must occur before the `StartOS()` call for restart to be possible.



It is only possible to restart RTA-OS3.x once it has been shutdown. You can only call `Os_Restart()` from the `ShutdownHook()`.

Using this feature allows you to jump back to any arbitrary point in your pre-`StartOS()` initialization, so you can place code to initialize other parts of the system outside the OS.

Code Example 12.5 shows how you might use structure of the main program when using `Os_SetRestartPoint()` to place a marker.

```

AppModeType StartupAppMode;
OS_MAIN(){
    InitializeTarget();
    /* Set up normal application mode */
    StartupAppMode = NormalOperation;
    Os_SetRestartPoint(); /* We will return here on restart */
}

```

```

switch (StartupAppMode) {
  case NormalOperation:
    /* Do mode-specific initialization */
    break;
  case LimpHome:
    /* Do mode-specific initialization */
    break;
  ...
}
StartOS(StartupAppMode);
}

FUNC(void, OS_APPL_CODE) ShutdownHook(StatusType Error){
  ...
  if (FailureDetected == True) {
    StartupAppMode = LimpHomeMode;
    Os_Restart();
    /* Never reach here */
  }
  ...
}

```

Code Example 12.5: Using `Os_SetRestartPoint()` and `Os_Restart()`

12.5 Summary

- RTA-OS3.x will not work unless everything is located in the right place in memory.
- The target hardware must be initialized before RTA-OS3.x can run.
- RTA-OS3.x does not run until the `StartOS()` call is made.
- RTA-OS3.x can be stopped at any time using the `ShutdownOS()` call.
- RTA-OS3.x can be restarted by using the `Os_SetRestartPoint()` call to place a restart marker before `StartOS()`, using the `Os_Restart()` call to jump back to the marker and calling `StartOS()` to restart RTA-OS3.x.
- Tasks, alarms and schedule tables can be auto-started in different application modes.

13 Error Handling

Many of the RTA-OS3.x API calls return an error code at runtime which tells you whether the OS detected an error during the execution of the API call or not. The set of error codes that are returned depend on two things:

1. the build status of the OS
2. the API call itself

The OS provides two types of build status:

Standard status does a minimum amount of runtime error checking and is intended for production builds of your application (i.e. the build that you will send into series production after you have gained sufficient confidence that your application is free from errors). Four classes of error are detected:

1. E_OK - no error was detected. It is possible that this is because no error checking was done. In this case the call will not have modified the state of the OS (it will have silently failed).
2. E_OS_LIMIT - an internal limit of the OS was reached, for example you tried to activate a task more often than your configuration allows.
3. E_OS_NOFUNC - the call cannot be made
4. E_OS_STATE - the call cannot be made because the object is not in a valid state

Extended status performs the checks as standard build, but adds a significant amount of extended error checking to check for all reasonable violations of OS API usage. There are too many errors to list here, but they fall into 3 classes:

1. E_OK - no error was detected. It is possible that this is because no error checking was done. In this case the call will not have modified the state of the OS (it will have silently failed).
2. E_OS_<standard_code> - an error case defined by the AUTOSAR (or OSEK) OS standard occurred.
3. E_OS_SYS_<vendor_code> - an error case defined by ETAS occurred, in addition to the cases identified by the E_OS_<standard_code> codes occurred.



You are strongly encouraged to use extended status during in the early stages of development so that you can debug any problems arising from incorrect use of the RTA-OS3.x API. When you are sure that you are using the OS correctly, you can use standard status to check non-functional properties of your application like production memory sizes and performance.

Each API call that returns an error code will return a different set of values depending on what type of errors can occur when the call is made.

Common (extended build) errors are:

Error Code	Meaning
E_OS_ID	You made an API call on the wrong type of object
E_OS_VALUE	A parameter is outside a permitted range
E_OS_CALLEVEL	You made an API call from the wrong place

You can find out which API calls return which error codes, and what each code means for the specific API (and therefore what you might have to do to fix the error) by referring to the *Reference Guide*.

13.1 Centralized Error Handling - the ErrorHandler()

The common way of checking errors from either standard or extended status builds is to use the error hook which provides a “catch all” error handler. If the ErrorHandler is enabled, then it is called by RTA-OS3.x when any API call is about to return an error code that is not E_OK. The error code is passed into the ErrorHandler routine and you can use it to work out which error has occurred.

Figure 20.11 shows how the ErrorHandler is enabled.



If you enable the error hook then you must provide an implementation. If you do not provide an implementation then your program will not link correctly.

Depending on the severity of the error, you can decide whether to terminate (by calling ShutdownOS()) or to resume (by handling or logging the error and then returning from ErrorHandler()). Code Example 13.1 shows you the usual structure of the Error Hook.

```
FUNC(void, OS_APPL_CODE) ErrorHandler(StatusType status) {
    switch (status) {
        case E_OS_ACCESS:
            /* Handle error then return. */
            break;
        case E_OS_LIMIT:
            /* Terminate. */
            ShutdownOS(status);
    }
}
```

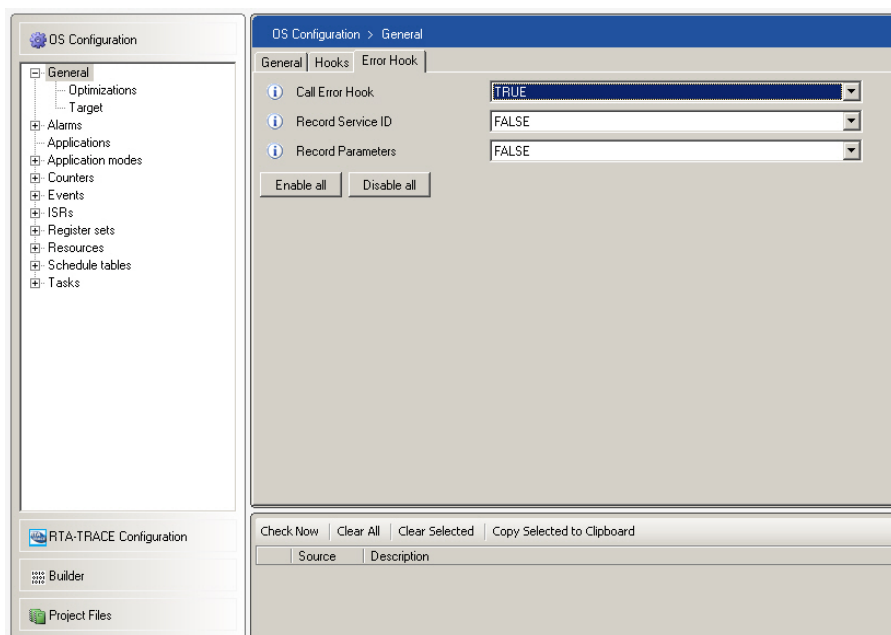


Figure 13.1: Configuring the Error Hook

```

default:
    break;
}
}

```

Code Example 13.1: Suggested Structure of the Error Hook

The Error Hook is adequate for coarse debugging - it tells you that something has gone wrong. For example, if you get `E_OS_CALLEVEL`, then you know that you have made an API call from the wrong context *somewhere* in your code but you have no indication where it might be. You really need to know more about the error so that you can remove the bug. In this case, you need to know which API call resulted in the error being generated. You might find in some cases that knowing which parameters were passed to an API call when it failed helps you to debug a problem. This information is available at run-time by configuring advanced error logging.

13.1.1 Configuring Advanced Error Logging

Three levels of detail are available:

1. Do not record the service details (default)
2. Record the API name only.
3. Record the API name and the associated parameters.

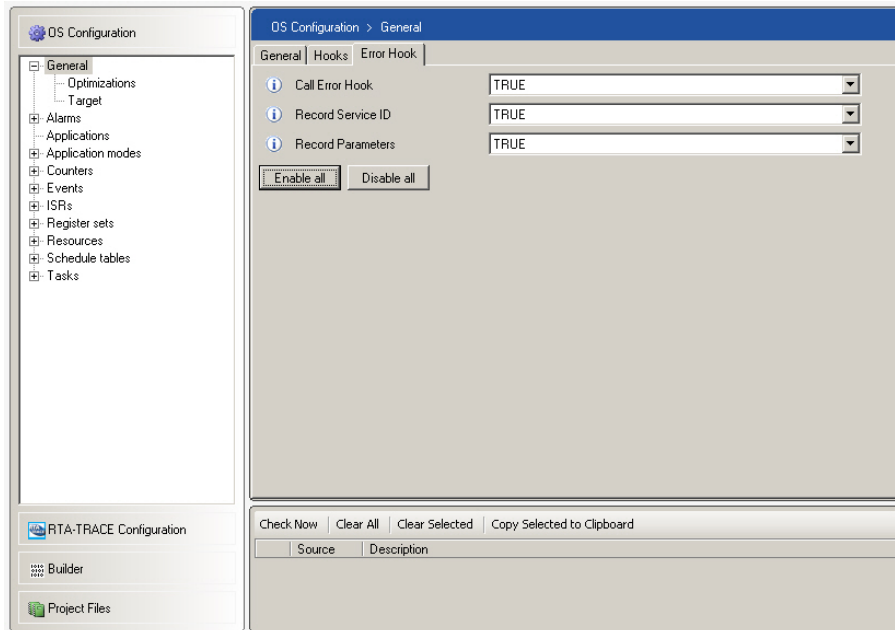


Figure 13.2: Configuring Advanced Error Logging

Figure 13.2 shows how the level of detail is defined in `rtaoscfg`.

If you choose not to record the service details, your application does not need to pay the additional overheads associated with collecting this information.

Using Advanced Error Logging

When error logging is enabled, RTA-OS3.x provides a set of macros for accessing the name and the associated parameters of the API call that caused the error.

You can find out which API call caused the error using the `OSErrorGetServiceId()` macro. This macro returns an `OSServiceIdType` of the form `OSServiceId_<API name>`. If, for instance, an `ActivateTask()` call results in an error, `OSErrorGetServiceId` will return `OSServiceId_ActivateTask`.

The parameters to the API call are available using macros in the form shown in Code Example 13.2. A macro is defined for each parameter of each API call.

```
OSError_<API Name>_<API Parameter Name>
```

Code Example 13.2: Advanced Error Logging

Using the `ActivateTask()` example again, `OSError_ActivateTask_TaskId` will return the `TaskId` parameter passed to `ActivateTask()`. This additional

error logging information can be usefully incorporated into the `ErrorHook()` code. This is shown in Code Example 13.3.

```
FUNC(void, OS_APPL_CODE) ErrorHook(StatusType status) {
    OSServiceIdType callee;
    switch (status) {
        case E_OS_ID:
            /* API call called with invalid handle. */
            callee = OSErrorGetServiceId();
            switch (callee) {
                case OSServiceId_ActivateTask:
                    /* Handle error. */
                    break;
                case OSServiceId_ChainTask:
                    /* Handle error. */
                    break;
                case OSServiceId_SetRelAlarm:
                    /* Handle error. */
                    break;
                default:
                    break;
            }
            break;
        case E_OS_LIMIT:
            /* Terminate. */
            ShutdownOS();
        default:
            break;
    }
}
```

Code Example 13.3: Additional Error Logging Information

The macros for obtaining the API name and the associated parameters should only be used from within the Error Hook. The values they represent do not persist outside the scope of the hook.



When you use extended error logging, the value returned by `OSErrorGetServiceId()` may be misleading. This generally happens when API calls have a side effect. For example if you activate a task from a schedule table expiry point and that task activation results in an error, then `OSErrorGetServiceId()` will return `OSServiceId_ActivateTask` even though the API call that you made was `Os_AdvanceCounter()`.

13.1.2 Working out which Task is Running

When debugging your RTA-OS3.x applications, you will probably want to know which task or Category 2 ISR is responsible for raising the error. OSEK OS provides the `GetTaskID()` API call to tell you which task is running.

Code Example 13.4 shows you how to do this.

```
TaskType CurrentTaskID;
/* Pass a TaskRefType for the return value of GetTaskID() */
GetTaskID(&CurrentTaskID);
if (CurrentTaskID == Task1) {
    /* Code for task 1 */
} else {
    if (CurrentTaskID == Task2) {
        /* Code for task 2 */
    }
    ...
}
```

Code Example 13.4: Using GetTaskID()

13.1.3 Working out which ISR is Running

AUTOSAR OS extends the OSEK scheme to Category 2 ISRs with the GetISRID() API call.

Unlike GetTaskID(), GetISRID() returns the ID of the ISR through the return value of the function rather than as an out parameter to the function call. If you call GetISRID() and a task is executing, then the function returns INVALID_ISR.

The following code shows how to use GetISRID() together with GetTaskID().

```
ISRType CurrentISRID
TaskType CurrentTaskID;
/* Is an ISR running? */
CurrentISRID = GetISRID();
if ( CurrentISRID != INVALID_ISR ) {
    if (CurrentISRID == ISR1) {
        /* Work out which ISR */
    }
} else {
    GetTaskID(&CurrentTaskID);
    if ( CurrentTaskID == Task1 ) {
        /* Work out which task */
    }
}
}
```

13.1.4 Generating a Skeleton ErrorHandler()

Writing error hooks that trap the types of errors that your configure may generate can be time consuming and error-prone. RTA-OS3.x can help this activity by generating the framework for the ErrorHandler() that includes checking for all types of error, for all API calls.

The framework `ErrorHook()` is generated using the following **rtaosgen** command line:

```
C:\>rtaosgen --samples:[ErrorHook] MyConfig.xml
```

This generates an error hook in `Samples\Hooks\ErrorHook.c` that you can use in your application. If the file is already present, then **rtaosgen** will generate a warning. If you want to overwrite an existing file, then you can use:

```
C:\>rtaosgen --samples:[ErrorHook]overwrite MyConfig.xml
```

13.2 In-line Error Handling

An alternative to the `ErrorHook()` is to check the API return codes in-line with calling. This means that you can build some degree of run-time fault tolerance into your application.

This may be useful if you want to check for error conditions that can occur in the Standard status (such as `ActivateTask()` returning `E_OS_LIMIT`). Code Example 13.5 shows you how this can be done.

```
TASK(FaultTolerant){
    /* Do some work */
    if (ActivateTask(HelperTask) != E_OK) {
        /* Handle error during task activation. */
    }
    TerminateTask();
}
```

Code Example 13.5: In-line Error Checking

13.3 Conditional Inclusion of Error Checking Code

If you are adding code to check for runtime errors that only occur in extended status, then you do not want to go through your application by hand to remove this code at when you change to standard status.

RTA-OS3.x provides two macros that allow you conditionally include/exclude code during development:

```
OS_STANDARD_STATUS    is defined when standard status is configured
OS_EXTENDED_STATUS    is defined when extended status is configured
```



The macros `OS_STANDARD_STATUS` and `OS_EXTENDED_STATUS` are provided by RTA-OS3.x only and are not necessarily portable to other implementations.

13.4 Summary

- AUTOSAR OS provides facilities for debugging through the Error Hook which provides a mechanism for trapping exceptional conditions at runtime. It can provide a resumption model of exception handling.
- Further information on the source of an error is available through macros accessible in the `ErrorHook()`.

14 Measuring and Monitoring Stack Usage

RTA-OS3.x provides stack monitoring features that can be used during development to check whether you get any unexpected stack overruns.

When stack monitoring is configured, RTA-OS3.x also provides features for measuring the stack usage of each task and ISRs at runtime. This can be used to identify which tasks consume what stack space and can help provide information that might be useful for optimizations (for example, identifying which tasks could share an internal resource to reduce the amount of stack required).

You may also want to collect accurate stack usage information for each task so that the stack allocations you specify are not pessimistic - i.e. you don't tell RTA-OS3.x that tasks use more stack space than is really necessary.

14.1 Stack Monitoring

A common problem when building embedded systems is that of stack overrun, i.e. tasks and or ISRs consuming too much stack space at runtime.

AUTOSAR OS allows you to monitor the stack for overruns. When stack monitoring is enabled, RTA-OS3.x checks on each context switch whether the stack has exceeded its pre-configured stack allocation value (see Section 4.6.3).



Category 1 ISRs in your system bypass RTA-OS3.x completely and therefore consume stack without OS knowledge. If your Category 1 ISRs result in stack problems then these will not be detected by RTA-OS3.x.

RTA-OS3.x calls ShutdownOS(E_OS_STACKFAULT) when a stack fault is identified. This is the behavior required for AUTOSAR but this is not very useful because it does not allow you to try and identify what has failed and by how much the stack has been overrun¹ In RTA-OS3.x you can you can override this behavior and trap problems with the `Os_Cbk_StackOverrunHook()` instead. Section 14.2 provides more details.

Stack monitoring impacts both the memory footprint and the run-time performance of RTA-OS3.x and is therefore disabled by default. Stack monitoring is enabled in **General → Stack Monitoring Enabled**. Figure 14.1 shows how to select your chosen option.

When you configure Stack Monitoring you need to define a stack allocation budget for each task and Category 2 ISR. This figure must include the stack

¹In extreme cases, it may not be possible for you do anything, but one advantage of the single stack model used by RTA-OS3.x is that you can add a system-wide stack safety-margin at link time and then use this 'spare' stack space for debugging if a stack fault occurs.

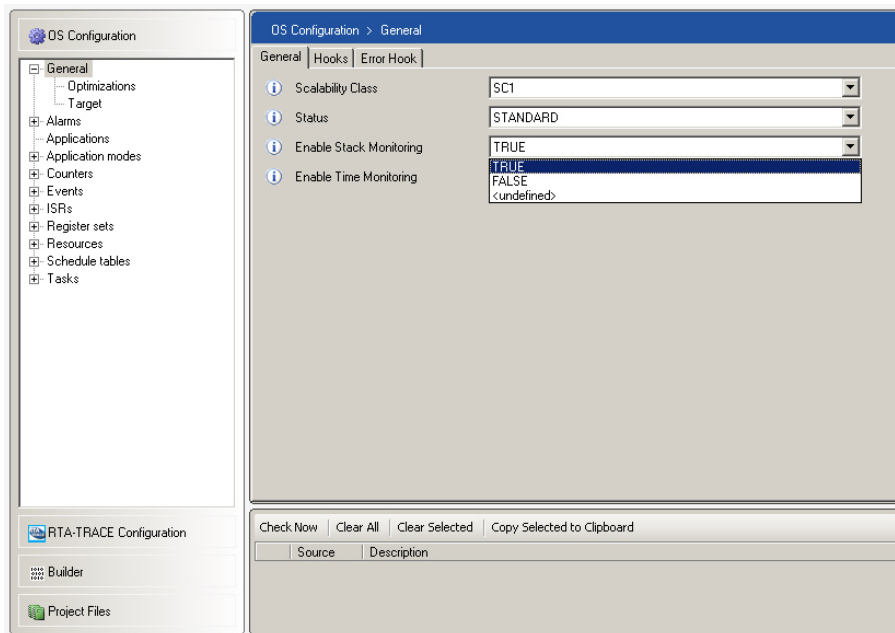


Figure 14.1: Enabling Stack Monitoring

required by your application and the stack required for the RTA-OS3.x context. Section 14.3 explains how to use RTA-OS3.x's stack measurement features to get this data.

RTA-OS3.x provides 2 ways to define the stack allocation:

1. Task/ISR defaults
2. Per task/ISR configuration

If a per task/ISR value is configured for a task/ISR, then this overrides the default value.

14.1.1 Setting Defaults

Default settings set the stack allocation for all tasks, all Category 2 ISRs and all Category 1 ISRs. You can see how to do this in Figure 14.2. If no other stack allocation is specified elsewhere, then RTA-OS3.x uses the default value.

14.1.2 Configuring Stack Allocation per Task/ISR

Each task and ISR can specify its own stack allocation as part of the task/ISR configuration. Figure 14.3 shows how this is configured for tasks, ISRs have a similar configuration element. Whenever you specify a stack allocation value for a task/ISR the value configured overrides any default value that you might have set.

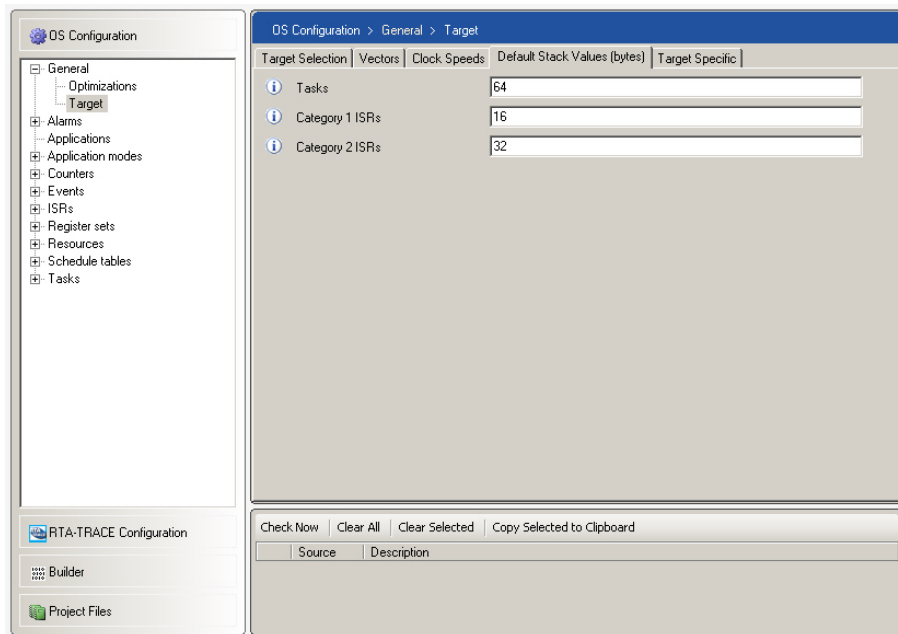


Figure 14.2: Setting default stack allocation

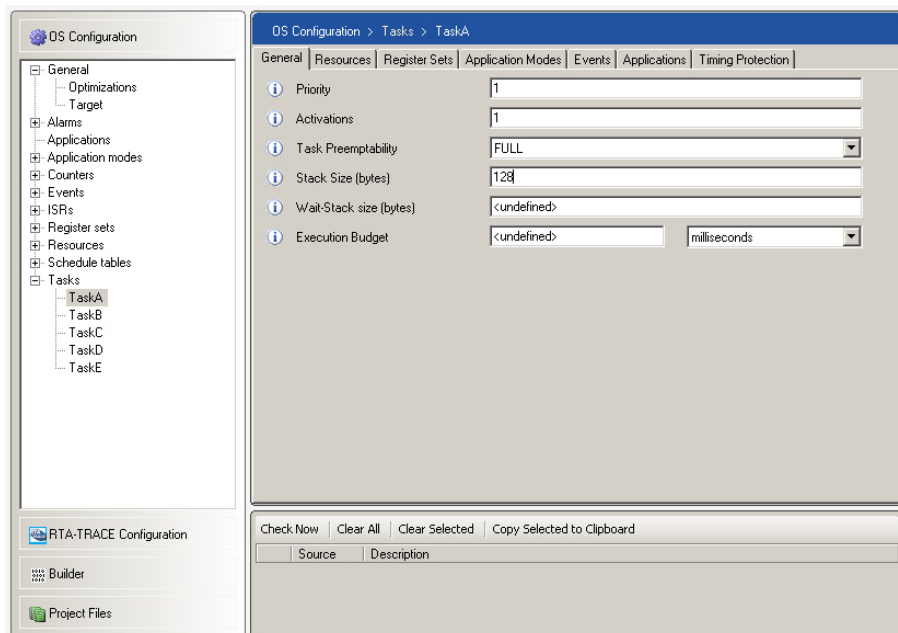


Figure 14.3: Setting Stack Allocation for Tasks

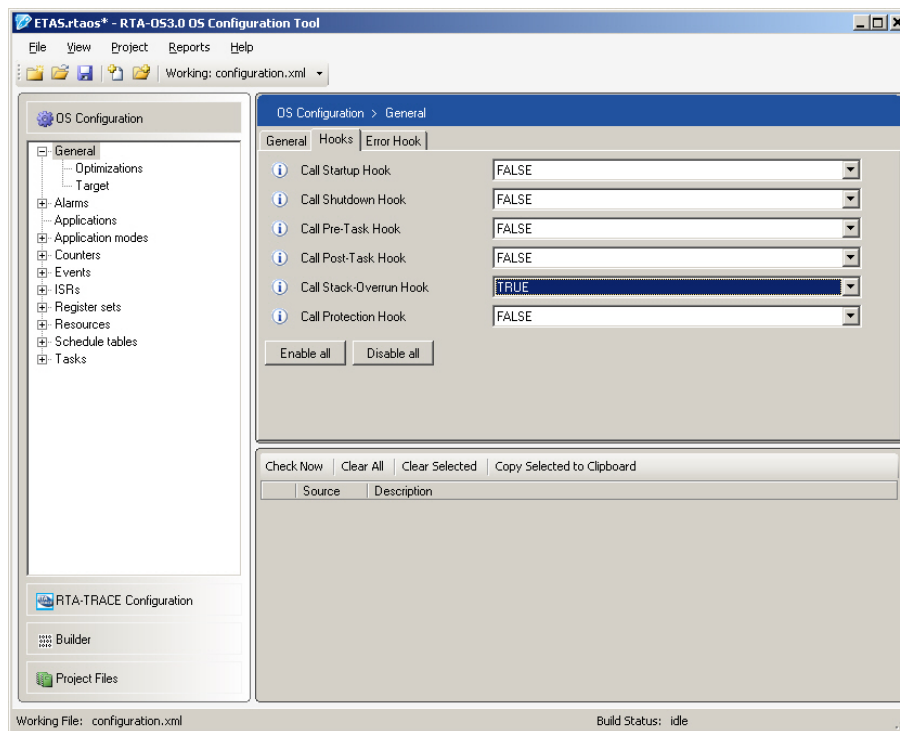


Figure 14.4: Configuring the stack overrun hook

14.2 Using the `Os_Cbk_StackOverrunHook()`

Recall from Section 4.6.5 that RTA-OS3.x can be configured to call the `Os_Cbk_StackOverrunHook()` when problems with extended task management are detected at runtime. The same hook is used by RTA-OS3.x for reporting stack overruns detected by stack monitoring.

If you configure RTA-OS3.x to use the `Os_Cbk_StackOverrunHook()` as shown in Figure 14.4 then RTA-OS3.x will call the hook when a problem is detected by stack monitoring.

ETAS *Calling `Os_Cbk_StackOverrunHook()` when a problem is detected by stack monitoring is an RTA-OS3.x extension to AUTOSAR OS and is not portable to other implementations.*

The hook is passed a parameter indicating the number of bytes by which the stack has overrun and a reason for the problem. Stack monitoring adds another reason - `OS_BUDGET` - to those presented in Section 4.6.5. `OS_BUDGET` indicates that a task has exceeded its stack allocation.

`OS_BUDGET` is similar to `OS_ECC_START` - it identifies a situation where the stack has overrun. The difference between the two cases is that `OS_ECC_START` only occurs when an extended task is started (basic tasks that exceed their configured stack allocation do not result in this error) whereas

OS_BUDGET problems are detected on every context switch for every type of task and ISR.

As with the `ErrorHook()` you can make calls to `GetTaskID()` and `GetISRID()` to identify what was executing at the point the problem occurred. Code Example 14.1 shows an example `Os_Cbk_StackOverrunHook()`.

```
FUNC(void, OS_APPL_CODE) Os_Cbk_StackOverrunHook(Os_StackSizeType
    Overrun, Os_StackOverrunType Reason) {
    ISRType CurrentISRID
    TaskType CurrentTaskID;

    /* Work out what has failed */
    CurrentISRID = GetISRID();
    if ( CurrentISRID != INVALID_ISR ) {
        /* An ISR has overrun */
        if (CurrentISRID == ISR1) {
            /* Work out which ISR */
        }
    } else {
        /* It must be a task that has overrun */
        GetTaskID(&CurrentTaskID);
        if ( CurrentTaskID == Task1 ) {
            /* Work out which task */
        }
    }

    /* Work out why */
    switch (Reason) {
        case OS_BUDGET:
            /* Problem: The task/ISR exceeded its stack
             allocation */
            /* Solution: Add Overrun to the stack allocation */
            break;
        case OS_ECC_START:
            /* Problem: Some lower priority task on the stack
             has used too much stack space */
            /* Solution: Enable stack monitoring to find out
             which task */
            break;
        case OS_ECC_WAIT:
            /* Problem: The extended task had consumed too much
             stack space then executing WaitEvent() */
            /* Solution: Add Overrun to the WaitEvent() stack
             allocation */
            break;
    }
}
```

Code Example 14.1: The Stack Overrun Hook



When `Os_Cbk_StackOverrunHook()` is entered this indicates that your system is not behaving as expected. You should not return from the `Os_Cbk_StackOverrunHook()`. Entering the hook usually means that your stack is corrupt. If you do return from the hook then the behavior of your application is undefined.

14.3 Measuring Stack Usage

The figures that you supply for stack monitoring represent the worst-case stack used by each task and should be the sum of the space required by the task. This includes the context for RTA-OS3.x and the space required for worst-case function call tree made by the task (where worst-case means the tree that results in the most stack space being used by the task).



Stack measurement is a feature of RTA-OS3.x and is not portable to other implementations of the OSEK or AUTOSAR OS standards.

When stack monitoring is enabled, RTA-OS3.x automatically logs the worst case stack usage seen at runtime for each task and Category 2 ISR. The API calls `Os_GetTaskMaxStackUsage()` and `Os_GetISRMaxStackUsage()` are provided to allow you to find out what has been logged. The maximum values are logged on a context switch. If the stack usage is greater than the currently recorded value then the maximum is updated.



If your task or Category 2 ISR has not been preempted, terminated or entered the waiting state at least once then RTA-OS3.x will not yet have logged a value and the `Os_Get[Task|ISR]MaxStackUsage()` calls will return zero.

If a context switch does not happen at the worst-case stack depth then the values reported by `Os_GetTaskMaxStackUsage()` and `Os_GetISRMaxStackUsage()` will be wrong. To avoid this problem, RTA-OS3.x provides the `Os_GetStackUsage()` API call that can be used to instrument your application to ensure that the worst cases are logged. `Os_GetStackUsage()` returns the current stack usage for the caller but it also updates the maximum observed value each time it is called if and only if it is higher than the currently recorded value. Section 14.3.1 provides additional information on how to use `Os_GetStackUsage()` to instrument your application to worst-case stack usage measurement.



It is highly recommended that you use the `Os_GetTaskMaxStackUsage()`, `Os_GetISRMaxStackUsage()` and `Os_GetStackUsage()` to measure the stack values you need to use in extended task management.

On targets that have a single stack, the stack measurement API calls return a scalar value indicating the number of bytes of stack space consumed by the calling Task/Category 2 ISR. On targets with multiple stacks the calls return

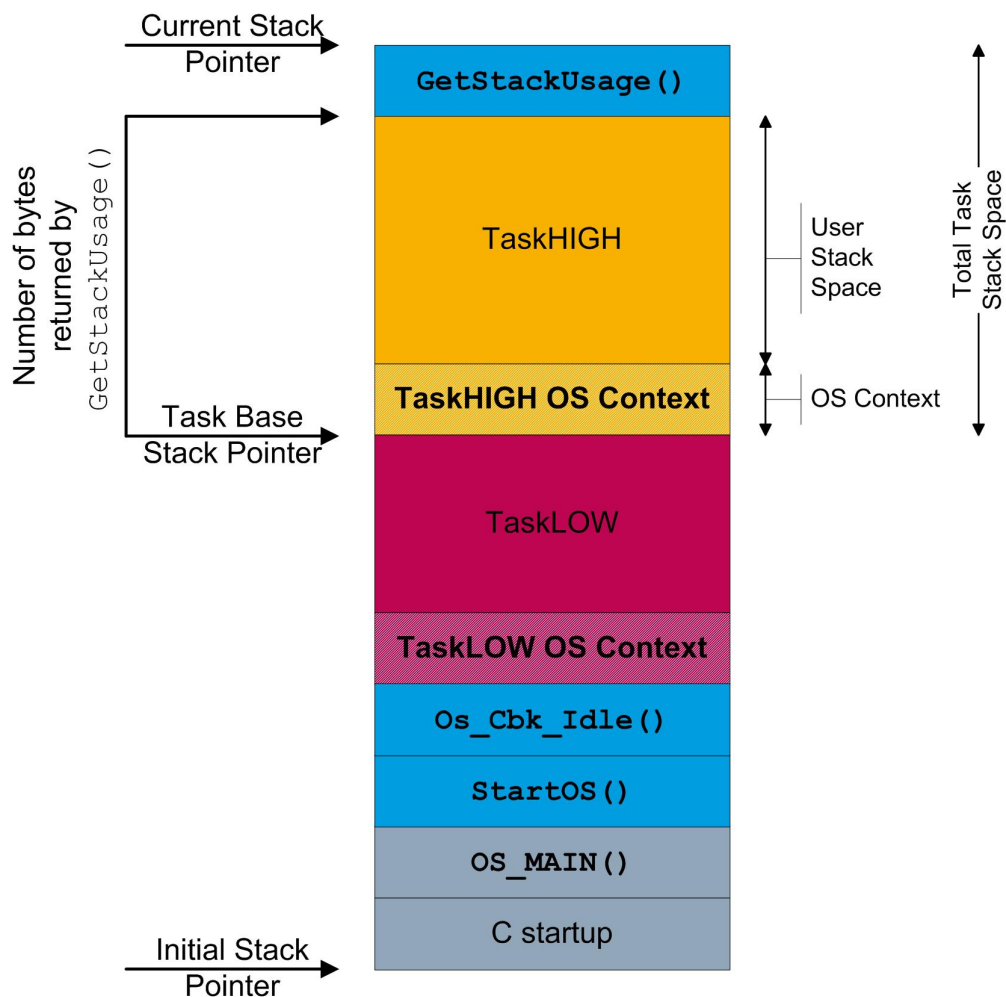


Figure 14.5: Stack Diagram

a data structure containing the number of bytes used on each stack. The *Target/Compiler Port Guide* for your port will tell you how to extract stack space information from this data structure.

The values returned are measured from the initial value of the stack pointer at the point RTA-OS3.x starts the task/ISR. This means that measurements include the stack context required by RTA-OS3.x. However, the stack values returned do not include the stack space required for the calls themselves. Figure 14.5 shows the size returned by `Os_GetStackUsage()` when it is called from task TaskHIGH.

14.3.1 Marking the Worst Case for Function Calls

To measure the worst-case stack usage for each task or ISR, you need to place a call `Os_GetStackUsage()` call at each leaf of your function call hierarchy.

If you have leaves that are library functions then you will need to make a `Os_GetStackUsage()` call in the parent function and determine the worst-case stack space of the library call. The worst-case stack space requirement for the RTA-OS3.x API is provided in the *Target/Compiler Port Guide* for your port. If you make calls to other libraries at the leaves of your call hierarchy, you must contact the vendor to obtain the worst-case stack requirements for the library calls you make.

Code Example 14.2 shows a task that makes a number of function calls. It shows the placement of `Os_GetStackUsage()` calls required to measure stack usage.

```
#include <Os.h>

Os_StackSizeType Measurement1;
Os_StackSizeType Measurement2;
Os_StackSizeType Measurement3;

void Function1(void) {
    ...
    Measurement1 = Os_GetStackUsage();
    ActivateTask(Higher);
    ...
}

void Function2(void) {
    ...
    Function3();
    Measurement2 = Os_GetStackUsage();
    ...
}

void Function3(void) {
    ...
    Measurement3 = Os_GetStackUsage();
    ...
}

TASK(Low) {
    Function1();
    ...
    Function2();
    TerminateTask();
}
```

Code Example 14.2: Measuring Stack Usage

The worst-case stack usage (WCSU) for Code Example 14.2 will be the maximum value of `Measurement1`, `Measurement2` and `Measurement3`. Figure 14.6

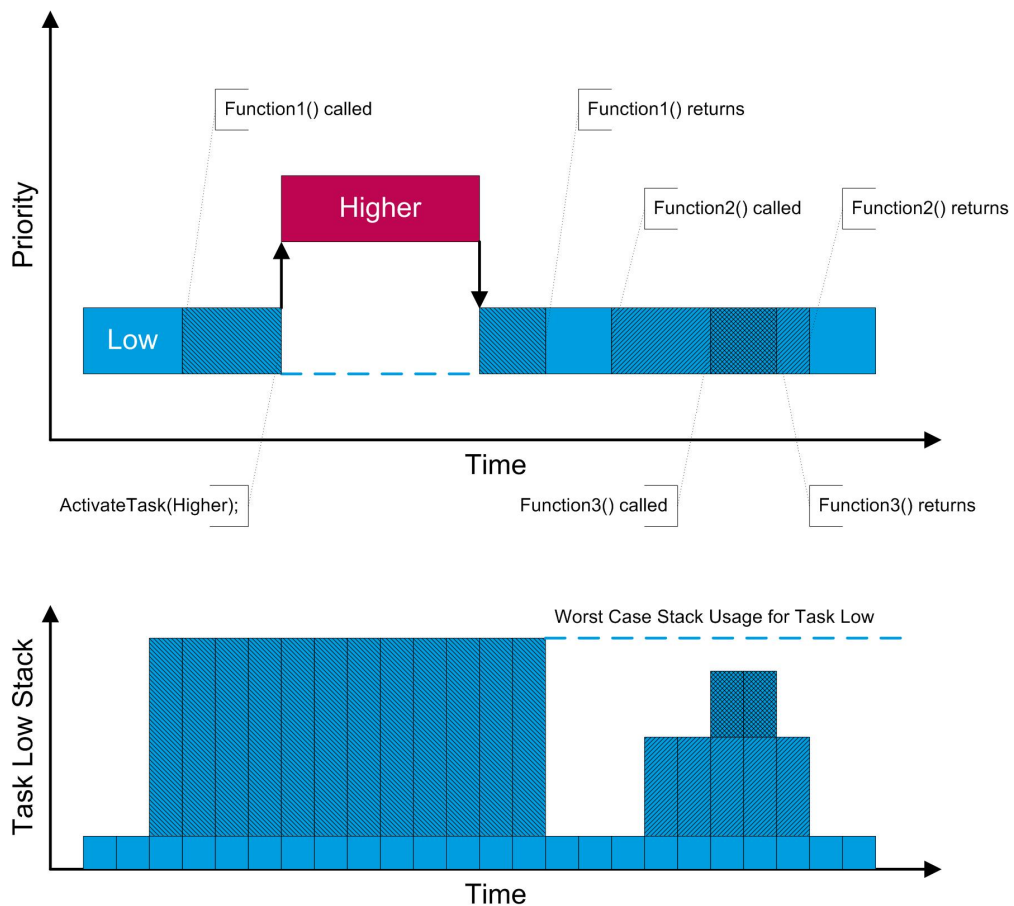


Figure 14.6: Measuring the worst-case stack for the program call tree

show Code Example 14.2 executing. In this case, the WCSU is when task Low calls Function1().

14.4 Summary

- RTA-OS3.x provides in-kernel features that allow you to measure and monitor stack usage at runtime.
- Each task and ISR must specify a stack allocation in bytes for each stack used.
- Arbitrary measurements of the current stack pointer value can be made using the `GetStackOffset()` API call.
- Stack faults can be handled by calling `ShutdownOS()` (as specified by the AUTOSAR OS standard) or can alternatively be re-directed to RTA-OS3.x's `Os_Cbk_StackOverrunHook()` for diagnosis.

15 Measuring and Monitoring Execution Time



All timing monitoring and measuring facilities provided by RTA-OS3.x are not part of the OSEK or AUTOSAR OS standards and are therefore not portable to other implementations.

RTA-OS3.x provides facilities for measuring the execution times of user code at the kernel level.

15.1 Enabling Time Monitoring

Time monitoring can be used in both standard and extended builds and is enabled by setting **Time Monitoring Enabled** to true in the **General/Timing** tabbed control. The feature needs access to a free running hardware timer, ideally one that runs at the same speed as your CPU clock because this will allow RTA-OS3.x to carry out cycle-accurate measurements.

Before you can use time monitoring you need to tell RTA-OS3.x some details about the timing of the target hardware. There are two values to provide:

1. the instruction cycle rate

This is the rate at which instructions are executed on your target hardware (sometimes called the clock speed).

2. the stopwatch speed

This is the rate at which the stopwatch timer runs. Ideally, this will be the same speed as the instruction cycle rate. However, it might be slower than the CPU instruction rate because your timer module might use some kind of pre-scaler.

RTA-OS3.x generates a set of macros that encapsulate this information to allow you to scale timing measurements:

Macro	Description
OSCYCLEDURATION	The duration of a CPU instruction in nanoseconds.
OSCYCLESERSECOND	The number of CPU instructions in a second.
OSSWTICKDURATION	The duration of a stopwatch tick in nanoseconds.
OSSWTICKSERSECOND	The number of stopwatch instructions in a second.

15.1.1 Providing a Stopwatch

The free running timer you provide is called the “stopwatch” and is used by RTA-OS3.x to measure execution times. RTA-OS3.x gets access to the stopwatch using a callback function called `Os_Cbk_GetStopwatch()`.



An implementation of `Os_Cbk_GetStopwatch()` must be provided if you are using RTA-OS3.x's time monitoring functionality. Your program will not link correctly if you do not provide this function.

Any code that your application uses to obtain execution times should be conditionally compiled. RTA-OS3.x provides the macro `OS_TIME_MONITORING`, which allows you to do this. Code Example 15.3 shows an example of conditional compilation when getting the time that a resource is held.

Code Example 15.1 shows a typical example.

```
#ifndef OS_TIME_MONITORING
FUNC(Os_StopwatchTickType, OS_APPL_CODE)
    Os_Cbk_GetStopwatch(void) {
        return (Os_StopwatchTickType)TIMER_CHANNEL_0;
    }
#endif /* OS_TIME_MONITORING */
```

Code Example 15.1: Providing a stopwatch

The stopwatch returns ticks and any values reported by RTA-OS3.x are in terms of ticks on the stopwatch time base. You can use the macros provided by RTA-OS3.x to convert stopwatch measurements into ‘clock time’ units like milliseconds, microseconds etc.

Uncertainty in Stopwatch Measurements

If the stopwatch runs slower than the CPU clock, then when RTA-OS3.x reads the stopwatch, there is a possibility that the time is less than the real amount of time that has elapsed. This occurs because of the difference in resolution of the CPU clock and the stopwatch. Figure 15.1 shows the basic issue - you might read the lower resolution stopwatch just before it will be incremented by the CPU clock.

This difference is called the uncertainty and you will need to compensate for this in any calculations you do that use time measurement.

This does not occur for stopwatches that run at the same rate as the CPU clock because you are already using the maximum possible resolution of time. The stopwatch uncertainty is equal to zero if the instruction cycle rate and the stopwatch speed are equal. In most other cases the uncertainty is one (but see Section 15.1.2).

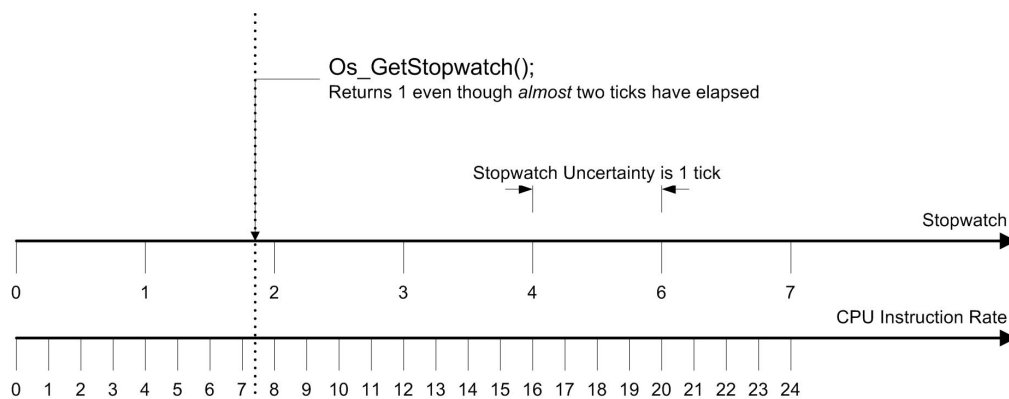


Figure 15.1: Uncertainty in stopwatch measurements

15.1.2 Scaling the Stopwatch

In most cases your, `Os_Cbk_GetStopwatch()` will return a value read directly from a hardware timer and you will convert timing measurements into 'real' time after measurement.

However, you may prefer to scale the stopwatch directly in the `Os_Cbk_GetStopwatch()` callback so that all times reported by RTA-OS3.x are already in the units you require. For example, Code Example 15.2 shows how to scale the stopwatch from Code Example 15.1 so that the stopwatch returns a value in nanoseconds.

```
FUNC(Os_StopwatchTickType, OS_APPL_CODE)
    Os_Cbk_GetStopwatch(void) {
        return (Os_StopwatchTickType)(TIMER_CHANNEL_0 *
            OSSWTICKDURATION);
    }
```

Code Example 15.2: Providing a stopwatch

Scaling the stopwatch also has an impact on the stopwatch uncertainty as shown in Figure 15.2.

An appropriate modification to the stopwatch uncertainty calculations you make is to multiply the uncertainty by the scaling factor.

15.2 Automatic Measurement of Task and ISR Execution Times

When your application uses time monitoring, RTA-OS3.x measures the execution times of each task and Category 2 ISR in your application.

RTA-OS3.x maintains a log of the longest observed execution time over all executions for each task and Category 2 ISR. The execution time for tasks is measured as follows:

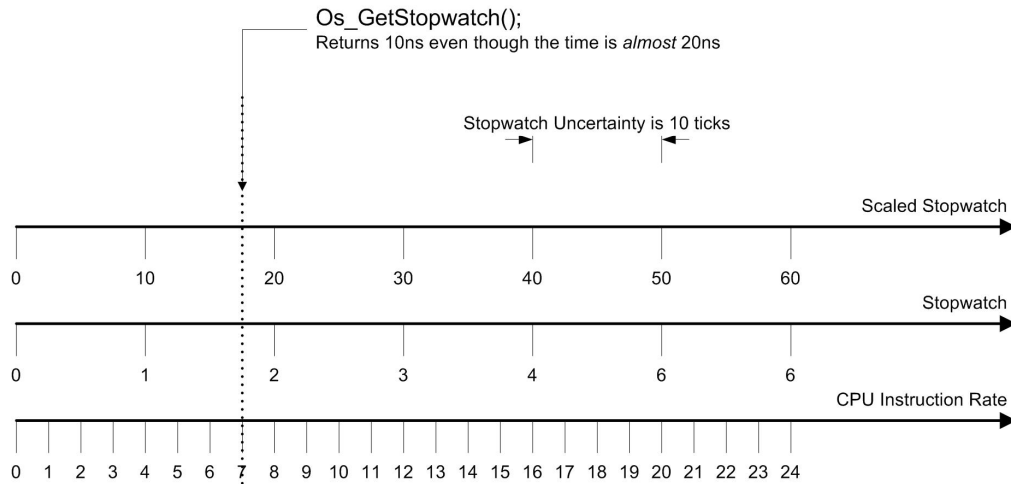


Figure 15.2: Uncertainty in scaled stopwatch measurements

Basic Tasks are measured from their first instruction to the completion of the `TerminateTask()` API call.

Extended Tasks are measured from their first instruction to the first `WaitEvent()`, between adjacent `WaitEvent()` calls and from `WaitEvent()` to the `TerminateTask()` API call.

Pre- and post-task hooks, if configured, are not included in the execution time measurement.

Execution times are measured using the stopwatch provided by `Os_Cbk_GetStopwatch()`. RTA-OS3.x automatically compensates for preemption during measurement. When a task is preempted then the measurement for the preempted task stops and measurement for the preempting task starts as shown in Figure 15.3.



Measurements are taken on a context switch (or, in the case of extended tasks, the possibility for a context switch). This means that a switch must occur for a time to be recorded. Therefore, a basic task must terminate at least once for a timing measurement to be made and an extended task must either terminate or make a `WaitEvent()` call.

The largest observed execution time for each task/ISR can be read using `Os_GetLargest[Task|ISR]ExecutionTime()` API call. The call returns zero if the task/ISR has not yet completed an execution.

The best place to record task and ISR execution times is in `Os_Cbk_Idle()` since, if the code here executes, you can be guaranteed that there are not tasks or ISRs that are ready to run. Code Example 15.3 shows a typical example.

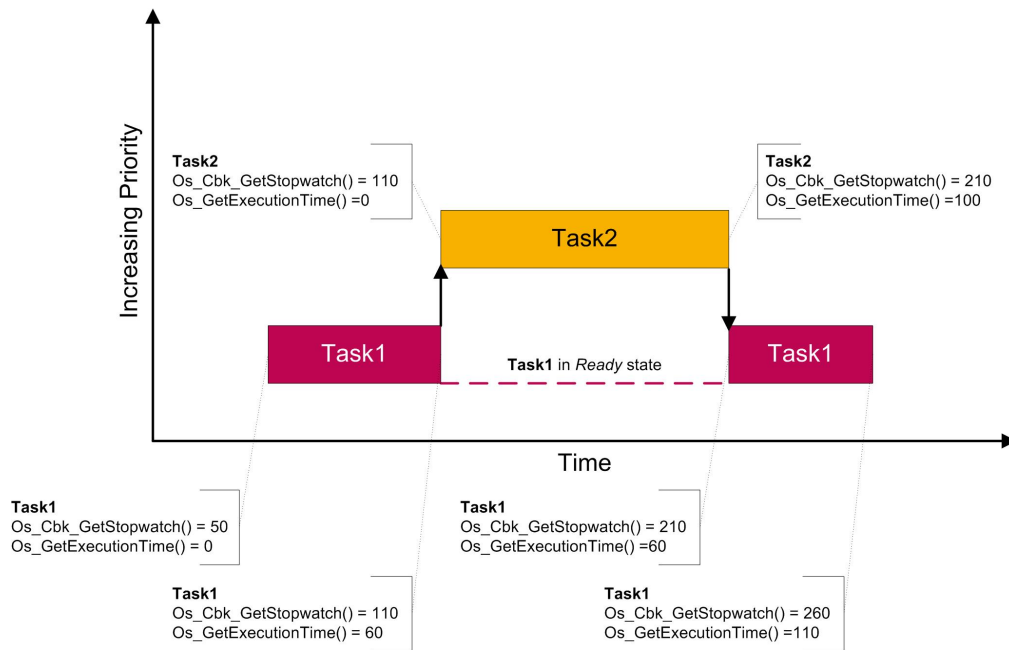


Figure 15.3: Compensating for preemption in timing measurements

```

FUNC(boolean, OS_APPL_CODE) Os_Cbk_Idle() {
    #if defined(OS_TIME_MONITORING)
        Os_StopwatchTickType TaskTime;
        Os_StopwatchTickType ISRTime;
        TaskTime = GetTaskMaxExecutionTime(MyTask);
        ISRTime = GetISRMaxExecutionTime(MyISR);
    #endif
    return TRUE;
}

```

Code Example 15.3: Reading the longest observed execution times

You can reset a largest time using the `Os_ResetLargest[Task|ISR]ExecutionTime()` API call.

15.3 Manual Time Measurement

RTA-OS3.x's time monitoring provides a API called `Os_GetExecutionTime()` that can be used to get the current stopwatch value. By placing this call before and after any section of code, you can measure the execution time of any fragment of your program. For example:

- you might want to profile the execution of some 3rd party library code
- you may want to debug exactly where time is being consumed by your own applications

- you might want to measure the blocking due to resource locking or the disabling of interrupts

Code Example 15.4 shows how you can measure blocking times. The same principle applies to any code section that you need to measure.

```
TASK(Task1) {
    Os_StopwatchTickType start,finish,correction;
    ...
    #if defined(OS_TIME_MONITORING)
        /* Get time for Os_GetExecutionTime() call itself. */
        start = Os_GetExecutionTime();
        finish = Os_GetExecutionTime();
        correction = finish - start -
            Os_Cbk_GetStopwatchUncertainty();
        /* Measure resource lock time. */
        start = Os_GetExecutionTime();
    #endif
    /* The section of code to measure */
    GetResource(Resource1);
    /* Critical section. */
    ReleaseResource(Resource1);
    #if defined(OS_TIME_MONITORING)
        finish = Os_GetExecutionTime();
        /* Calculate amount of time used. */
        used = finish - start - correction +
            Os_Cbk_GetStopwatchUncertainty();
    #endif
}
```

Code Example 15.4: Measuring Blocking Times

15.4 Imprecise Computation

Because the overheads imposed by time monitoring are small, it can be used for production code. You can exploit this fact to perform imprecise computation.

Imprecise computation is useful in applications that interactively converge on a result. For example, you might use Newton-Raphson to converge on a value.

If a task has not traveled down the worst-case path, then it will not have run in the worst-case execution time. If this is the case, any ‘spare’ CPU cycles available to the task can be used to refine a result. This technique is illustrated in Code Example 15.5.

```
TASK(NewtonRaphson) {
    TickType Budget = CONFIGURED_EXECUTION_BUDGET;
```

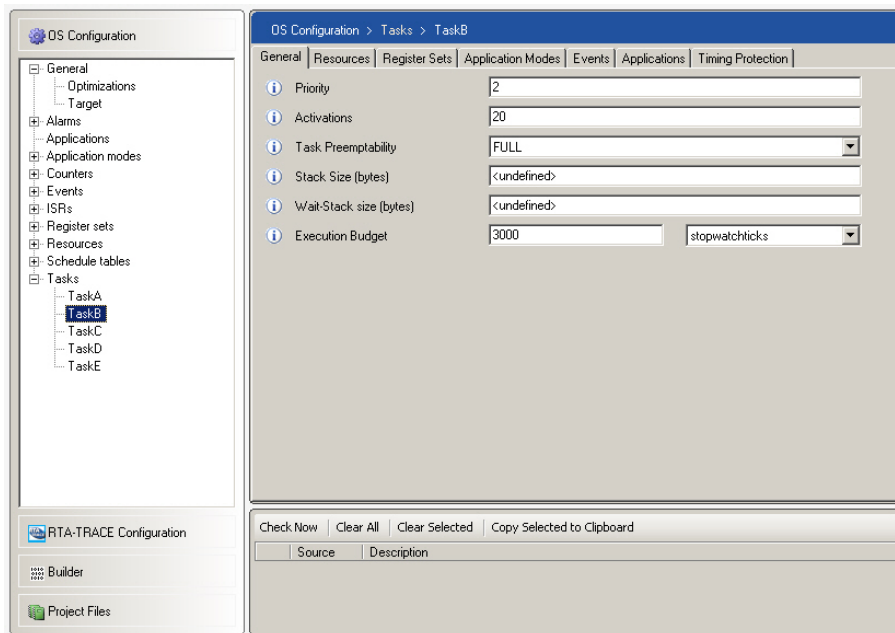


Figure 15.4: Specifying the Execution Time Budgets

```

TickType LoopTime = TIME_FOR_ONE_ITERATION;
...
Result = ...;
while ((Budget - Os_GetExecutionTime()) > LoopTime) {
    /* Perform iterative refinement of output. */
    Result = Result - (Function(Result)/Derivative(Function,
        Result));
}
...
}

```

Code Example 15.5: Imprecise Computation

15.5 Monitoring Execution Times against Budgets

Time monitoring also allows you to set budgets for execution times and let RTA-OS3.x check for violations at runtime. The execution time budgets for each task and Category 2 ISR can be set in your application. These values are optional and do not have to be supplied. Configuration of an execution budget is shown in Figure 15.4.

The type of the budget value can be set as 'clock time' or in terms of stopwatch ticks or CPU cycles. RTA-OS3.x uses the target timing characteristics to perform any necessary conversions. Figure 15.5 shows how these values are set.

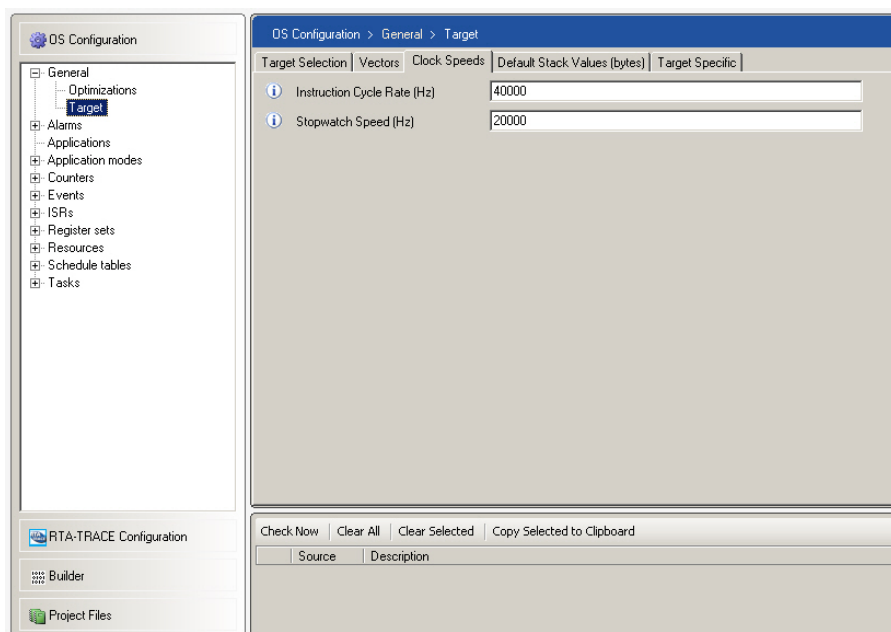


Figure 15.5: Specifying the Instruction Rate and Stopwatch Speed

When time monitoring is enabled, RTA-OS3.x will check to see whether tasks or Category 2 ISRs consume more time than is specified in the budget. If the budget is exceeded, then RTA-OS3.x will call the `Os_Cbk_TimeOverrunHook()` when the task terminates (or, in the case of an extended task, when it calls `WaitEvent()`). This allows you to log the budget overrun. As budgets are checked on a context switch there is the potential for a task or Category 2 ISR to overrun by a large margin before this is actually detected. Figure 15.6 shows what happens when a task overruns.



The `Os_Cbk_TimeOverrunHook()` is mandatory if time monitoring is configured in RTA-OS3.x. Your program will not link correctly if you do not provide this function.

The prototype for `Os_Cbk_TimeOverrunHook()` is shown in Code Example 15.6.

```
#ifndef OS_TIME_MONITORING
FUNC(void, OS_APPL_CODE)
    Os_Cbk_TimeOverrunHook(Os_StopwatchTickType Overrun) {
    /* Log budget overruns. */
    }
#endif
```

Code Example 15.6: The `Os_Cbk_TimeOverrunHook` Prototype

You should be aware that, for extended tasks, the execution time is reset to zero at the start of the task and when resuming from `WaitEvent()`. Nor-

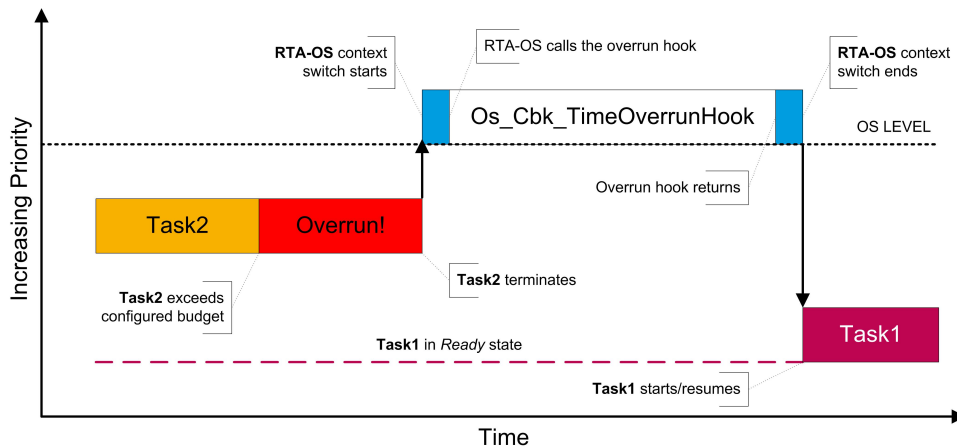


Figure 15.6: Call of the `Os_Cbk_TimeOverrunHook()`

mally the budget is used to check the execution time between consecutive `WaitEvent()` calls.

You should also be aware that the execution time is only sampled by RTA-OS3.x when a task is preempted by another task or ISR or when the task/ISR terminates.



In some unusual circumstances, it is possible for a budget overrun to be missed. This could happen when the interval between preemptions approaches the maximum interval that can be measured by a `Os_StopwatchTickType`. The range of a `Os_StopwatchTickType` is target dependent, but is normally 2^{16} or 2^{32} .

15.6 Summary

- RTA-OS3.x provide in-kernel features that allow you to measure the execution time of tasks and ISRs at runtime.
- You need to provide access to a free-running timer for RTA-OS3.x to use as a stopwatch.
- The worst-case execution time of tasks and ISRs is logged automatically.
- Arbitrary measurements can be made using the `Os_GetExecutionTime()` API.
- If an execution budget is specified for a task or ISR, then RTA-OS3.x will automatically monitor the task or ISR and generate an error at context switch time if the budget is exceeded.

16 Using an ORTI-Compatible Debugger

ORTI is an acronym that stands for ‘OSEK Run Time Interface’. ORTI was designed to provide a standardized and extensible way for an OSEK operating system to provide internal details of its behavior to a debugger. The design of the ORTI is sufficiently general that it can support operating systems other than OSEK and in RTA-OS3.x ORTI support is provided for OSEK OS and AUTOSAR OS features.

ORTI provides a small language that captures two things:

1. how to find objects and variables within the running operating system; and
2. how to interpret or display their values.

An OS can generate an ORTI file that contains a description of this information that the debugger can use. This means that ORTI is like a symbol table - telling the debugger which things in memory mean which objects in the OS.



ORTI is not supported by all debuggers. A list of compatible ORTI debuggers is provided in the Target/Compiler Port Guide for your port. If no ORTI-compatible debugger is listed, then ORTI generation is not support for your debugger. RTA-OS3.x can, in most cases, be customized to support ORTI for unsupported debuggers. Please contact ETAS for further details.

In the cases where RTA-OS3.x can generate an ORTI file for your debugger, this chapter shows you how to configure the generation of ORTI information for your debugger so that, during execution of the application, you can observe values of key operating system variables for applications based on RTA-OS3.x.

For details of how to view ORTI information at runtime you should consult your debugger documentation.

16.1 Development Process

The following steps describe how to use ORTI with your program.

Step 1 Use `rtaoscfg` to enable ORTI debugger support. As ORTI is target-specific, the configuration is done in the “Target Specific” settings. Figure 16.1 shows how this is done.

Step 2 Build the RTA-OS3.x library. The kernel is instrumented with ORTI support when generated. The ORTI file that you need for your debugger is generated as a file called `<projectname>.orti`.

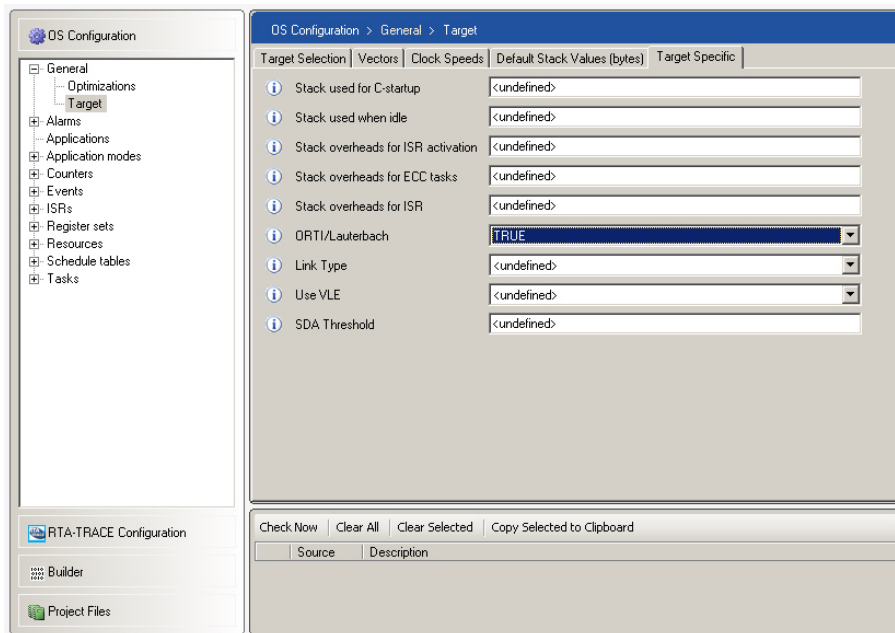


Figure 16.1: Enabling ORTI Support

Step 3 Build the application.

Step 4 Start the debugger, load the application and then load the ORTI file. For details of how to do this, please consult the documentation for your debugger.

The debugger will then display the information shown by the ORTI file. The format of this information depends upon the debugger.

16.2 Intrusiveness

ORTI relies upon reading values from the memory of the running application. This means that the presence of ORTI can affect the operation of the application. It is useful to know the extent to which this might happen. ORTI can acquire data via four routes:

1. Constant values within the ORTI file. These are used for quantities that will not change during the execution of an application. These have no impact on the running application.
2. Values generated as part of the normal operation of the application. Data is read from variables that would be present even if ORTI were not. These have no additional impact on the application.
3. Values generated specifically for ORTI support. Such variables constitute a very small extra overhead in the application.

4. Constants generated only for ORTI support. This data amounts to a small overhead in the application. These constants are only generated for debuggers that cannot obtain the information by other means. They are only present when you specify that you are using a debugger, so you may wish to disable debugger support in your final production release.

16.3 Validity

Many of the values reported by ORTI are simply those contained in the application's memory. Using ORTI to inspect the system before it has been fully initialized will lead to misleading results. RTA-OS3.x is fully initialized when, as a result of calling `StartOS()`, `Os_Cbk_Idle()` the first task or Category 2 ISR is entered.

Care should be taken where a variable may be cached in a register for a significant portion of its lifetime, especially in the case of register-rich processors. ORTI can only look at the data stored in the variable's memory location. This could be out of date if the register-based copy has been updated recently.

16.4 Interactions

The ORTI output will be correct when the program is stopped at a breakpoint that is:

- In code executed by a task or Category 2 ISR that is outside of any AUTOSAR OS API call.

The ORTI output may be misleading if the application is stopped at a breakpoint that is:

- Within an AUTOSAR OS API call.
- In code executed by a Category 1 interrupt handler.

The output may be misleading because the OSEK data used by ORTI could be in a partially updated state. Normally it is possible to tell if the program is part way through an AUTOSAR OS call by the debugger reporting the name of the function in which the processor stopped.

On a platform with more than two interrupt priority levels, however, a Category 1 interrupt can occur part way through an OSEK call. If the program is stopped at a breakpoint in a Category 1 interrupt handler, it is necessary to use the debugger's stack trace facility to determine the name of the function that was interrupted. The ORTI output can be relied upon, provided that the Category 1 interrupt did not occur within an AUTOSAR OS API call.

16.5 Summary

- RTA-OS3.x can optionally generate ORTI information for use with a third-party ORTI compatible debugger.
- ORTI support is port-specific functionality. Additional details on the exact nature of ORTI support for your port can be found in the relevant *Target/Compiler Port Guide*.

17 RTA-TRACE2.1 Integration

RTA-TRACE2.1 is a software logic analyzer for embedded systems which provides a set of services to assist in debugging and testing a system. RTA-TRACE2.1 also has the ability to see exactly what is happening in a system at runtime with a production build of the application software.



RTA-TRACE2.1 is a separate product to RTA-OS3.x and is not supplied with your RTA-OS3.x installation. For further details about how to obtain RTA-TRACE2.1 please contact your local ETAS Sales Office (see Section 22.2).

RTA-TRACE2.1 logs trace records to an on-target trace buffer. Each trace record contains information about what happened, when it happened and which object(s) were involved. RTA-TRACE2.1 relies on an instrumented OS to gather tracing data. While hand-instrumentation is possible, **rtaosgen** can automatically add RTA-TRACE2.1 instrumentation to the generated OS kernel. This chapter explains how to use the RTA-TRACE2.1 configuration editor provided with the **rtaoscfg** tool. Section 17.1 describes the basic configuration. RTA-TRACE2.1 also provides extensive control on which data is traced and allows you to configure user-defined trace information.

Further details about RTA-TRACE2.1 are provided in the RTA-TRACE2.1 user documentation. However, you should note the following:

- the information presented in Sections 17.1 and 17.3 augments the information provided in your *RTA-TRACE2.1 Configuration Guide* for configuration with RTA-OS3.x's **rtaoscfg** tool.
- RTA-OS3.x makes some changes to how the RTA-TRACE2.1 ECU link works. The information presented in Section 17.4 augments the information provided in your *RTA-TRACE Configuration Guide*.
- for RTA-OS3.x, all RTA-TRACE2.1 API calls, callbacks, macros and types adopt the AUTOSAR naming convention. The changes are as follows:

API Feature	RTA-TRACE2.1	RTA-TRACE2.1 with RTA-OS3.x
Call	<name>	Os_<name>
Callback	osTrace<name>	Os_Cbk_Trace<name>
Type	osTrace<name>	Os_Trace<name>
Macro	OSTRACE_ENABLED	OS_TRACE

A complete reference for the modified RTA-TRACE2.1 API is provided in the *Reference Guide*.

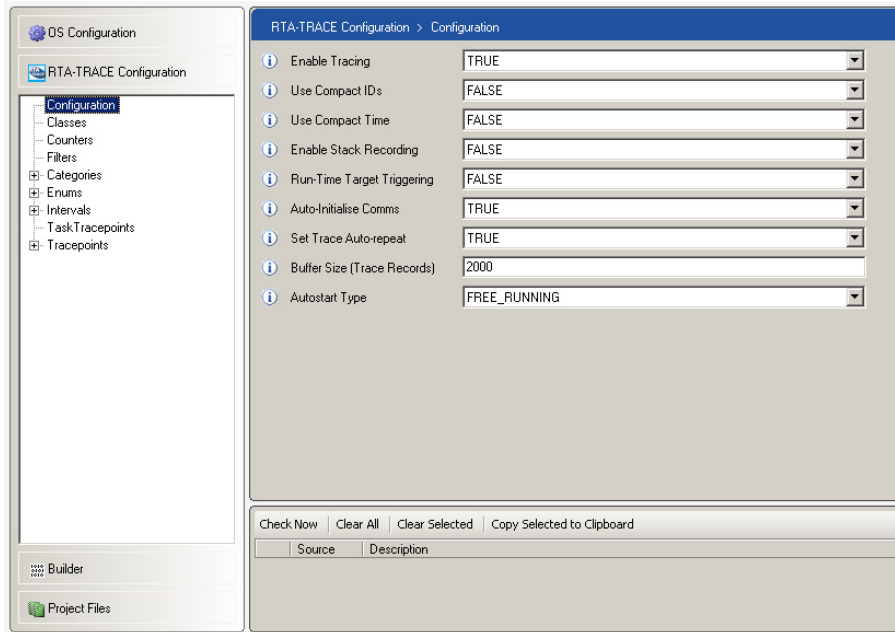


Figure 17.1: Configuring RTA-TRACE2.1

17.1 Basic Configuration

The basic configuration parameters RTA-TRACE2.1 are shown in Figure 17.1.

Enable Tracing selects whether RTA-TRACE2.1 instrumented is added to RTA-OS3.x or not. If this is not set, then no instrumentation is added to RTA-OS3.x.

Use Compact IDs selects compact trace format which reduces the size of a trace record stored in the trace buffer:

Identifier	Regular IDs	Compact IDs
Task Tracepoint	12-bit (max 4096 IDs)	4-bit (max 16 IDs)
Tracepoint	12-bit (max 4096 IDs)	8-bit (max 256 IDs)
Interval	12-bit (max 4096 IDs)	8-bit (max 256 IDs)
OS Objects	16-bit (max 65536 IDs)	8-bit (max 256 IDs)

For most common applications it is safe to use compact identifiers.

Use Compact Time selects compact (16-bit) or extended (32-bit) time format. This option may not be available for every RTA-OS3.x port.

Enable Stack Recording selects whether or not to record stack usage or not. When enabled, this logs two trace records for each trace event: one for the event itself and another for the stack size. Enabling this option therefore doubles the amount of trace data that is recorded.

Run-Time Target Triggering selects whether or not runtime target triggering is available.

Auto-Initialize Comms selects whether the RTA-TRACE2.1 communications link is initialized automatically during StartOS(). Setting this configuration item to TRUE means that RTA-OS3.x will automatically call Os_TraceCommInit() to initialize the communications link. If set to FALSE then Os_TraceCommInit() must be called elsewhere in your application. This field should be set to FALSE when a debugger link is used to upload trace data from the target to the host PC.

Buffer Size sets the size of the trace buffer reserved on the target for the tracing information. The size is specified in trace records *not* bytes. A trace buffer of 2000 records is recommended as a default setting.

Auto-start Type selects whether tracing is started automatically during StartOS() and which tracing mode is used (Bursting, Free-Running or Triggering). See Section 17.2.

17.2 Controlling RTA-TRACE2.1

RTA-TRACE2.1 can be used in three different modes:

Bursting mode treats the buffer as a *linear* buffer and logs trace data until the buffer is full. When the buffer is full, tracing stops and the buffer is made available for uploading to the RTA-TRACE2.1 host PC. This is useful for capturing a 'one-shot' log of data. RTA-TRACE2.1 is started in this mode using Os_StartBurstingTrace(). Tracing can be automatically re-started after the upload if the call Os_SetTraceRepeat(TRUE) has been made.

Free-running mode treats the buffer as a *circular* buffer and makes data available for uploading to the RTA-TRACE2.1 host PC as soon as it has been logged. If data can be uploaded sufficiently often that the buffer is never full, then free-running tracing provides a continuous stream of trace data. If the buffer becomes full, tracing is suspended until space becomes available again. RTA-TRACE2.1 is started in this mode using Os_StartFreerunningTrace().

Triggering mode treats the buffer as a *circular* buffer and logs continuously. If the buffer overflows then old data is overwritten by new data. Data is not made available for upload until one or more user-specified triggers occur. When a trigger occurs, a number of pre-trigger trace records is locked in the buffer (the number of pre-trigger records to be stored is user-specified) and tracing continues until a number of post-trigger trace records has been logged (the number of post-trigger records is

also user-specified). When the post-trigger number of records has been logged, tracing stops and the buffer is made available for uploading to the RTA-TRACE2.1 host PC. RTA-TRACE2.1 is started in this mode using `Os_StartTriggeringTrace()`. The pre and post-trigger windows are set using `Os_SetTriggerWindow(pre,post)`. Tracing can be automatically re-started after the upload if the call `Os_SetTraceRepeat(TRUE)` has been made.

If you have configured RTA-TRACE2.1 to auto-start then RTA-OS3.x will make the correct RTA-TRACE2.1 `Os_Start...()` API automatically during `StartOS()`. If RTA-TRACE2.1 is already running when the call to `Os_Start...()` is made then the trace buffer is cleared and RTA-TRACE2.1 re-starts in the chosen mode.

The `Os_StopTrace()` API call is used to stop RTA-TRACE2.1.

RTA-OS3.x defines the macro `OS_TRACE` when RTA-TRACE2.1 is enabled. You can use this macro to conditionally compile RTA-TRACE2.1 code into your application as shown in Code Example 17.1.

```
FUNC(void, OS_APPL_CODE) StartupHook(void)
    ...
    #ifdef OS_TRACE
        SetTraceRepeat(TRUE);
        StartBurstingTrace();
    #endif
    ...
}
```

Code Example 17.1: Using the `OS_TRACE` macro

17.2.1 Controlling with Objects are Traced

By default, RTA-TRACE2.1 traces every type of OS object. Sometimes this is not appropriate - you may be interested in only a subset of tasks or you may need to reduce the amount of data being logged because your data-link has low bandwidth.

RTA-TRACE2.1 allows you control over data collection using *classes* and *filters*.

Classes

RTA-TRACE2.1 groups trace objects into classes. By default, all classes are traced at runtime. However, to minimize the amount of trace data that is gathered (and therefore minimize the amount of time spent uploading data) you might choose to switch off some classes of tracing.

Each class can be configured as:

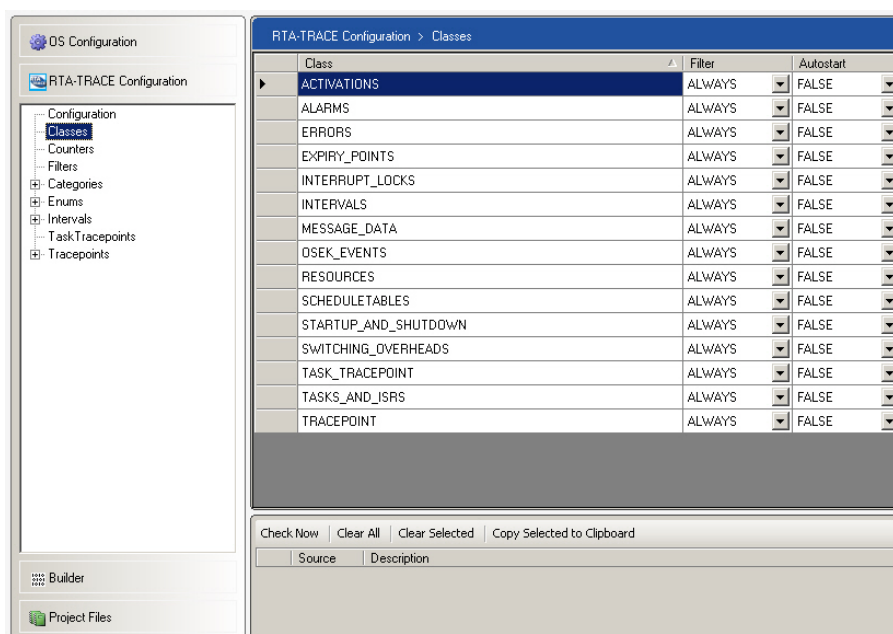


Figure 17.2: Configuring RTA-TRACE2.1 classes

Always the class is always traced.

Never the class is never traced.

Runtime the tracing of the class can be enabled/disabled at runtime using the API calls `Os_EnableTraceClasses()` and `Os_DisableTraceClasses()`.

Figure 17.2 shows how trace classes can be configured.

Any trace class configured as runtime is disabled when RTA-TRACE2.1 starts. However, it is possible to set the runtime classes to be auto-started when RTA-TRACE2.1 starts by setting the Autostart value to TRUE.

Filters

Filters allow individual Tasks and ISRs to be excluded from tracing. As with trace classes, all Tasks and ISRs are traced by default, but can be configured as:

Always the task/ISR is always traced.

Never the task/ISR is never traced.

Runtime the tracing of the task/ISR is controlled by the runtime state of the `OS_TRACE_TASKS_AND_ISR_CLASSES`.

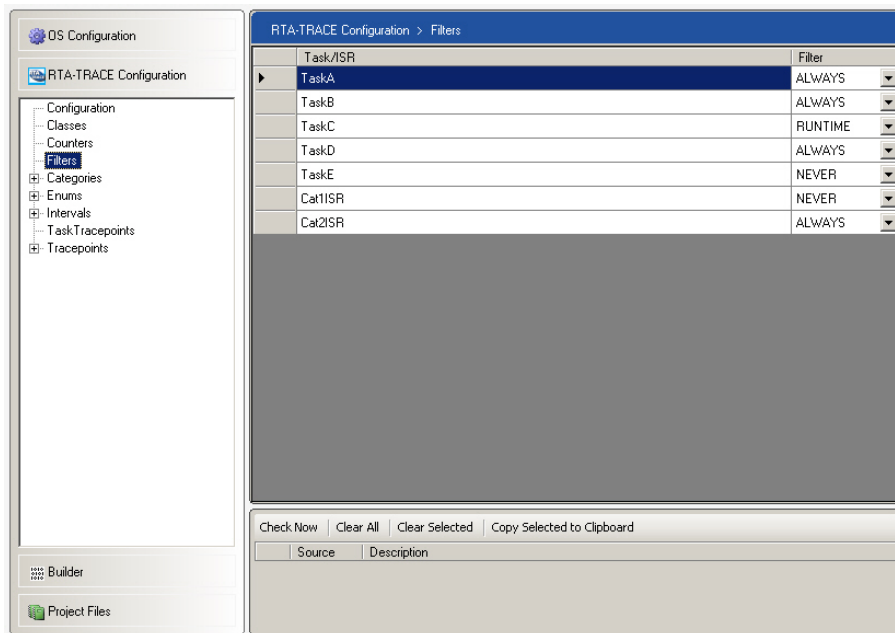


Figure 17.3: Configuring RTA-TRACE2.1 filters

Figure 17.3 shows how trace filters can be configured.

The setting of the `OS_TRACE_TASKS_AND_ISRS_CLASS` is applied *before* a filter is applied. This means that filter settings for a task/ISR interact with trace classes in the following way:

Class Setting	Filter	Task/ISR Traced?
Never	Never	X
Never	Runtime	X
Never	Always	X
Runtime [Disabled]	Never	X
Runtime [Disabled]	Runtime	X
Runtime [Disabled]	Always	X
Runtime [Enabled]	Never	X
Runtime [Enabled]	Runtime	✓
Runtime [Enabled]	Always	✓
Always	Never	X
Always	Runtime	✓
Always	Always	✓

17.3 User-Defined Trace Objects

RTA-TRACE2.1 provides 3 different types of objects that you can configure to help with debugging your application:

260 RTA-TRACE2.1 Integration

Tracepoints are used to log arbitrary data values (for example the value of a variable or content of a data structure) to the trace buffer. Each tracepoint is logged with a timestamp so you can see on the RTA-TRACE2.1 visualization the value of a data item at a moment in time. A tracepoint can be logged from anywhere in the application.

Task Tracepoints are similar to tracepoints but are displayed on the RTA-TRACE2.1 visualization next to the task which logs them

Intervals are used to measure durations of time. An interval has a start and an end marker that can be logged from anywhere in your application. Intervals are particularly useful for measuring end-to-end response times over multiple tasks during program execution.

The following sections describe how to configure these user-defined objects and how to control whether or not they are logged at runtime.

17.3.1 Tracepoints

Each tracepoint requires a unique integer identifier. The maximum number of tracepoints that can be configured depends on the setting of “Use Compact IDs” (see Section 17.1). RTA-OS3.x automatically allocates a unique ID for the tracepoint if its ID is set to zero.

Each tracepoint can also be associated with a discrete data value or a block of data. RTA-TRACE2.1 needs to know how to format the data value supplied and this is configured by specifying a format-string (see Section 17.3.5 for more information about format strings). The format-string controls how RTA-TRACE2.1 will display the data value in the RTA-TRACE2.1 GUI. Figure 17.4 shows the configuration of three tracepoints that log data as a signed integer, a hexadecimal value and an unsigned integer respectively.

Any task in the application can log a tracepoint using the following API calls:

- `Os_LogTracepoint()` - log the tracepoint without any associated data;
- `Os_LogTracepointValue()` - log the tracepoint with an associated value;
- `Os_LogTracepointData()` - log the tracepoint with an associated block of data (specified using a base/bound scheme).

For further details, see the *Reference Guide*.

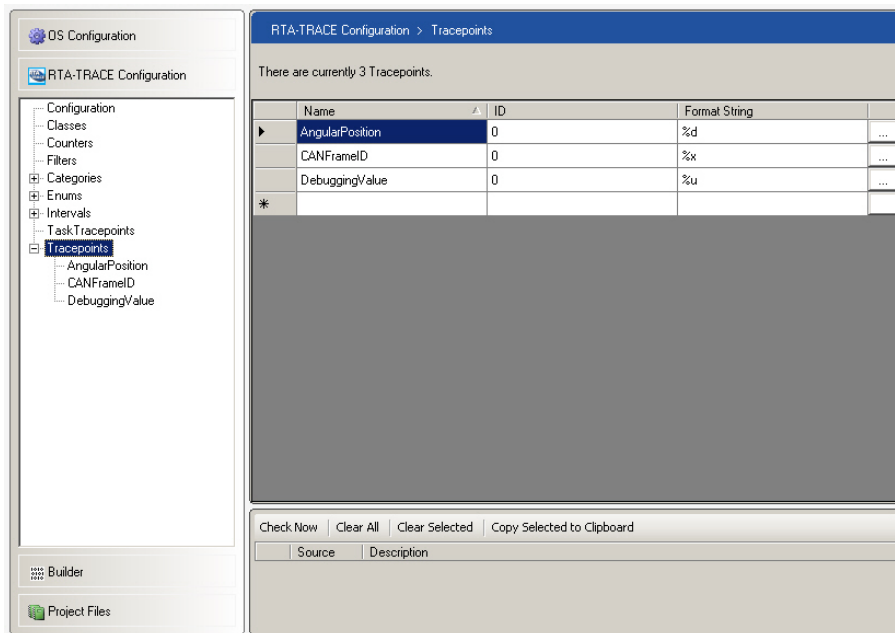


Figure 17.4: Configuring RTA-TRACE2.1 tracepoints

17.3.2 Task Tracepoints

Task-tracepoints are configured just like normal tracepoints. See Section 17.3.1 for further details.

Logging a task tracepoint uses a different set of API calls to normal tracepoints:

- `Os_LogTaskTracepoint()` - log the tracepoint against the calling tasks without any associated data;
- `Os_LogTaskTracepointValue()` - log the tracepoint against the calling tasks with an associated value;
- `Os_LogTaskTracepointData()` - log the tracepoint against the calling tasks with an associated block of data (specified using a base/bound scheme).

For further details, see the *Reference Guide*.

17.3.3 Intervals

Intervals are used to measure arbitrary times in the application, for example an end-to-end response time. Each interval must be named and allocated an unique identifier. As with tracepoints, an interval identifier is an integer which

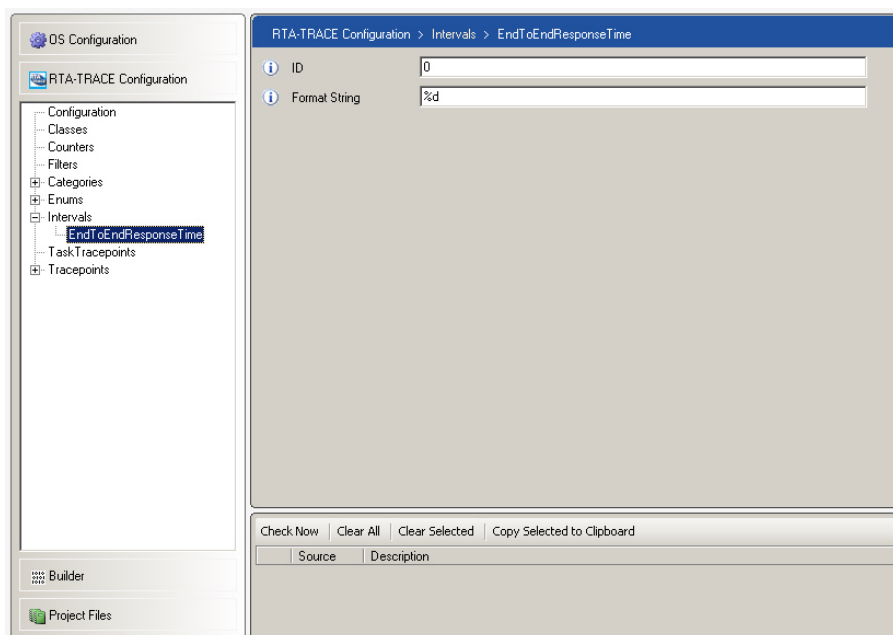


Figure 17.5: Configuring RTA-TRACE2.1 intervals

is specified at configuration time. If a value of zero is configured, RTA-OS3.x automatically allocates a unique identifier to each interval.

Figure 17.5 shows how an interval is configured.

Each interval can also be associated with a discrete data value or a block of data. RTA-TRACE2.1 needs to know how to format the data value supplied and this is configured by specifying a format-string (see Section 17.3.5 for more information about format-strings).

Logging an interval requires you to mark the start and the end of the interval using the following API calls:

- `0s_LogIntervalStart()` - log the start of the interval without any associated data;
- `0s_LogIntervalStartValue()` - log the start of the interval with an associated value;
- `0s_LogIntervalStartData()` - log the start of the interval with an associated block of data (specified using a base/bound scheme);
- `0s_LogIntervalEnd()` - log the end of the interval without any associated data;

- `Os_LogIntervalEndValue()` - log the end of the interval with an associated value;
- `Os_LogIntervalEndData()` - log the end of the interval with an associated block of data (specified using a base/bound scheme).

Calls with and without data or values can be mixed, as shown in Code Example 17.2.

```
#include <Os.h>
#include "ThirdPartyLibrary.h"
TASK(A) {
    ...
    Os_LogIntervalStart(LibraryCallMeasurement, OS_TRACE_CATEGORY_ALWAYS);
    x = CallToLibraryFunction(y,z);
    Os_LogIntervalEndValue(LibraryCallMeasurement, x, OS_TRACE_CATEGORY_ALWAYS);
    ...
}
```

Code Example 17.2: Mixing `Os_LogInterval...` calls

For further details, see the *Reference Guide*.

17.3.4 Controlling which User-Defined Objects are Traced

User-defined objects are logged in the RTA-TRACE2.1 trace buffer at runtime. Each API to log a user-specified object takes a parameter defining the trace category for which is logged:

```
Os_Log[[Task]Tracepoint|Interval[Start|End]][Data|Value](..., Os_TraceCategoriesType
    CategoryMask)
```

Trace categories are user-defined names that allow you control whether a user-defined trace object is traced or not at runtime.

Each category has a category bit-mask. The mask is an integer that represents a unique identifier for the category in the trace buffer. The mask can be set to a specific integer value, but it is recommended that you set the mask to zero and let RTA-OS3.x generate the category mask automatically.



If you choose to set your own mask values then you must ensure that the integer representing the mask is a power of two i.e. 1,2,4,8,16 etc.

As with classes, each trace category can be filtered:

Always the category is always traced.

Never the category is never traced.

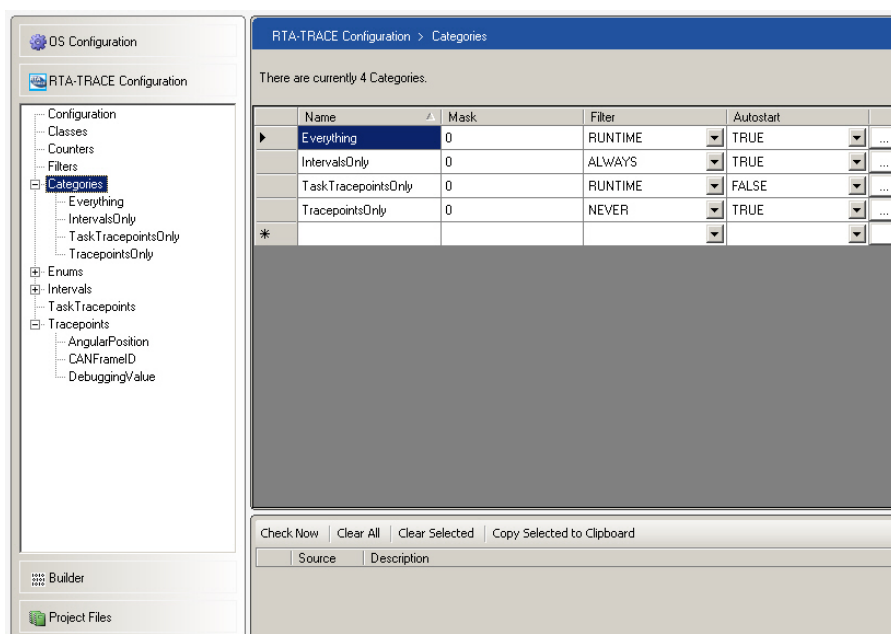


Figure 17.6: Configuring RTA-TRACE2.1 categories

Runtime the tracing of the category can be enabled/disabled at runtime.

By default, runtime trace categories are disabled when RTA-TRACE2.1 starts. The initial categories configuration allows you to control which of the run-time are enabled when tracing starts.

Figure 17.6 shows how trace categories can be configured.

RTA-TRACE2.1 also defines two constant category masks:

1. `OS_TRACE_CATEGORY_ALWAYS` is always be traced.
2. `OS_TRACE_CATEGORY_NEVER` is never be traced.

Runtime control for categories is provided though the RTA-TRACE2.1 API calls `Os_EnableTraceCategories()` and `Os_DisableTraceCategories()`. Each call takes a category mask (or a bit-wise OR of category masks) as input. All user tracing can be disabled by calling `Os_DisableTraceCategories(OS_TRACE_CATEGORY_ALWAYS)` and re-enabled by calling `Os_EnableTraceCategories(OS_TRACE_CATEGORY_ALWAYS)`.

17.3.5 Format Strings

Format strings are used to tell RTA-TRACE2.1 how to display a user-defined trace item's data. Simple numeric data can be displayed using a single format specifier. More complex data, e.g. a C **struct**, can be displayed by repeatedly moving a cursor around the data block and emitting data according to more complex format specifiers.

If a format string is not supplied, data is displayed in the following manner:

- If the data size is no greater than the size of the target's integer type, data is decoded as if "%d" had been specified.
- Otherwise the data is displayed in a hex dump, e.g.

```
0000 00 01 02 03 04 05 06 07 08 09 0a 0b 0c 0d 0e 0f
0010 10 11 12 13 14 15 16 17 18 19 1a 1b 1c 1d 1e 1f
```

The hex dump has a maximum size of 256 bytes.



When format specifiers are given, the target's endian-ness is taken into account. When a hex dump is shown, the target's memory is dumped byte-for-byte. In particular, you may not get the same output from a hex dump as from the %x format specifier.

Rules

Format strings are similar to the first parameter to the C function `printf()`:

- Format strings are surrounded by double-quote (") symbols.
- A format string may contain two types of object: ordinary characters, which are copied to the output stream, and format elements, each of which causes conversion and printing of data supplied with the event.
- A format element comprises a percent sign, zero or more digits and a single non-digit character, with the exception of the %E element.
- The format element is decoded according to the rules in the table below, and the resulting text is added to the output string.
- The special format element %% emits a %.
- In addition to ordinary characters and conversion specifications, certain characters may be emitted by using a 'backslash-escape sequence'. To emit a double-quote " character, \" is used, and to emit a \ character, \\ is used.

- The optional size parameter to integer format specifiers defines the field's width in bytes. Valid values are 1, 2, 4 or 8.



An important difference from `printf()` is that the cursor does not automatically move on from the current field when a field is emitted. This is to facilitate multi-format output of a single field.

Format	Element Meaning
<code>%offset@</code>	Moves the cursor offset bytes into the data. This can be used to extract values from multiple fields in a structure.
<code>%[size]d</code>	Interpret the current item as a signed integer. Output the value as signed decimal.
<code>%[size]u</code>	Interpret the current item as an unsigned integer. Output the value as unsigned decimal.
<code>%[size]x</code>	Interpret the current item as unsigned integer. Output the value as unsigned hexadecimal.
<code>%[size]b</code>	Interpret the current item as an unsigned integer. Output the value as unsigned binary.
<code>%enum[:size]E</code>	Interpret the current item as an index into the enumeration class who's ID is enum. Emit the text in that enumeration class that corresponds with the item's value. The enumeration class should be defined using ENUM directives.
<code>%F</code>	Treat the current item as an IEEE 'double'. Output the value as a double, in exponent format if necessary.
<code>%?</code>	Emit in the form of a hex dump.
<code>%%</code>	No conversion is carried out; emit a %.

Enumerations

Sometime you may want RTA-TRACE2.1 to display symbolic data for a a given trace value. This is possible in a number of ways with format strings, but one possibility is to use a value to reference an enumeration of symbolic values. Each enumeration you need must be configured before it can be referenced from a format string.

An enumeration is given a name and contains a set of name/value pairs that define the mapping between the value and the associated symbolic name. Figure 17.7 shows how an enumeration of 'Gear' has been configured with a simple mapping between an integer value and the symbolic names of the gears.

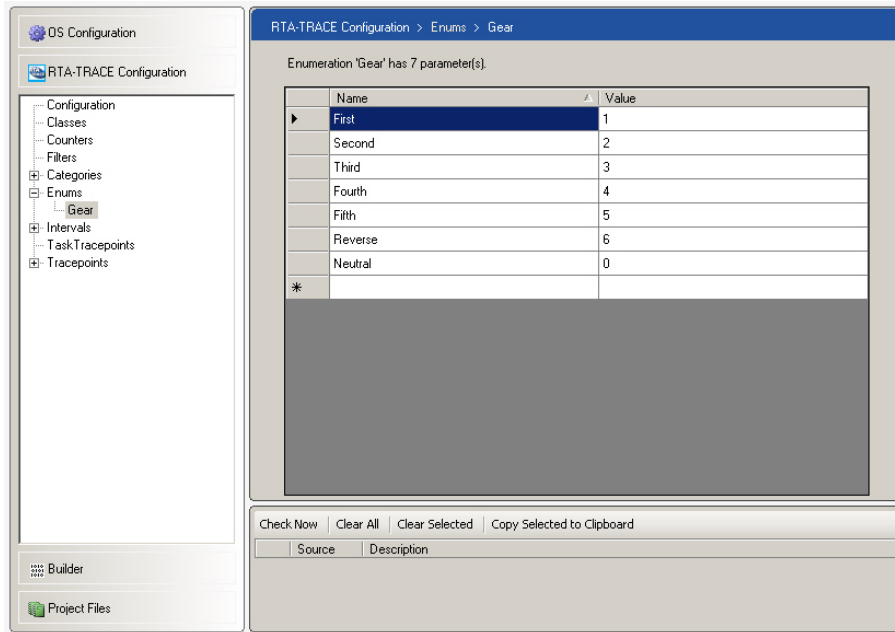


Figure 17.7: Configuring RTA-TRACE2.1 enumerations

Example Format Strings

Description	Format	Example	Notes
A native integer displayed in decimal and hexadecimal	"%d 0x%x"	10 0xA	The "0x" is not emitted by the %x format specifier but is specified in literal characters in the string. Absence of size specifier means the target's integer size is assumed. This example is a 16-bit processor.
A single unsigned byte representing a percentage.	"%1u%%"	73%	Use of size specifier of 1 byte. Use of %% to emit %.
struct { int x; int y; }; on a 32-bit processor.	"(%d,%4@%d)"	(20,-15)	Use of %offset@ to move to byte-offset within the structure.
A value of type enum e_Rainbow, (defined as the colors of the rainbow!)	"%1E"	Yellow	The number 1 refers to the ID of the enum class in the ENUM directives, not to the width of the field.

17.4 ECU Links

RTA-TRACE2.1 provides two standard ways to get data from the ECU to the host PC:

- **Debugger Link** - This is a passive data link - it does not require any supporting code in your application. However, you will need to use your debugger¹ to "pull" the contents of the trace buffer from the target to the PC running the RTA-TRACE2.1 Server.
- **Serial Link** - This is an active link - you need to provide code in your application to "push" the contents of the trace buffer to from the target to the PC running the RTA-TRACE2.1 Server. Both polled and interrupt-driven serial communication is possible.

¹A debugger is not supplied with RTA-OS3.x or RTA-TRACE2.1. A list of compatible debuggers can be found in the *Target/Compiler Port Guide* for your port.

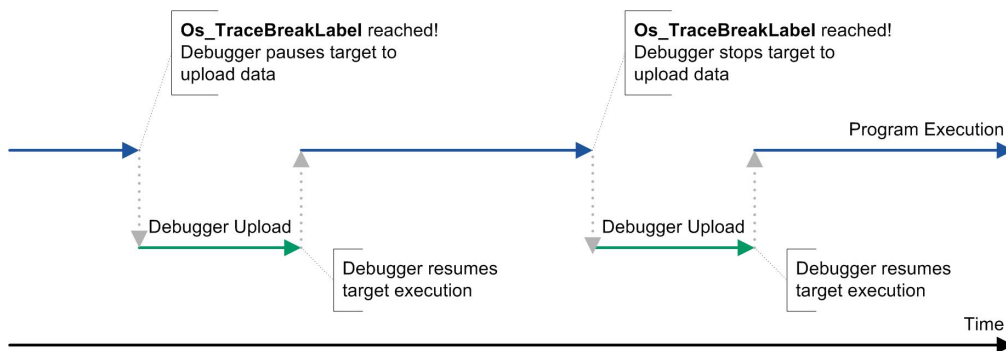


Figure 17.8: Impact of the debugger link with target execution

Other data links may be available - please contact ETAS for details.

The following sections describe how to use the standard data-links in your application.

17.4.1 Debugger Links

The debugger link only transfers data to the RTA-TRACE2.1 server once there is a full buffer (or a full trigger window in the case of triggering mode) available for transmission.

When the buffer is full, RTA-TRACE2.1 calls the function `Os_TraceBreakLabel()`. You should use your debugger to place a breakpoint on this function so that each time the trace buffer is full, the target is paused and you can then upload the contents of the variable `Os_TraceBuffer[]` to the debugger. Many debuggers can be scripted to perform these steps automatically.

The RTA-TRACE2.1 server accepts data in two formats:

1. Lauterbach format
2. CrossView format

These formats are described in the *RTA-TRACE ECU Link Guide*.

Using the debugger link may impact interaction with the target. Each time the trace buffer is full then the target is paused by the debugger and only resumed once the trace buffer has been uploaded as shown in Figure 17.8.

The debugger link is therefore best used when you need to capture a snapshot trace, such as those obtained using bursting or triggering mode.

17.4.2 Serial Links

The serial link needs your application code to actively transmit data to the RTA-TRACE2.1 server over a serial connection. RTA-TRACE2.1 manages the emptying of the trace buffer and the state of the serial connection and uses a set of callbacks to control the serial hardware itself.

Initializing the Serial Device

Serial communication is initialized by called `Os_TraceCommInit`. The call must be made before RTA-TRACE2.1 is started. If “Auto-Initialize Comms” has been configured then RTA-OS3.x will automatically call this API during `StartOS()`.

`Os_TraceCommInit` requires you to provide the `Os_Cbk_TraceCommInitTarget` callback that should initialize the serial hardware for your target. The callback should return `OS_TRACE_STATUS_OK` if the initialization was successful and `OS_TRACE_STATUS_COMM_INIT_FAILURE` otherwise. Code Example 17.3 shows how the callback might look in your code.

```
FUNC(Os_TraceStatusType, OS_APPL_CODE)
    Os_TraceCommInitTarget(void)
{
    /* Set baud rate */
    SERIAL_BAUD_REGISTER = 9600; /* baud */

    /* Set transmit enable bit in control register 2 */
    SERIAL_CTRL_REGISTER |= TE_BIT;

    return OS_TRACE_STATUS_OK;
}
```

Code Example 17.3: Initializing the serial hardware

Data Transmission

Data transmission is a two stage process:

1. Check if there is any data to transmit.
2. If data is available then transmit it.

RTA-TRACE2.1 will automatically detect when the buffer is full and use this information to make the trace buffer available for transmission. This mechanism is sufficient when using bursting or triggering mode.

If you are free-running mode then this behavior may result in you losing trace records when the buffer becomes full because RTA-TRACE2.1 will suspend

tracing until the buffer is emptied. However, you can tell RTA-TRACE2.1 to check for available data and make it ready for transmission *before* the trace buffer becomes full by calling `Os_CheckTraceOutput()`.

To ensure that the trace buffer is uploaded as quickly as possible you should call `Os_CheckTraceOutput()` as often as you can. A good place to make the call² is in the `Os_Cbk_Idle()` callback as shown in Code Example 17.3 shows how the callback might look in your code.

```
FUNC(boolean, OS_APPL_CODE) Os_Cbk_Idle(void)
{
    #ifdef OS_TRACE
        Os_CheckTraceOutput();
    #endif
    return TRUE;
}
```

Code Example 17.4: Checking if data is available for transmission

`Os_CheckTraceOutput()` has a short execution time so there is no significant overhead on the application if it gets called more frequently than strictly necessary.

When data is available for transmission RTA-TRACE2.1 signals this through the callback `Os_Cbk_TraceCommDataReady()`. An implementation of this callback is provided in the RTA-OS3.x library. However, when using a serial link it is recommended that you provide your own implementation of `Os_Cbk_TraceCommDataReady()` to start the communication process.

RTA-TRACE2.1 provides two mechanisms to transmit data from the target:

1. Asynchronous Dump - transmit the available buffer in single operation.
2. Byte-wise - transmit the available buffer a byte at a time.

Asynchronous Dump

A trace buffer dump is made using the RTA-TRACE2.1 API `Os_TraceDumpAsync()`. The call takes a function name as a parameter. The function must be able to transmit a byte of data over the serial line.

You should only call `Os_TraceDumpAsync()` when there is data available which means the call should be made from the `Os_Cbk_TraceCommDataReady()` callback. This means you need to provide an implementation of `Os_Cbk_TraceCommDataReady()` to override the one provided in the RTA-OS3.x library. Code Example 17.5 shows the implementation of a trace buffer dump.

²Assuming that there is enough slack time available in your system that the callback runs.


```

void TransmitByte(uint8 val) {
    while(!tx_ready) {/* Wait for space in serial device */}
        transmit(val) ;
    }

FUNC(void, OS_APPL_CODE) Os_Cbk_TraceCommDataReady(void) {
    Os_TraceDumpAsync(TransmitByte);
}

```

Code Example 17.5: Implementing a Trace Buffer Dump

Byte-Wise Transmission

Byte-wise transmission feeds one byte of the trace buffer at a time to the serial device. The RTA-TRACE2.1 API call `Os_UploadTraceData()` is made to transfer a byte of the trace buffer to the serial device. You need to make the call often enough to ensure that data is transmitted. The call can be made from anywhere in your application code, but you need to be aware that if it is made from a higher priority task then it will affect the responsiveness of lower priority tasks.

If there is data waiting then `Os_UploadTraceData()` attempts to queue the byte for sending as follows:

1. The callback `Os_Cbk_TraceCommTxReady()` is made to check if there is space in the serial device's buffer.
2. If space is available, then the callback `Os_Cbk_TraceCommTxStart()` is made to signal that transmission is about to start. If there is no space then the call returns immediately.
3. The callback `Os_Cbk_TraceCommTxByte()` is made to actually transmit the byte
4. The callback `Os_Cbk_TraceCommTxEnd()` is made to signal that transmission has completed

The RTA-TRACE2.1 serial ECU link can operate in either interrupt or polling mode. Interrupt mode prioritizes communication at the expense of the application's timing characteristics. Polling mode prioritizes the application's timing behavior at the possible risk of some loss of trace data. In general it is recommended to use polling mode and if necessary set target-side triggers and filters to generate a smaller volume of data (see Section 17.2.1).

Whichever transmission mode you choose, you always need to provide implementations of the four callback functions:

1. `Os_Cbk_TraceCommTxReady()`
2. `Os_Cbk_TraceCommTxStart()`
3. `Os_Cbk_TraceCommTxByte()`
4. `Os_Cbk_TraceCommTxEnd()`

The following sections explain what these callbacks need to do and how to construct the polled or interrupt mode driver.

Polling Mode

Polled mode requires that you make regular calls to `Os_CheckTraceOutput()` and `Os_UploadTraceData()` to ensure data in the trace buffer is made available to upload and then uploaded before the trace buffer becomes full. Typically, it is sufficient in most system to do this from RTA-OS3.x's idle mechanism, `Os_Cbk_Idle()`, so that when your application has nothing else to do it can be uploading the trace buffer as a 'background' activity.



If you are using RTA-TRACE2.1 in free-running mode then you must call `Os_CheckTraceOutput()` regularly. If this is not called in a timely fashion then `Os_UploadTraceData()` will not have any data to transmit. Failing to call `Os_CheckTraceOutput()` regularly will result in the trace buffer becoming full. If this occurs then RTA-TRACE2.1 will suspend tracing until the buffer has been emptied or partially emptied and `Os_CheckTraceOutput()` has been called.

You need to provide implementations of the callbacks `Os_Cbk_TraceCommTxReady()` and `Os_Cbk_TraceCommTxByte()`. It is not necessary to do anything for `Os_Cbk_TraceCommTxStart()` and `Os_Cbk_TraceCommTxEnd()` callbacks, but 'dummy' implementations need to be provided.

Code Example 17.6 shows a typical polled driver implementation.

```

FUNC(BOOLEAN, OS_APPL_CODE) Os_Cbk_Idle(VOID)
{
    #ifdef OS_TRACE
        Os_CheckTraceOutput();
        Os_UploadTraceData();
    #endif
    return TRUE;
}

FUNC(VOID, OS_APPL_CODE) Os_Cbk_TraceCommTxStart(VOID){
    /* Do nothing */
}

```

```

FUNC(boolean, OS_APPL_CODE) Os_Cbk_TraceCommTxReady(void){
    return (serial_device_has_space());
}

FUNC(void, OS_APPL_CODE) Os_Cbk_TraceCommTxByte(uint8 byte){
    serial_device_transmit_byte(byte);
}

FUNC(void, OS_APPL_CODE) Os_Cbk_TraceCommTxEnd(void){
    /* Do nothing */
}

```

Code Example 17.6: Polled Transmission

Interrupt Mode

Trace data throughput can be optimized by using the serial module's 'Transmit Complete' interrupt and a user-supplied interrupt handler that calls `Os_UploadTraceData()`. This means that data transmission takes precedence over task execution. Interrupt mode is therefore best suited to bursting and triggered modes where data transmission takes place after trace recording has stopped.



It is not recommended to use interrupt transmission in free-running mode because handling the transmit complete interrupt will affect the timing behavior of the system.

When RTA-TRACE2.1 detects that the trace data buffer is ready for transmission in the callback `Os_Cbk_TraceCommDataReady()` is called. You must call `Os_UploadTraceData()` to start the transmission of the trace data.

You will need to configure an RTA-OS3.x interrupt (either Category 1 or Category 2) using **rtaoscfg** and provide an implementation of the handler.

As with polled mode, implementations of the callbacks `Os_Cbk_TraceCommTxReady()` and `Os_Cbk_TraceCommTxByte()` are required. The functionality of these callbacks will be identical to the ones you would write for a polled mode driver.

Interrupt mode uses the callbacks `Os_Cbk_TraceCommTxStart()` and `Os_Cbk_TraceCommTxEnd()` to enable and disable the transmit interrupt.

Code Example 17.7 shows a typical polled driver implementation.

```

ISR(SerialTxInterrupt){
    Os_UploadTraceData();
    dismiss_serial_tx_interrupt();
}

FUNC(void, OS_APPL_CODE) Os_Cbk_TraceCommDataReady(void) {

```

```

    Os_UploadTraceData();
}

FUNC(void, OS_APPL_CODE) Os_Cbk_TraceCommTxStart(void){
    enable_serial_tx_interrupt();
}

FUNC(boolean, OS_APPL_CODE) Os_Cbk_TraceCommTxReady(void){
    return (serial_device_has_space());
}

FUNC(void, OS_APPL_CODE) Os_Cbk_TraceCommTxByte(uint8 byte){
    serial_device_transmit_byte(byte);
}

FUNC(void, OS_APPL_CODE) Os_Cbk_TraceCommTxEnd(void){
    disable_serial_tx_interrupt();
}

```

Code Example 17.7: Interrupt Transmission

Mode Summary

The following table gives a summary of what needs to be implemented for polling and interrupt-driven modes of operation.

Callback	Polled-Mode	Interrupt Mode
Os_Cbk_TraceCommTxStart	empty	Enable Tx interrupt
Os_Cbk_TraceCommTxReady	Check for space in serial device	Check for space in serial device
Os_Cbk_TraceCommTxByte	Transmit a byte	Transmit a byte
Os_Cbk_TraceCommTxEnd	empty	Disable Tx interrupt
Os_Cbk_TraceCommDataReady	empty	Call Os_UploadTraceData();

17.5 Summary

- RTA-OS3.x can automatically instrument the kernel library to generate RTA-TRACE2.1 profiling information.
- The instrumented kernel logs trace data to an on-target memory buffer.
- The buffer can be emptied by ‘pulling’ the data out using a third-party debugger or by ‘pushing’ the data out over a serial communication link.
- Further information about RTA-TRACE2.1 ships with your RTA-TRACE2.1 product.

18 Protected Systems

The AUTOSAR architecture has been designed to support the integration of software from multiple suppliers onto the same ECU. This has many advantages in vehicle systems design, principally that fewer more powerful ECUs can be used in a vehicle, reducing cost and increasing electrical system reliability.

However, such benefits also bring new challenges. When an ECU contains multiple, logically distinct applications that share the same processor then it is possible that one application may fail due to a fault in another application.

A badly behaved application may (accidentally or deliberately):

- overwrite memory belonging to another application
- access objects used by another application (for example activate a task, cancel an important alarm)
- overwrite memory used by the OS, resulting in system-wide instability or other problems
- use more time than was assumed when the applications were integrated, either by running for too long or by running too often
- use peripherals that it should not access
- try and use OS API calls that have system-wide impact (for example `ShutdownOS()`)

For safety-critical systems, such fault propagation should be minimized (and preferably eliminated entirely). The development of a safety-case is much easier if individual application safety cases can be integrated into an overall safety case. This is only feasible if it can be demonstrated that a fault in one application cannot propagate beyond its own boundary and cause a fault in another, unrelated, application.

Even if a safety-critical system is not being built, application suppliers can only be expected to take responsibility (and liability) for their applications failing if they can be assured that their applications cannot be incorrectly blamed for failures.

To prevent unexpected behavior it is necessary to provide mechanisms to protect applications in the system. AUTOSAR OS R3.x provides three types of protection:

Timing Protection can be used to enforce temporal separation between tasks and ISRs by enforcing pre-defined timing constraints on their runtime behavior. This is described in Chapter 19

Service Protection can be used to prevent applications from using API calls to manipulate the objects owned by other applications. This is described in Section 20

Memory Protection can be used to enforce spatial separation between tasks and ISRs. This prevents one task or ISR from corrupting data used by other tasks and ISRs. This is described in Chapter 21

Service and memory protection need to define the protection boundary (i.e. the scope of the protection scheme). In AUTOSAR OS R3.x the protection boundary is defined by using an *OS-Application*. OS-Applications are described in Chapter 20.1.

Protection is only possible for objects that are managed by the OS. This means that it is not possible to provide any protection against faults in Category 1 ISRs.



It is highly recommended that Category 1 ISRs are not used in protected systems.

18.1 Customized Protection Schemes

Recall from Section 2.1.2 that all OS features are assigned to Scalability Classes in AUTOSAR OS R3.x. The following table summarizes the permitted combinations of protection features as follows:

Feature	SC1	SC2	SC3	SC4
Core OS	✓	✓	✓	✓
Memory Protection	X	X	✓	✓
Service Protection	X	X	✓	✓
Timing Protection	X	✓	X	✓

In addition to the AUTOSAR OS R3.x Scalability Classes, RTA-OS3.x allows the selection of a higher class and then the de-selection of certain features in the class. This gives four additional protection subclasses:

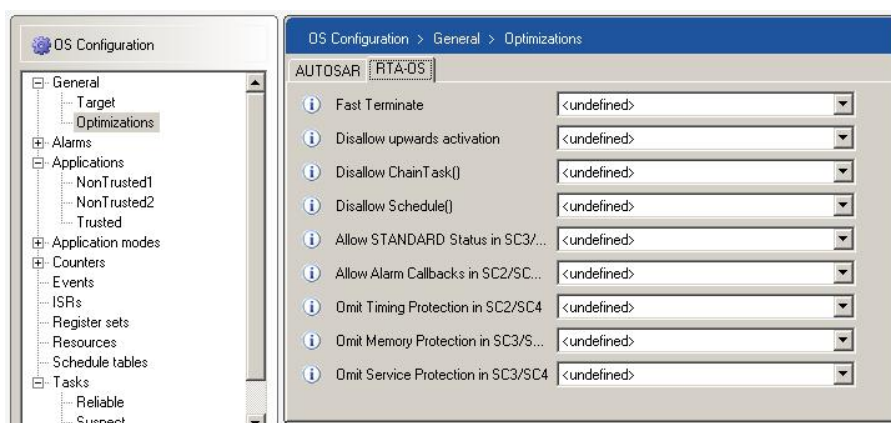


Figure 18.1: Optimizing Scalability Class Features

Feature	SC1	SC2	SC3'	SC3''	SC3	SC4'	SC4''	SC4
Core OS	✓	✓	✓	✓	✓	✓	✓	✓
Timing Protection	X	✓	X	X	X	✓	✓	✓
Service Protection	X	X	X	✓	✓	X	✓	✓
Memory Protection	X	X	✓	X	✓	✓	X	✓

RTA-OS3.x provides this flexibility so that applications can optimize their use of protection features. For example, if you need a system to run tasks in different processor modes because this provides adequate protection against non-privileged tasks accessing critical parts of your hardware, but you do not need memory protection, then this can be achieved by configuring Scalability Class 3 and then choosing to omit memory protection.

Figure 18.1 shows how the Scalability Class optimizations are selected in RTA-OS3.x.

This also means that when you are building critical systems, but your safety case does not require the use of certain OS features, these features can be removed from the OS, avoiding any problems with justifying the use of unused code in the kernel build.

18.2 Handling Protection Errors

When RTA-OS3.x detects that an application has attempted to violate a defined protection boundary, the default reaction is to call `ShutdownOS()`, passing in the type of protection error that has occurred. AUTOSAR OS R3.x defines five protection errors:

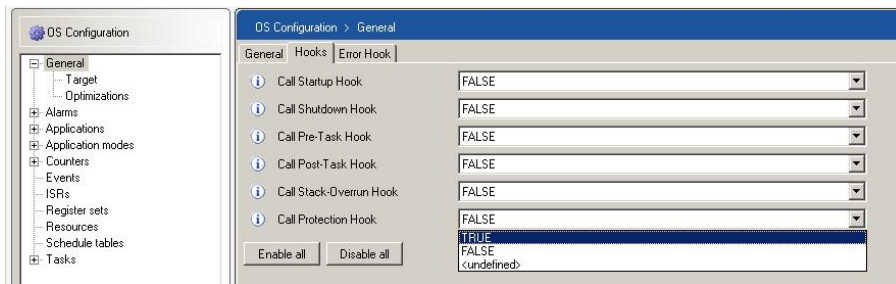


Figure 18.2: Configuring the Protection Hook

Error Code	Description
E_OS_PROTECTION_MEMORY	A memory protection violation has occurred.
E_OS_PROTECTION_TIME	A timing execution time error has occurred. A task or ISR has executed for too long.
E_OS_PROTECTION_ARRIVAL	A timing arrival rate error has occurred. A task or ISR has been activated or arrived more frequently than allowed.
E_OS_PROTECTION_LOCKED	A timing locking error has occurred. A resource may have been locked for too long, or interrupts may have been disabled for too long.
E_OS_PROTECTION_EXCEPTION	A software exception has occurred.

However, it is not always appropriate to shutdown the entire OS. If you have multiple applications executing then you may want to simply shut down the faulty application and leave the correctly functioning applications running. You might even just want to log the error (for example when debugging) but have the OS ignore the problem.

These cases can be handled by configuring and implementing a *Protection Hook* (`ProtectionHook()`) which is called instead of directly calling `ShutdownOS()`. When called, the protection hook is passed the type of error which has occurred so that appropriate corrective action can be taken. Figure 18.2 shows how the protection hook is enabled.



If you enable the protection hook then you must provide an implementation. If you do not provide an implementation then your program will not link correctly. The protection hook runs in the context of the OS and must therefore be trusted code.

The protection hook needs to identify which error has occurred and then tell RTA-OS3.x what action to take in response. AUTOSAR OS R3.x defines a set

of actions, however certain actions are only possible in response to certain errors:

Action	Description
PRO_IGNORE	Ignore the error. This is only possible for E_OS_PROTECTION_ARRIVAL.
PRO_TERMINATETASKISR	Forcibly terminate the task or ISR that caused the error.
PRO_TERMINATEAPPL	Forcibly terminate the application which owns the task or ISR that caused the problem, but do not activate the application's restart task.
PRO_TERMINATEAPPL_RESTART	Forcibly terminate the application that owns the task or ISR that caused the problem. Activate the application's restart task.
PRO_SHUTDOWN	Shutdown the OS by calling ShutdownOS().

The action required is passed back to the OS using the return value of the ProtectionHook(). RTA-OS3.x will do the requested action. The protection hook can also use OS API calls to work out which task or ISR caused the error and to which OS-Application it belongs .

Placing the choice of fault reaction in the user domain provides a high degree of flexibility in how the errors are handled because the action can be determined based on:

- which error has occurred
- the task or ISR that cause the error
- the OS-Application to which the task or ISR belongs

Code Example 18.1 shows you the usual structure of the ProtectionHook().

```

FUNC(ProtectionReturnType, OS_APPL_CODE)
ProtectionHook(StatusType FatalError) {
    ProtectionReturnType Action = PRO_SHUTDOWN;

    if (GetISRID() != INVALID_ISR) {
        /* Always shutdown if a faulty interrupt */
        Action = PRO_SHUTDOWN;
    } else {
        switch (FatalError) {
            case E_OS_PROTECTION_MEMORY:
                Action = PRO_TERMINATETASKISR;

```

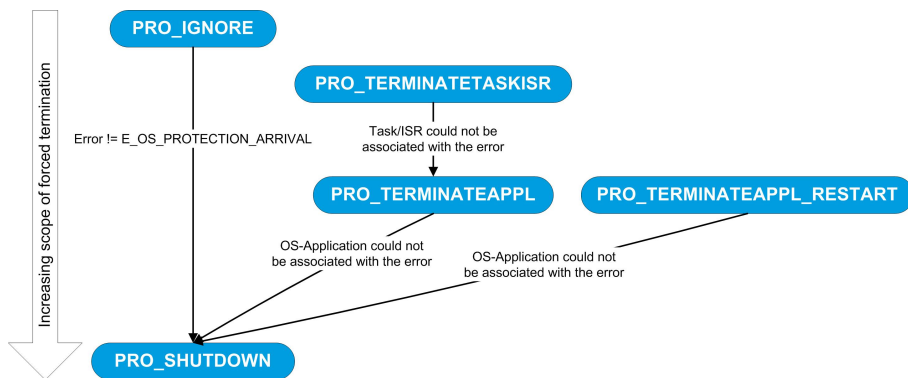


Figure 18.3: Protection Hook fallback actions

```

        break;
    case E_OS_PROTECTION_TIME:
        Action = PRO_TERMINATEAPPL;
        break;
    case E_OS_PROTECTION_ARRIVAL:
        Action = PRO_IGNORE;
        break;
    case E_OS_PROTECTION_LOCKED:
        Action = PRO_TERMINATEAPPL;
        break;
    case E_OS_PROTECTION_EXCEPTION:
        Action = PRO_SHUTDOWN;
        break;
    }
}
return Action;
}

```

Code Example 18.1: Suggested Structure of the Protection Hook

18.2.1 Fallbacks

When an object cannot be associated with an error, or when the wrong action is returned in response to an error, AUTOSAR OS R3.x will *fall back* to an alternative action. Figure 18.3 shows the possible fallback cases.

18.3 Forced Termination

18.3.1 Tasks and ISRs

When a Task or ISR is forcibly terminated, RTA-OS3.x will automatically release all locked resources and remove all interrupt locks held by that Task or ISR.

In addition, when a currently-running ISR is forcibly terminated RTA-OS3.x calls `Os_Cbk_Terminated_<ISRName>` which must then clear the pending flag for the terminated ISR. Failing to clear the interrupt-pending flag will cause the interrupt to be re-entered when the processor priority is lowered.

An implementation of `Os_Cbk_Terminated_<ISRName>` must be provided for each ISR in the following circumstances:

- Timing protection is configured;
- Memory protection is configured;
- a system uses `TerminateApplication()`.


18.3.2 OS-Applications

The following actions are done when an OS-Application is terminated:

- each task and ISR owned by the OS-Application is forcibly terminated (see Section [18.3.1](#)).
- all running alarms owned by the OS-Application are canceled.
- all running schedule tables owned by the OS-Application are stopped (note that this may result in a next-ed schedule table owned by the OS-Application being left in the `SCHEDULETABLE_NEXT` state).
- all interrupt sources of ISRs owned by the OS-Application are disabled (see Section [18.3.3](#)).

18.3.3 Disabling Interrupt Sources

Disabling interrupt sources has the same issue as clearing the interrupt pending flag - RTA-OS3.x knows nothing about the peripheral itself because it is correctly decoupled from the hardware to provide minimum dependencies. However, the system integrator knows which peripheral (or peripherals) is handled by a particular ISR, so RTA-OS3.x expects to find a callback that can be used to disable the interrupt source.

 *Interrupt source control is not standardized in AUTOSAR OS R3.x so this model is not necessarily portable to other implementations.*

The callback has the form `Os_Cbk_Disable_<ISRName>`. An implementation of this callback must be provided for each ISR in a system where `TerminateApplication()` is used. Code Example [18.2](#) shows how such a callback might be programmed.

```

FUNC(void, OS_APPL_CODE) Os_Cbk_Disable_MyISR(void){
    DEVICE_CTRL |= SET_MASK_BIT;
}

```

Code Example 18.2: Disabling an ISR source



Interrupt disable callbacks need to be trusted code. RTA-OS3.x executes the callbacks at OS level and with the same access rights as the OS itself.

18.4 Generating a Skeleton ProtectionHook()

RTA-OS3.x can generate the framework for the ProtectionHook() that includes checking for all types of protection error.

The framework protectionHook() is generated using the following **rtaosgen** command line:

```
C:>rtaosgen --samples:[ProtectionHook] MyConfig.xml
```

This generates a protection hook in Samples\Hooks\ProtectionHook.c that you can use in your application. If the file is already present, then **rtaosgen** will generate a warning. If you want to overwrite an existing file, then you can use:

```
C:>rtaosgen --samples:[ProtectionHook]overwrite MyConfig.xml
```

18.5 Summary

- When applications share the same processor, faults in one application can propagate to another application, causing it to fail.
- AUTOSAR OS R3.x provides three types of protection mechanisms to prevent fault propagation.
 - Timing Protection
 - Service Protection
 - Memory Protection
- Attempted violations of the protection boundaries result are detected by RTA-OS3.x
- The reaction to protection faults is user-programmable using the ProtectionHook()

19 Timing Protection

Timing protection is provided in Scalability Classes 2 and 4. It is independent of whether or not OS-Applications are configured.

In a Scalability Class 2 configuration, timing protection can be applied to tasks and ISRs as required - it is permitted to omit the configuration of protection for objects that work correctly.

In a Scalability Class 4 configuration, all tasks and ISRs in non-trusted OS-Applications must define timing protection attributes. Tasks and ISRs in trusted OS-Applications can optionally define timing attributes.

19.1 What is a timing fault?

A timing fault occurs in a real-time system when a task or ISR misses a deadline. A deadline is a statement of the longest acceptable time between the task or ISR being released into the system and it completing its execution. For tasks, this is typically the time between activation and termination. For ISRs, this is typically the time between the interrupt occurring and the interrupt handler returning.

The time required for a task or ISR to complete its processing is called the *response time*. Typically we are interested in the worst-case response time - the longest time it will ever take the task or ISR to complete. A deadline is satisfied if the response time is less than or equal to the deadline. Figure 19.1 shows how response time is related to the deadline.

In a preemptive real-time system, the response time for a task or ISR comprises three different aspects of time:

Execution Time which is the time the task (or ISR) itself takes to execute, excluding all preemptions.

Interference Time which is the amount of time for which a task (or ISR) is preempted by higher priority tasks (or ISRs) executing in preference. This is determined by how often higher priority tasks execute (their arrival rate) and how long they run for (their execution times).

Blocking Time which is how long a task (or ISR) is prevented from running because a lower priority task (or ISR) has locked a shared resource or disabled interrupts

You can calculate response times, and check whether deadlines can be met, using an analytical technique called *schedulability analysis*. A full discussion of schedulability analysis is outside the scope of this guide, but you can learn more about this topic in the *Analysis Visualizer User Guide* which explains

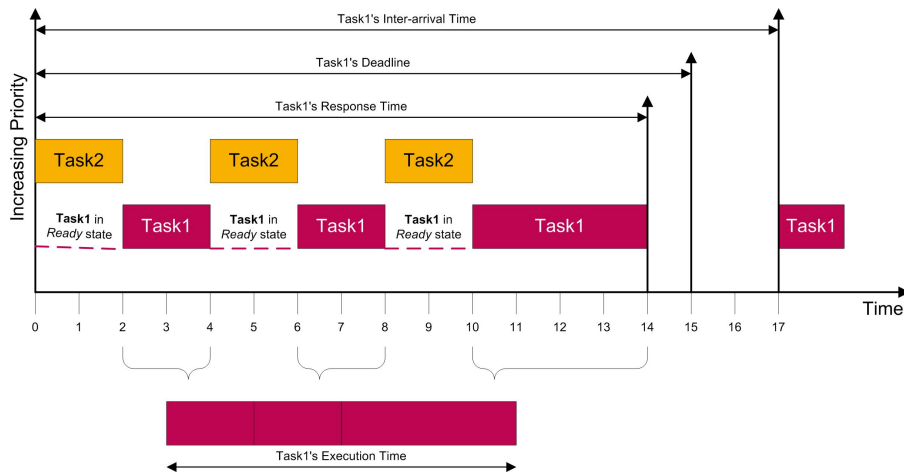


Figure 19.1: Relationship between the response time and the deadline

how to model systems for analysis and use the schedulability analysis tools supplied with RTA-OS3.x.

19.2 Meeting Deadlines at Runtime

19.2.1 Why deadline monitoring is not sufficient

One solution to the problem of missed deadlines is *deadline monitoring*, where each task and ISR is associated with a deadline and the OS checks whether the deadline is exceeded at runtime. In a preemptive OS like AUTOSAR OS R3.x this solution can tell you that a task or ISR misses a deadline, but cannot tell you the cause of the overrun.

Deadline monitoring cannot identify which task is at fault because the response time includes the interference and blocking times - both of which are dependent on other tasks in the system. If any of these other tasks execute for longer than expected, then the task being monitored will have a longer response time. In consequence, a task may miss its deadline even though it is executing correctly. Note that it is impossible to tell which higher priority task (or tasks) is at fault - deadline monitoring simply cannot tell you.

Consider a system with three tasks with the following characteristics:

Task	Priority	Execution Time	Period/Deadline
A	High	1ms	5ms
B	Medium	3ms	10ms
C	Low	5ms	15ms

Assume that all tasks are released at the same time zero and that the deadlines are equal to the period of the tasks. This system models an AUTOSAR OS system with basic, non-queued tasks (BCC1) driven by alarms with the

specified period and which were all started at the same time (e.g. using `SetAbsAlarm(<Alarm>, <SameTime>, <Period>)`).

Figure 19.2(a) shows the intended behavior: all tasks execute correctly, and all meet their deadlines.

Figure 19.2(b) shows the same set of tasks, but this time Tasks A and B sometimes execute for too long and the second arrival of Task B occurs 2 ticks before specified. In this case, Tasks A and B both meet their respective deadlines, *even though they demonstrate incorrect behavior*, but Task C fails to meet its deadline *even though it is the only task in the system that executes correctly*.

Deadline monitoring cannot help to prevent timing problems in a preemptive OS because it allows timing faults to propagate around the system until they appear as failures (missed deadlines). In AUTOSAR OS R3.x, where timing violations are handled by terminating the offending object, this is a significant safety risk.

19.2.2 Meeting Deadlines and Identifying Violators

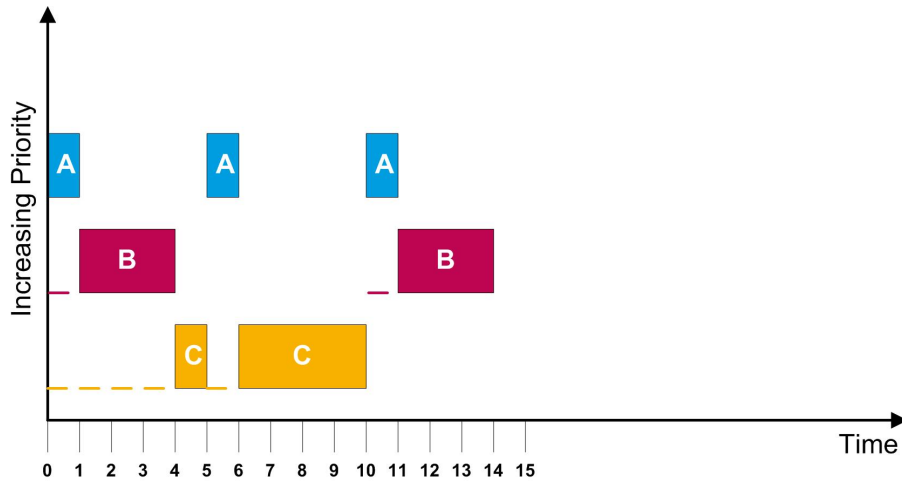
An alternative approach to deadline monitoring is to guarantee upper (or lower) bounds on the factors determining response times:

1. the worst case (longest) execution time of tasks and ISRs
2. the worst case (longest) time for which each standard resource is held
3. the worst case (longest) time for which interrupts are disabled
4. the worst case (shortest) time between task activations (or resumptions from waiting)
5. the worst case (shortest) time between occurrences of an ISR

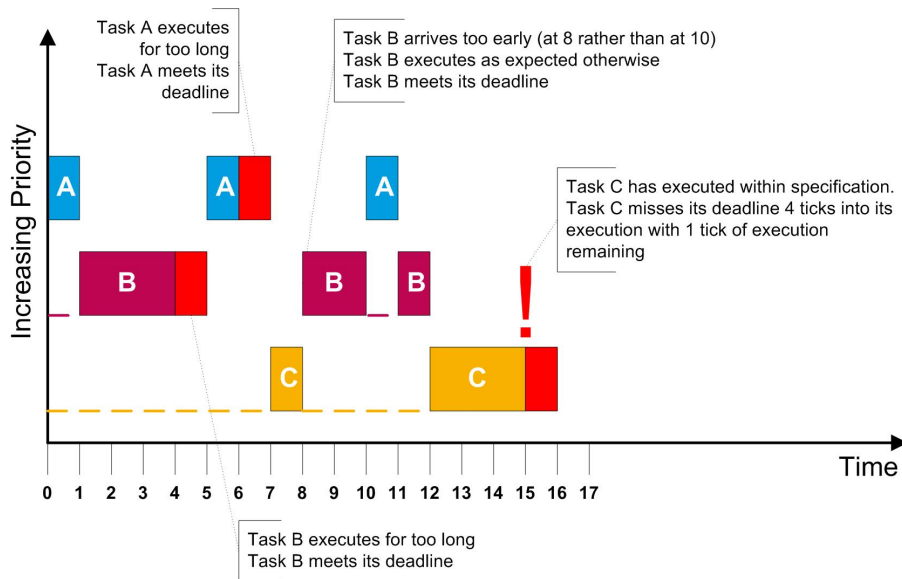
This is the scheme used by AUTOSAR OS R3.x to provide timing protection. Even though there are 5 aspects to control, there are two types of protection mechanism that need to be used:

execution budget protection to prevent tasks and ISRs from executing for longer than allowed, locking resources for longer than allowed or disabling interrupts for longer than allowed.

time frame protection to prevent tasks and ISRs from occurring more frequently than allowed.



(a) Tasks A, B and C run correctly



(b) Tasks A and B at fault, but failure occurs in Task C

Figure 19.2: Impact of faulty timing behavior on lower priority objects

This scheme provides protection against missed deadlines by actively checking that each time a task or ISR executes it does so within the permitted bounds of the timing model. Furthermore, the scheme will immediately identify the faulty object - timing faults are therefore stopped at the point they occur instead of propagating through the system.

19.3 Execution Budgets and Time Frames

Each task and ISR can optionally define timing protection parameters. By default, these parameters are undefined and no timing protection is applied by RTA-OS3.x.

Basic timing protection requires that you define:

1. the execution budget
2. the time frame



All AUTOSAR OS R3.x timing protection parameters are measured in seconds.

The timing protection model of AUTOSAR OS R3.x is intrinsically linked with the state model of the OS. Figure 19.3 shows how these two models interact.

Timing protection for ISRS has a similar three-state model to basic tasks.

19.3.1 Execution Budgets

The execution budget captures the worst-case execution time of a task or ISR for a single invocation. This is the longest time for which the task or ISR executes. It includes the time spent in hook routines and the time spent making OS API calls.

Execution time is measured as follows:

Basic Tasks The net time a task spends in the RUNNING state without entering the SUSPENDED state.

Extended Tasks The net time a task spends in the RUNNING state without entering the SUSPENDED or WAITING state. An extended task executing the `WaitEvent()` API call to wait on an event which is already set is said to have entered the WAITING state.

Note that the execution budget for an extended task must only specify the longest time between starting and the first call to `WaitEvent()`, successive calls to `WaitEvent()` or between `WaitEvent()` and `TerminateTask()`. For example, in Code Example 19.1 you would set

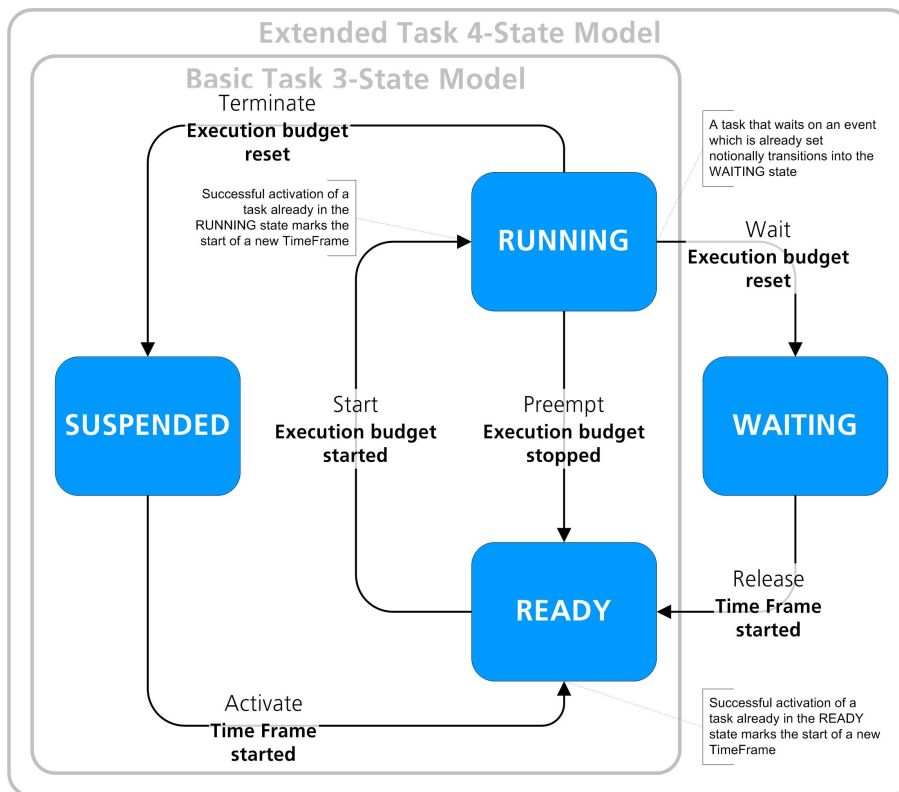


Figure 19.3: Execution time protection and OS state transitions

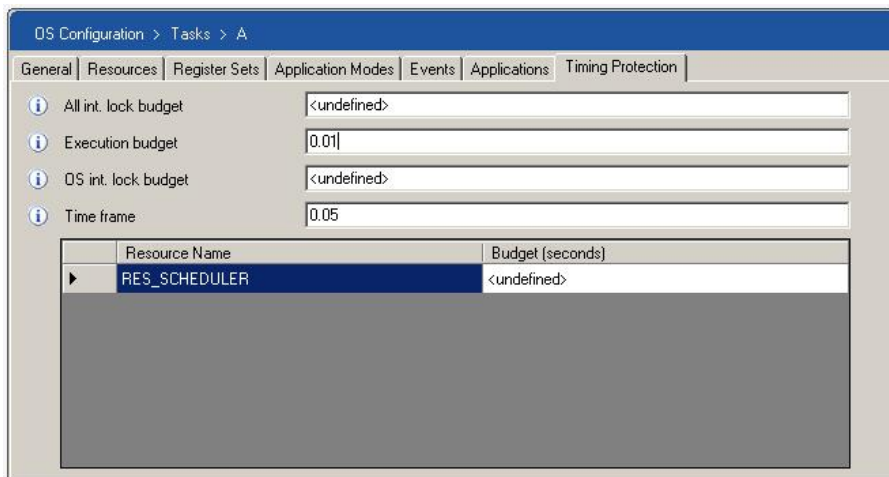


Figure 19.4: Configuration of the execution budget and time frame

the execution budget for the task to 200ms as this is the maximum time that the task will execute without entering the SUSPENDED or WAITING state.

```
#include <Os.h>
TASK(ExtendedTask){
    /* 10ms */
    WaitEvent(Event1);
    /* 200ms */
    WaitEvent(Event2);
    /* 50ms */
    WaitEvent(Event3);
    /* 120ms */
    TerminateTask();
}
```

Code Example 19.1: Execution budget for extended tasks

ISRs The net time from the first to the last instruction of the user-provided Category 2 interrupt handler excluding all preemptions due to higher priority ISRs executing in preference.

Figure 19.4 shows the configuration of a 10ms execution budget.

Each time a task or ISR executes, RTA-OS3.x will check its execution time against the configured execution budget. If the execution budget is exceeded then a protection error occurs and RTA-OS3.x will call the ProtectionHook() with E_OS_PROTECTION_TIME.



Detection of an overrun will occur at the next tick of the timing protection time reference. It is not possible to detect violations at a resolution greater than that provided by the time reference.

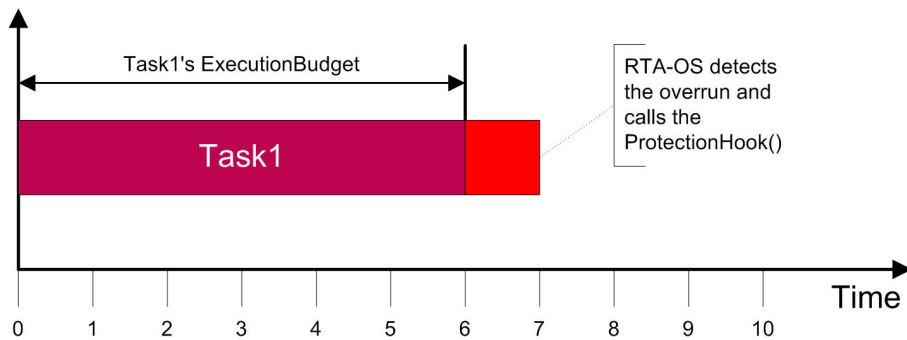


Figure 19.5: ProtectionHook() invocation

19.3.2 Execution Budget Detection Modes

RTA-OS3.x provides two ways to detect execution budget timing faults:

1. Active Detection is the default model in AUTOSAR OS R3.x. A budget overrun attempt is detected as soon as possible and reported through the ProtectionHook(). This is shown in Figure 19.3.2.
2. Passive Detection is a secondary model provided by RTA-OS3.x. A budget overrun is detected on the next appropriate context switch. For a task which is preempted, ProtectionHook() may be called upon resumption of the task; for a task which terminates, the ProtectionHook() may be called at task termination.



The Passive Detection model is not as rigorous as Active detection, and will still allow timing errors to propagate - it is provided as a lower-overhead optimization.



Passive detection is not part of the AUTOSAR OS R3.x standard.

Passive detection uses the same RTA-OS3.x functionality as time-monitoring, but ProtectionHook() is called instead of Os_Cbk_TimeOverrunHook(). Although passive mode detects errors later, it consumes fewer system resources because it can be provided using the same time reference (stop-watch) that is used for time frame protection.

19.3.3 Time Frames

The time frame captures the worst-case inter-arrival time. An inter-arrival time is used to indicate the minimum time between task executions and is measured as follows:

Basic Tasks The time between successively entering the READY state from the SUSPENDED state. Activation of a task always represents a new arrival. This applies in the case of multiple activations, even if an existing instance of the task is in the RUNNING or READY state.

Extended Tasks The time between successively entering the READY state from the SUSPENDED or WAITING states. Setting an event for a task in the WAITING state represents a new arrival if the task is waiting on the event. Waiting for an event in the RUNNING state which is already set represents a new arrival.

ISRs The time between successive occurrences of an interrupt.

Figure 19.4 shows the configuration of a 50ms time frame.

Note that time frame protection still allows you to have queued tasks activations. Task queuing captures how many pending activations are allowed independently of when the activation occur. Time frame protection controls how often things can be added to the queue.

Each time a task or ISR arrives, RTA-OS3.x logs the arrival time. This marks the start of a new time frame. If the time since the last arrival is shorter than the configured time frame then a protection error occurs and RTA-OS3.x will call the ProtectionHook() with E_OS_PROTECTION_ARRIVAL.



The execution budget should always be less than the time frame.

If the execution budget is longer than the timeframe then there will never be enough time to execute the task or ISR to completion. The ratio of the execution budget to the time frame gives the worst case CPU utilization, so you can do the following simple checks to identify whether your configuration is incorrect:

- Is the utilization of each task and ISR less than 100%?
- Is the sum of utilizations for all tasks and ISRs less than 100%?

If the answer to either of these question is “Yes” then you have defined a system which will take more time than there is available and you should re-check your configuration¹.

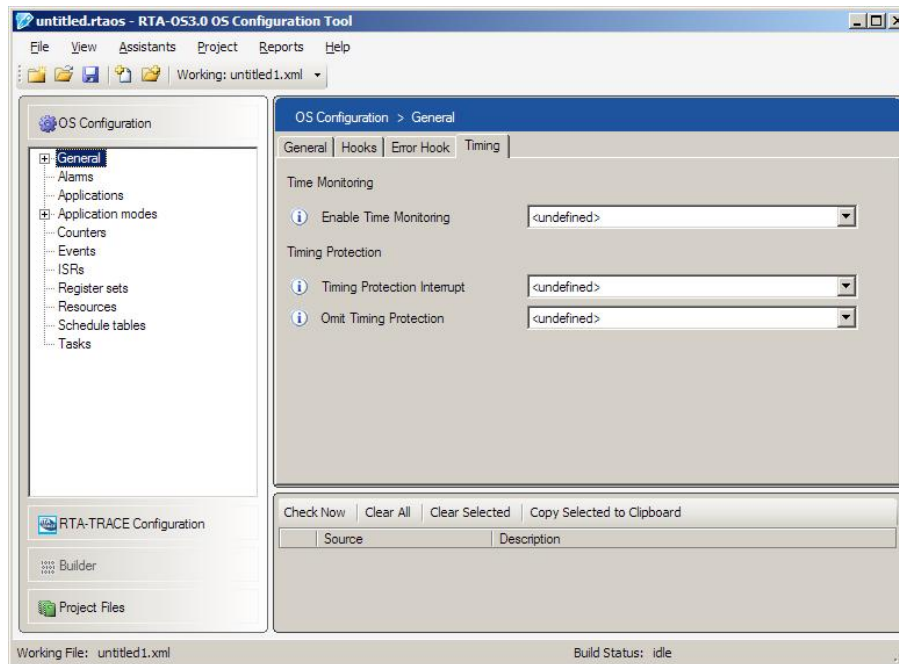


Figure 19.6: Global configuration of timing protection

19.4 Configuring Timing Protection

Figure 19.4 shows the global timing-protection configuration page. The controls function as follows:

Timing Protection Interrupt Declare a *Category 1* ISR to be used to support timing protection. If an ISR is declared then *Active timing protection* is enabled, otherwise *Passive timing protection* is used (See Section 19.3.2). The ISR declared here must call `Os_TimingFaultDetected`. You are also responsible for implementing the `Os_Cbk_SetTimeLimit` and `Os_Cbk_SuspendTimeLimit` callbacks as appropriate for the interrupt source.

Omit Timing Protection Under normal circumstances, timing protection is enabled when any Task or ISR is configured with timing protection values (see Section 19.5). This option allows timing protection to be globally disabled without removing budgets individually from each Task and ISR.

¹A utilization less than 100% is not a guarantee that the system is schedulable, but a utilization more that 100% is definitely not schedulable.

19.5 Lock Budgets

Any task or ISR that disables interrupts or locks resources needs to declare the worst-case (longest) time for the lock. If a lock is made multiple times during execution, then only the longest lock needs to be specified. `rtaoscfg` will show the resources that each task or ISR locks and you need only provide budget.

The list will also include `RES_SCHEDULER` if it is available (see Section 6.7 for more details). If the task or ISR does not lock `RES_SCHEDULER` then a budget of zero should be set.

Figure 19.7 shows the lock times of the task shown in Code Example 19.2.

```
#include <Os.h>
TASK(ExtendedTask){

    if (State == Initial) {
        DisableOSInterrupts();
        /* 0.5ms */
        EnableOSInterrupts();
        ...
        SuspendOSInterrupts();
        /* 1ms - Longest OS lock */
        SuspendOSInterrupts();

        GetResource(Resource1);
        /* 2ms - Longest Resource1 lock */
        ReleaseResource(Resource1);

    } else
        GetResource(Resource2);
        /* 1ms */
        GetResource(Resource1);
        /* 0.5ms */
        ReleaseResource(Resource1);
        /* 1ms */
        ReleaseResource(Resource2);
        ...
        GetResource(Resource2);
        /* 3ms - Longest Resource2 lock */
        ReleaseResource(Resource2);
        ...
        DisableAllInterrupts();
        /* 1ms - Longest ALL lock */
        EnableAllInterrupts();
    }
    TerminateTask();
}
```

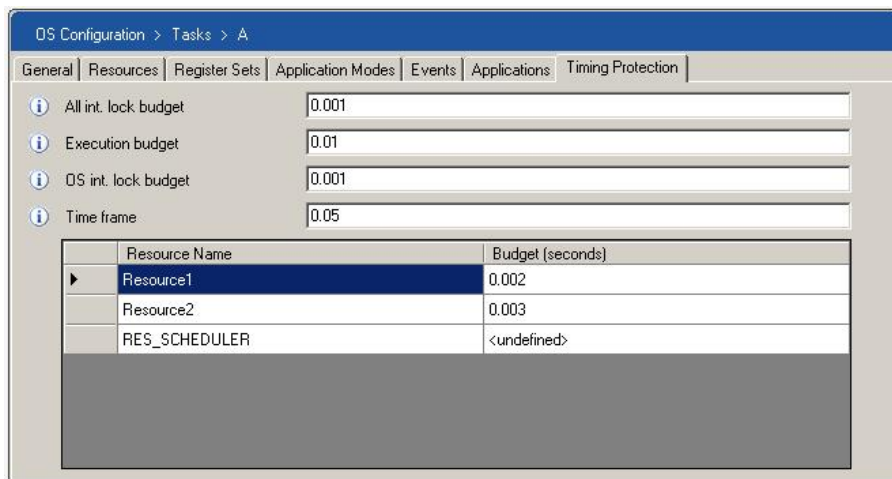


Figure 19.7: Configuration of the lock budgets

}

Code Example 19.2: Lock budgets

Each time a task or ISR locks an interrupt or resource, RTA-OS3.x starts to check the time it is held against the configured budget. If the lock is held for longer than configured, then RTA-OS3.x will call the `ProtectionHook()` with `E_OS_PROTECTION_LOCKED`.

19.6 Providing a Time Reference



Provision of a time reference for timing protection is not defined by the AUTOSAR OS R3.x standard. RTA-OS3.x uses the same philosophy here as elsewhere in the OS by providing a simple and convenient way for you to provide the time reference. However, this is not portable to other implementations.

RTA-OS3.x needs to be provided with a time reference that it can use for measuring times for timing protection.

The time reference required depends on whether you use the active or passive detection mode for execution budget overruns described in Section 19.3.2.

19.6.1 Passive Detection

When using passive detection, RTA-OS3.x uses the “stopwatch” as the free-running timer for monitoring time frames. You will need to provide an implementation of the `Os_Cbk_GetStopwatch()` which provide access to your timer of choice. The free running timer used for the stopwatch must have

sufficient resolution to monitor the shortest execution times and sufficient range to monitor the longest time frame.

Section 15.1.1 provides additional information on the “stopwatch”.

19.6.2 Active Detection

For active detection of budget overruns you need to provide two callbacks:

Os_Cbk_SetTimeLimit This callback is responsible for enabling the timing interrupt and setting it to fire after the specified number of ticks.

Os_Cbk_SuspendTimeLimit This callback must cancel the timing interrupt, returning the number of ticks remaining in the current time limit.

You must also provide the timing interrupt itself, which calls `Os_TimingFaultDetected` when it fires.

`Os_Cbk_SetTimeLimit`, `Os_Cbk_SuspendTimeLimit` and `Os_TimingFaultDetected` are described in greater detail (with examples) in the *Reference Guide*.

19.6.3 Rounding Errors

Whichever form of time reference you provide, RTA-OS3.x will need to convert the configured timing protection parameters, specified in seconds, into ticks of the time reference source.

AUTOSAR OS R3.x requires timing protection parameters to be rounded down to the nearest tick of the associated counter (or stopwatch).

This requirement is safe for time frames because if a system works with a shortened time frame then it will also be work with a longer time frame. In this case, timing protection will work on a tighter bound than the system will actually exhibit (it will be expecting more interference that will actually occur).

However, rounding an execution budget down causes it to become shorter than the task or ISR may actually require. This can result in ‘false positives’ during development as objects will cause timing protection faults even though they are actually behaving correctly.

To avoid any problems that occur due to this, execution times should be rounded up at configuration time to the nearest tick and then one additional tick added.

19.7 Function-Level Timing Protection

RTA-OS3.x extends the AUTOSAR OS R3.x specification to allow time-limited function calls. The option is enabled from the **General → Optimizations** page of the *OS Configuration* Workspace.

The API `CallTimeLimitedFunction()` is very similar to the standard AUTOSAR OS R3.x API `CallTrustedFunction()`, but takes a third parameter which specifies the time limit to be applied to the function call.

```
StatusType  
CallTimeLimitedFunction(  
    TrustedFunctionIndexType FunctionIndex,  
    TrustedFunctionParameterRefType FunctionParams,  
    Os_TimeLimitType TimeLimit)
```

Code Example 19.3: `CallTimeLimitedFunction` API

19.8 Summary

- Timing protection is used to prevent deadlines being missed at runtime.
- AUTOSAR OS R3.x uses a mode of enforcement, where tasks and ISRs cannot:
 - execute for longer than specified.
 - hold standard resources for longer than specified.
 - disable interrupts for longer than specified.
 - run more often than specified.
- Execution budgets, lock budgets and time frames need to be specified if they are to be protected.
- RTA-OS3.x must be given a time reference that it can use to enforce the protection scheme.

20 **Service Protection**

In a protected system it is important that when a user application interacts with the OS through API call this does not place the OS itself into an unknown or unsafe state.

In Scalability Classes 3 and 4, RTA-OS3.x performs additional runtime checks to prevent this from happening. There are three types of error which may occur:

1. calling an API called with an invalid parameter or an out of range value
e.g. calling `ActivateTask(ResourceName)` or `SetRelAlarm(Alarm,MAXALLOWEDVALUE+1),0`. This type of error is checked when RTA-OS3.x is built using extended status. Therefore, Scalability Classes 3 and 4 require that extended status is configured.
2. calling an API in the wrong context
e.g. calling `ActivateTask()` in the `StartupHook()` or making an API call from a Category 1 ISR. This type of error is prevented by additional error checking that is done only when Scalability Classes 3 or 4 is selected.
3. failing to call an API that is required for correct operation
e.g. a task not calling `TerminateTask()`. This type of error is identified by RTA-OS3.x at runtime and an appropriate default action is done, but only when Scalability Classes 3 or 4 is selected.

In addition to these errors, when there are multiple applications integrated onto the same ECU there it is also important that one application does not place another application into an unknown state. There are two additional types of error that can occur:

4. one application making an API call that impacts the behavior of other applications in the system
e.g. `ShutdownOS()`.
5. one application manipulating OS objects that belong to another application
e.g. setting or canceling an alarm in an unrelated application.

RTA-OS3.x reports the occurrence of service protection errors through the return value of API calls in the same way that other errors are returned. An appropriate corrective action can be programmed in the application itself by

checking the return value or as part of system-wide error handling using the `ErrorHook()` (see Section 13.1).



If you can ensure that your system is free from these type of errors (for example, by offline analysis) then RTA-OS3.x allows you to use combine Scalability Classes 3 and 4 with standard status to reduce the memory footprint of the OS and increase its runtime performance.

20.1 OS-Applications

Any protection scheme needs to define the scope of protection. This means that the protection boundary must be defined.

In the AUTOSAR OS systems we have seen so far, there has been no way to identify which Tasks, ISRs, Events, Resources, Alarms, Schedule Tables and Counters etc. belong to which applications, or even which applications have been integrated.

AUTOSAR OS R3.x provides a higher-level abstraction that allows OS objects to be grouped together into a cohesive functional unit called an *OS-Application*.



OS-Applications are not related in any way to Application Modes (see Section 12.2.2). OS-Applications create partitions between OS objects. Application Modes control the mode of the complete OS and which tasks, alarms and schedule tables are auto-started during `StartOS()`.

An OS-Application is said to *own* the OS objects in the group. An OS-Application is used to define the privilege level and the protection boundary of the tasks and ISRs it owns. There are two types of OS-Application:

Trusted OS-Applications run in privileged (supervisor) mode, when supported by the processor. They have unrestricted access to memory, all configured OS objects and the complete OS API. Timing protection is not required but can be optionally configured. All trusted OS-Applications have a different service protection boundary but share the same memory protection boundary - there is no difference in memory protection between declaring a multiple trusted OS-Applications and a single trusted OS-Application.

Non-Trusted OS-Applications run in non-privileged (user) mode, when supported by the processor. Memory protection is mandatory in Scalability Classes 3 and 4. Timing protection is mandatory in Scalability Classes 2 and 4. Tasks and ISRs in an OS-Application only have access to objects owned by the same OS-Application by default.

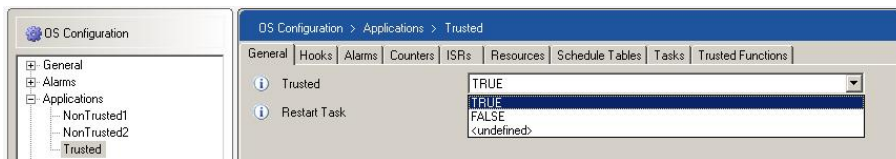


Figure 20.1: Specifying the type of an OS-Application

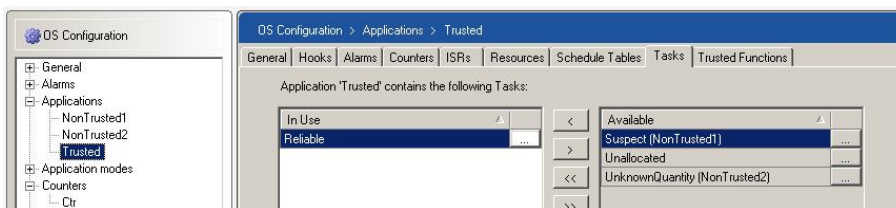


Figure 20.2: Allocating tasks to an OS-Application

The OS itself is trusted.

OS-Applications are mandatory when using service protection. OS-Applications are also used to define the scope of memory protection described in Chapter 21. This means that OS-Applications must always be configured then when building Scalability Class 3 or 4 systems.

There is limited control of OS-Applications at runtime - they can only terminate themselves (and optionally restart).

20.2 Configuring OS-Applications

OS-Applications are defined at configuration time and are allocated a unique name. An OS-Application must specify whether it is trusted or not. An OS configuration can define as many OS-Applications as it requires.

Figure 20.1 shows how an OS-Application is configured as “Trusted” or “Non-Trusted”.

20.2.1 Ownership of OS objects

When OS-Application are defined, each object in the system must be allocated to exactly one OS-Application. Figure 20.2 shows how tasks are assigned by way of example. RTA-OS3.x shows when an object has already been allocated to an OS-Application by identifying the current owner in brackets after the name of the object. Objects that have not yet been allocated are unmarked.

An object is moved from one OS-Application to another by moving it from the “Available” column to the “In Use” column.

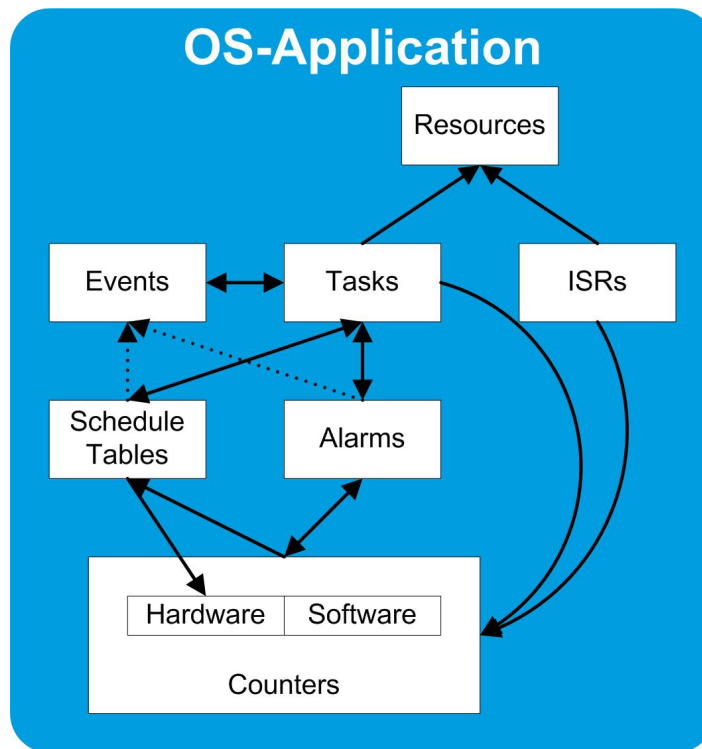


Figure 20.3: Access relationships for OS objects

Objects can be freely allocated to any OS-Application.



If you have ignored the warning in Section 18 and are using Category 1 ISRs then they must belong to a trusted OS-Application.

20.2.2 Access to OS objects

You will typically assign interacting objects to the same OS-Application(i.e. those that form a cohesive application)¹. In particular note that tasks released by alarm actions or from schedule table expiry points should be assigned to the same OS-Application as the alarm or schedule table otherwise they will not be accessible. Similarly, counters that drive alarms or schedule tables should be allocated to the same OS-Application.

Figure 20.3 shows the possible ownership relationships between OS objects. A solid arrow indicates that object X may need to access object Y. A dotted arrow indicates indirect ownership and applies to events only. Any event is accessible if the task using the event is accessible.

¹It is possible to share objects between OS-Applications. This described in Section 20.3.






Check Now		Clear All	Clear Selected	Copy Selected to Clipboard
Source	Description			
 Check	Starting XML Checks			
 Check	Starting Pre-Build Checks.			
 Check	Alarm aY cannot access Task Suspect			
 Check	Counter Ctr cannot access Alarm aY			
 Check	Task Unallocated must belong to an OS-Application			

Figure 20.4: Incorrectly configured ownership causes configuration errors

The “Check Now” feature of **rtaoscfg** will perform a consistency check on your configuration and report any conflicts in the error pane as shown in Figure 20.4.



It is not possible to identify all inconsistencies at configuration time. For example, if task A calls `ActivateTask(B)` then task A will need access to task B at runtime, but this cannot be identified from the OS configuration.

20.3 Accessing Applications

All objects owned by the same OS-Application have access to each other. Objects in the same OS-Application can use API calls to manipulate other objects without needing any special declaration or configuration.

By default, RTA-OS3.x prevents access to objects across the protection boundary, so each OS-Application, whether trusted or non-trusted, remains isolated from other OS-Applications as shown in Figure 20.5.

It is possible, however, for an object in one OS-Application to grant access to other OS-Applications at configuration time. An OS-Application which is granted access to an object is called an *accessing application*.



Access rights are granted for an OS-Application. This means that every object in the accessing OS-Application can manipulate the object which grants access. It is not possible in AUTOSAR OS R3.x to grant access on a per-object basis.

Events owned by tasks are a special case - if a task is accessible then so are all the events that can be set for the task. Events do not need to be configured as accessible individually.

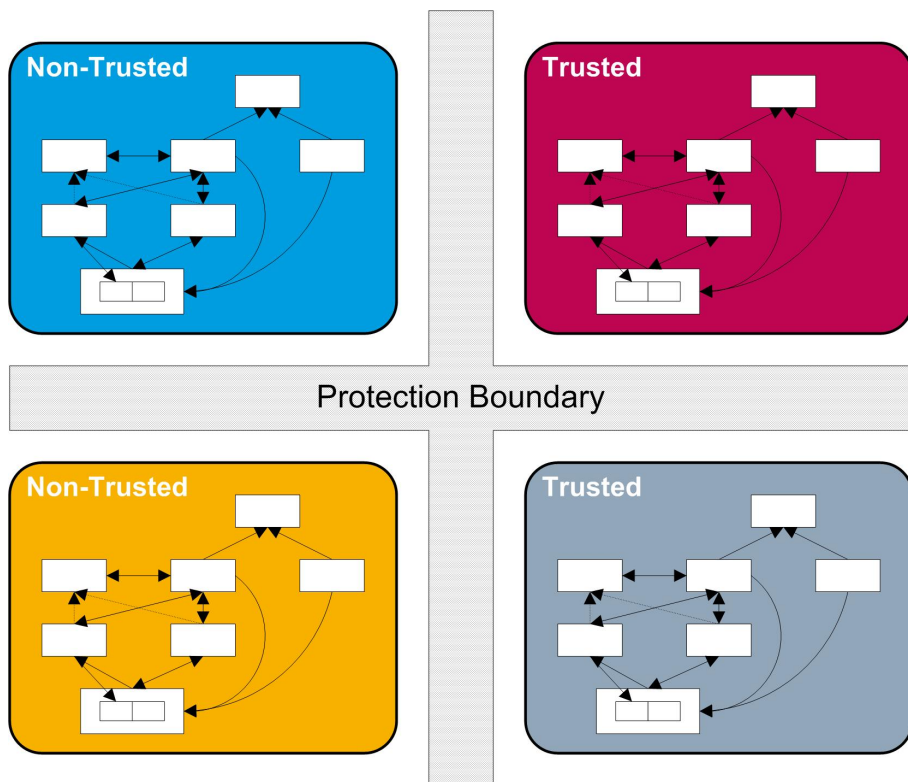


Figure 20.5: Isolated OS-Applications

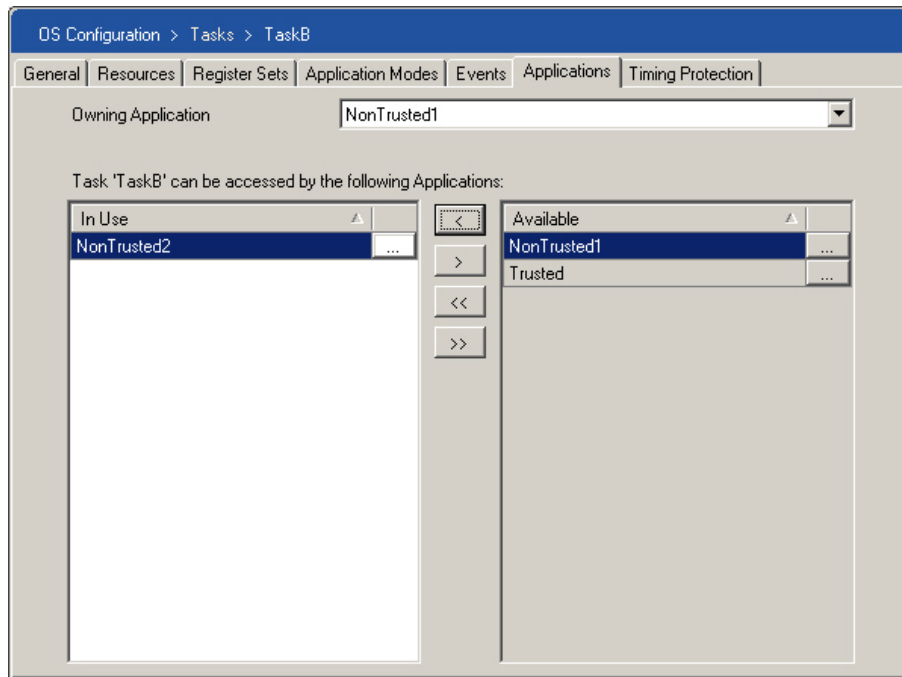


Figure 20.6: Configuring the accessing applications for a task

Any object in an accessing application can use the object granting access, either at configuration time or at runtime. For example, if a task TaskA grants access to OS-Application OtherApp then:

- an alarm in OtherApp can activate TaskA in the alarm action
- any task or ISR in OtherApp can call `ActivateTask(A)`

Figure 20.2 shows how the accessing applications are configured for a task. Access can be granted by moving a task from “Available” to “In Use”.



Granting access to other OS-Applications creates holes in the protection boundary. It is your responsibility to verify whether granting of access to an OS-Application is safe.

20.3.1 Hidden Accesses

All objects within an OS-Application are accessible to each other. When an object grants access to another OS-Application, and that object manipulates other objects in its owning OS-Application, there is a route by which the accessing application can indirectly manipulate those other objects.

For example, if an alarm which activates a task is made accessible, then any object in the alarm’s accessing OS-Application can use the alarm to indirectly

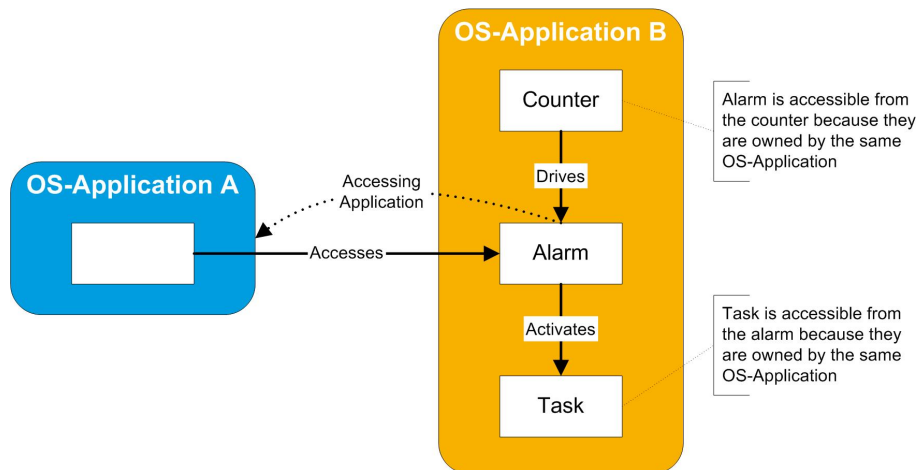


Figure 20.7: Objects in the same OS-Application do not need to declare transitive access

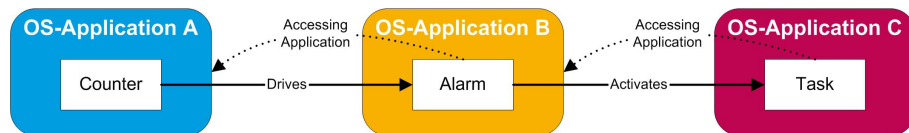


Figure 20.8: Accessing an alarm from a counter and a task from an alarm

control whether the task is activated (or not). This relationship is shown in Figure 20.7.

If the manipulated object is in another OS-Application then access needs to be defined. For example, consider the case where a counter is owned by one OS-Application, an alarm driven by the counter is owned by a second OS-Application and the task activated by the alarm is owned by a third OS-Application as shown in Figure 20.8.

In this case, the following accessing applications need to be defined:

- the alarm needs to be accessible to the counter by which it is driven.
- the task needs to be accessible by the alarm (so that the alarm can perform the action when it expires)

The same model is true for schedule tables, where every task in a different OS-Application to the schedule table will need to define an accessing application as illustrated by Figure 20.9.

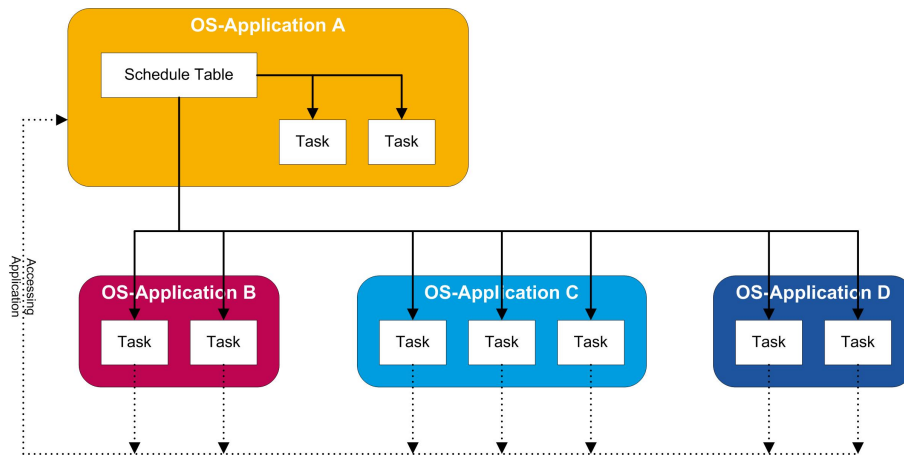


Figure 20.9: Accessing tasks from a schedule table

20.3.2 Safety Risks

Every time an object declares an accessing OS-Application, it is actively breaking the protection boundary provided by the OS-Application concept². This means that faults in the accessing application can propagate across the protection boundary and create secondary (possibly more serious) problems in the application granting access.



The ability to define accessing OS-Applications violates one of the basic principles safety critical systems engineering - do not provide any uncontrolled route through which faults can propagate. You should avoid defining accessing OS-Applications if you plan to build any type of critical system.

20.4 Identifying the Running OS-Application

OS-Applications do not have any state themselves, but are said to 'running' if a task or ISR owned by the OS-Application is in the RUNNING state.

The identity of the running OS-Application is available at runtime using `GetApplicationID()`. If this is called from a task or ISR then it will return the owning application. However the API call is of most use in system-wide hook routines. Code Example 20.1 shows how it might be used in the `PreTaskHook()`.

```

FUNC(void, OS_APPL_CODE) PreTaskHook(void) {
    ApplicationType RunningApplication;
    RunningApplication = GetApplicationID();

    switch (RunningApplication) {

```

²This is analogous to locking your house but providing people with a key. The more keys you distribute, the less safe your house becomes.

```

        case AppA:
            ...
        case AppB:
            ...
        default:
            ...
    }
}

```

Code Example 20.1: Identifying the running OS-Application

20.5 Access and Ownership

The accessibility and ownership of OS objects can be checked at runtime in two ways.

The first way is to check the return code of the API call. If a OS object is not accessible at runtime, the API call will return the `E_OS_ACCESS` status code. This can be checked at runtime and appropriate corrective action taken as shown in Code Example 20.2.

```

TASK(FaultTolerant){
    if (ActivateTask(MaybeAccessible) == E_OS_ACCESS) {
        /* Task could not be activated */
    }
    TerminateTask();
}

```

Code Example 20.2: Checking access using status codes

The second way to check success rights is before an API call is attempted. This can be done using `CheckObjectAccess()` as shown in Code Example 20.3. The call can be used to check access to any type of object, but must be told the type of the object that is passed in.

```

if CheckObjectAccess(GetApplicationID(),OBJECT_TASK,SomeTask) {
    /* SomeTask is accessible */
    ActivateTask(SomeTask);
}

```

Code Example 20.3: Checking access using an API

Similarly, the ownership of an OS object can be checked at runtime using `CheckObjectOwnership()` as shown in Code Example 20.4. The call returns the name of the OS-Application that owns an object.

```

if CheckObjectOwnership(OBJECT_TASK,SomeTask) ==
    GetApplicationID() {
    /* If we own SomeTask, then activate it */
    ActivateTask(SomeTask);
}

```

```
}
```

Code Example 20.4: Checking object ownership

20.6 Terminating OS-Applications

An OS-Application can terminate itself using the `TerminateApplication()` API call. This is similar in concept to `TerminateTask()`, but instead of terminating the current instance of a task, it terminates the current instance of an OS-Application.

The `TerminateApplication()` call takes a parameter which controls what happens after termination:

- `NO_RESTART` terminates the OS-Application
- `RESTART` also terminates the OS-Application but tells RTA-OS3.x to activate a *restart task* to 'tidy-up' the application's state

Code Example 20.5 shows how the call might be used.

```
TASK(ApplicationChecker){
    ...
    switch (ApplicationState) {
        case Completed:
            TerminateApplication(NO_RESTART);
            break;
        case TransientFaultDetected:
            TerminateApplication(RESTART);
            break;
        case PersistentFaultDetected:
            TerminateApplication(NO_RESTART);
            break;
    }
    ...
}
```

Code Example 20.5: Terminating an OS-Application

The current instance of an OS-application comprises the tasks and ISRs that are running or ready (or preempted) and all alarms and schedule tables that are running. When an OS-Application is terminated, RTA-OS3.x takes the following actions:

- each task and ISR owned by the OS-Application is forcibly terminated (see Section 18.3.1).
- all running alarms owned by the OS-Application are canceled.

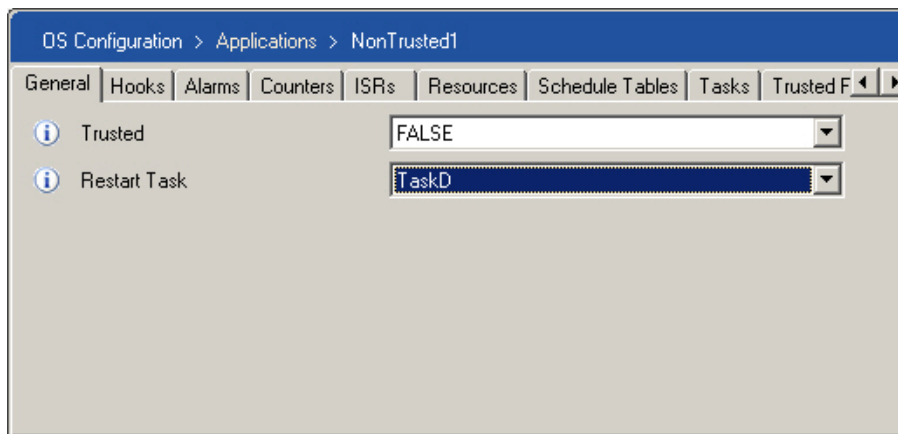


Figure 20.10: Configuring the restart task

- all running schedule tables owned by the OS-Application are stopped (note that this may result in a next-ed schedule table owned by the OS-Application being left in the SCHEDULETABLE_NEXT state).
- all interrupt sources of ISRs owned by the OS-Application are disabled (see Section 18.3.3).

20.6.1 The Restart Task

If an OS-Application does not have any accessing applications, then termination ensures that every OS object owned by the OS-Application will not be used until either the ECU is power-cycled or `StartOS()` is called again. In this case, termination disables the OS-Application.

Alternatively, you may want to terminate an OS-Application in response to some kind of internal error and the restart. However, because `TerminateApplication()` forcibly terminates tasks and ISRs, the state of the OS-application may be inconsistent. For example, the termination might have occurred during a critical section and data might not be in a consistent state.

RTA-OS3.x does not know how to make this state consistent, but it can activate a task owned by the OS-Application that can 'clean up' or re-initialize the internal state.

This task is known as the restart task. It is a normal OS task and will run according to its defined scheduling policy and priority level. Any task which is owned by the OS-Application can be configured as the restart task.

Figure 20.10 shows how the restart task is configured.

The restart task will be activated:

- when `TerminateApplication(RESTART)` is called.
- if the OS-Application has attempted to violate the memory or timing protection boundary and the `ProtectionHook()` returns `PRO_TERMINATEAPPL_RESTART`.

20.6.2 Interaction with Accessing Applications

If an OS-Application has accessing applications, then they can interact with the objects in a terminated OS-Application, for example by starting an alarm, or activating a task, or locking a shared resource to enter a shared critical section.

This is particularly dangerous because the accessing application does not know whether the accessed application objects are operating on corrupt data or not.

Even if the terminated OS-Application has a restart task and it has been activated, there is no guarantee that the restart task will run before an accessing OS-Application with higher priority tasks (or ISRs) interacts with objects in the terminated OS-Application. This can trigger failures in both the accessing application, the terminated but partially restarted OS-Application or the wider system.

It is possible with careful application design to avoid some of these problems, for example by ensuring that restart tasks are dedicated tasks and are allocated the highest priorities in the system (but note that this will not prevent interaction problems with ISRs).

However, these type of problems are best avoided by not introducing them into your OS-Application design - i.e. do not define any accessing applications.

20.7 OS-Application Hooks

An OS-Application is similar to a virtualized OS - it has its own lifecycle that involves startup, shutdown and error handling.

At integration time, OS-Application specific code could be added to the standard `StartupHook()`, `ShutdownHook()` and `ErrorHook()` for each OS-Application. However, each time the integration changes these hooks would need to be re-written.

AUTOSAR OS R3.x provides a more elegant solution to the problem of OS-Application specific hook functionality by allowing each OS-Application to define application-specific hooks. These hooks have the same parameters as

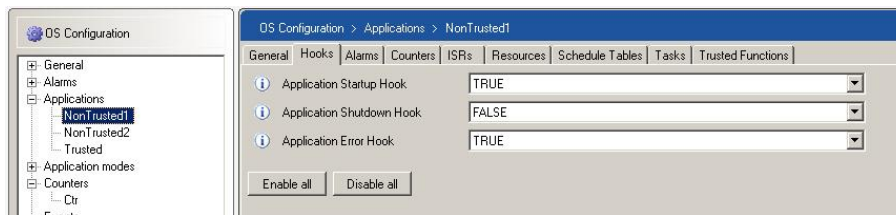


Figure 20.11: Configuring application-specific hooks



Figure 20.12: Execution of Application-specific Startup Hooks

the system-wide hooks but they have a unique name derived from the name of the OS-Application as follows:

<hook-name>_<os-application-name>

Figure 20.11 shows how the hooks can be enabled or disabled.



You must provide an implementation of any application-specific hook you configure. If you do not provide an implementation then your program will not link correctly.



The ordering of calls to OS-Application hooks is non-deterministic. Your system must therefore not rely on calls being executed in any particular order.

20.7.1 Startup Hook

Application-specific startup hooks are called after the system-wide StartupHook() returns and before the scheduler is running. If more than one OS-Application declares its own startup hook, then the ordering in which the hooks are called is non-deterministic.

The hooks run with the access rights of the owning OS-Application - they are trusted if the OS-Application is trusted and non-trusted otherwise.

Figure 20.12 shows the execution of the Startup Hook relative to the initialization of RTA-OS3.x.

Code Example 20.6 shows how an application-specific startup hook should appear in your code.

```
FUNC(void, OS_APPL_CODE) StartupHook_MyApplication(void) {
    /* Startup hook code. */
    EnableIOInterrupts();
    EnableTimerInterrupts();
}
```

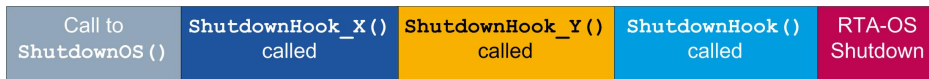



Figure 20.13: Execution of Application-specific Shutdown Hooks

```
...
}
```

Code Example 20.6: Application-specific Startup Hook

20.7.2 Shutdown Hook

Application-specific shutdown hooks are called when `ShutdownOS()` is called. The hooks run before the system-wide `ShutdownHook()` is called. As with the startup hook, if more than one OS-Application declares its own shutdown hook, then the ordering in which the hooks are called is non-deterministic. Similarly, the hooks run with the access rights of the owning OS-Application.

Figure 20.13 shows the execution of the Shutdown Hook relative to the call of `ShutdownOS()`.

Code Example 20.7 shows how an application-specific shutdown hook should appear in your code.

```
FUNC(void, OS_APPL_CODE) ShutdownHook_MyApplication(void) {
    /* Shutdown hook code. */
    DisableIOInterrupts();
    DisableTimerInterrupts();
    WritePersistantDataToNVRAM();
    ...
}
```

Code Example 20.7: Application-specific Startup Hook

20.7.3 Error Hook

Each OS-Application can do localized error handling by defining an application-specific error hook.

An application-specific error hook is called after the system-wide `ErrorHook()` is called. The hook runs with the access rights of the owning OS-Application.

Code Example 20.8 shows an application-specific error hook.

```
FUNC(void, OS_APPL_CODE) ErrorHandler_MyApplication(StatusType
    status) {
    switch (status) {
        case E_OS_ACCESS:
```

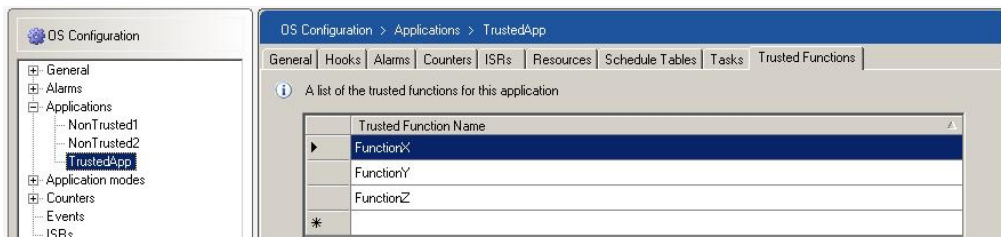


Figure 20.14: Declaring trusted functions

```

    /* Handle error then return. */
    break;
case E_OS_LIMIT:
    /* Terminate the OS-Application and then restart it */
    TerminateApplication(RESTART);
default:
    break;
}
}

```

Code Example 20.8: Application-specific error hook

20.8 Trusted Functions

Trusted OS-Applications can export functions, called *trusted functions*, whose purpose is to encapsulate functionality that needs to run in privileged mode and make it available to non-trusted OS-Applications. For example, a trusted function may be used to access peripheral hardware that can only be used when the processor is in a privileged mode.



Trusted functions are another case where AUTOSAR OS R3.x allows you to break the protection boundary of OS-Applications. You should exercise caution when using them.

RTA-OS3.x needs to be told the names of the trusted functions which will be exported by each trusted OS-Application. Figure 20.14 shows the declaration of three trusted functions; FunctionX(), FunctionY() and FunctionZ().

20.8.1 Calling Trusted Functions

Trusted functions are called indirectly through the CallTrustedFunction() API. This API provides the necessary wrapper functionality to move between non-privileged and privileged modes (e.g. when calling the trusted function from a non-trusted OS-Application). The call has the form:

```
CallTrustedFunction(<FunctionName>, <PointerToParameters>)
```

Note that RTA-OS3.x does not need to know either the number or the type of the parameters because the function itself is responsible for doing any

parameter marshalling required. If the trusted function is parameter-less, then this can be indicated using zero.

Code Example 20.9 shows some examples of how the mechanism can be used:

```
/* Parameter-less function */
CallTrustedFunction(FunctionX,
    (TrustedFunctionParameterRefType)0U);

/* Single parameter */
uint8 one_parameter 42;
CallTrustedFunction(FunctionY, &one_parameter);

/* Multiple parameters passed in a C struct */
struct {
    uint32 parameter1;
    bool   parameter2;
    float32 parameter3;
    sint16 parameter4;
} multiple_parameters = {99,true,0.1,-42};
CallTrustedFunction(FunctionZ, &multiple_parameters);
```

Code Example 20.9: Calling trusted functions

20.8.2 Implementing Trusted Functions

You are responsible for providing the implementation of each trusted function and marshalling the parameters correctly. All trusted functions have the following function prototype:

```
void FunctionName (<parameter type> *<name>)
```

They take a single reference parameter which is a pointer to the type of the parameters and return **void**.

If the function needs to provide some return value then this needs to be done using an 'out' parameter.

Code Example 20.10 shows how trusted function FunctionZ() from Code Example 20.9 might be implemented.

```
void FunctionZ (parameter_struct *local_struct)
{
    if (local_struct.parameter1 < 100) {
        ...
    }
}
```

Code Example 20.10: Implementing a trusted function



The trusted function itself can write to any memory location, so there is the possibility that a non-trusted caller can corrupt the system state using by calling a trusted function.

Some protection against this can be provided in the trusted function itself by validating whether or not the caller is permitted to write to selected locations using the `Check[Task|ISR]MemoryAccess()` API call.

20.9 Summary

- OS-Applications allow application software running on the same OS configuration to be partitioned to prevent unwanted interactions.
- OS-Applications can be trusted and run in privileged (supervisor) mode or can be non-trusted and run in non-privileged (user) mode.
- OS objects that share the same OS-Application have access to each other. OS objects in other OS-Applications are denied access by default.
- Access can be granted to other OS-Applications at configuration time.
- Access and ownership can be checked at runtime so that systems can be defensively programmed.

21 Memory Protection

Memory protection is provided when using Scalability Class 3 or 4. If a memory access violation is detected, RTA-OS3.x calls the `ProtectionHook()` with status code `E_OS_PROTECTION_MEMORY`.

AUTOSAR OS's memory protection model interacts with the OS-Application service protection boundary as follows:

Trusted Os-Applications: Tasks and ISRs in trusted OS-Applications have unrestricted access to all memory locations and can perform any operation on those locations that is permitted by the hardware¹.

Non-trusted Os-Applications: Tasks and ISRs in non-trusted applications have unrestricted access to their own memory locations and any data owned by their owning OS-Application. However, they have (optionally) restricted read, write and execute access to other memory locations. Figure 21.1 shows the mandatory and optional restrictions for non-trusted OS-Applications.

The following table summarizes the memory access restrictions defined by AUTOSAR OS R3.x for Tasks and ISRS in trusted and non-trusted OS-Applications:

¹They can read, write and execute the contents of RAM locations but only read and execute from ROM locations.

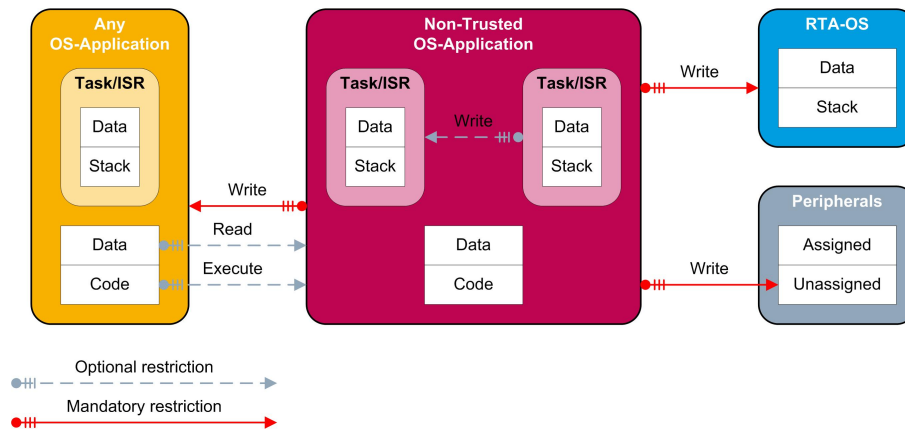


Figure 21.1: Memory protection restrictions for non-trusted OS-Applications

Access Type	Trusted	Non-Trusted
Read	Unrestricted	(Optionally) cannot read from private data belonging to other OS-Applications.
Write	Unrestricted	Cannot write to: <ul style="list-style-type: none"> • RTA-OS3.x's stack; • RTA-OS3.x's private data; • private data belonging to other OS-Applications; • private data belonging to Tasks/ISRs in different OS-Applications; • the stack of Tasks/ISRs in different OS-Applications; • non-assigned peripherals; • (optionally) private data belonging to another Task/ISR in the same OS-Application.
Execute	Unrestricted	(Optionally) cannot execute code from other OS-Applications.

21.1 Hardware Limitations

The capability of the MPU on your hardware will define what protection model is possible and this may mean that the AUTOSAR OS R3.x memory protection model cannot be implemented practically.

MPUs typically will define a fixed (usually small) number of protection regions called *contexts* and will define restrictions on the size of the memory region that can be protected. Similarly, on many devices it may not be possible to grant or deny access to peripherals on a per-peripheral basis.

Each context must be contiguous block of memory. The number of contexts provided by your MPU therefore controls the number of distinct blocks of memory that can be protected. For example, if you have a section of initialized and non-initialized data and locate them in different areas of the memory space, then two contexts will be required to protect them. However, if you locate the sections adjacently then you will be able to use a single context to protect both sections.

Even if there are large numbers of contexts are available, the minimum size restriction, called the *granularity*, of contexts will also need to be considered. On some devices, the granularity may be as little as 1 byte, on others it may be a few kilobytes. In the latter case, if only a few bytes need to be protected, then a block of memory equal to minimum granularity of a context would

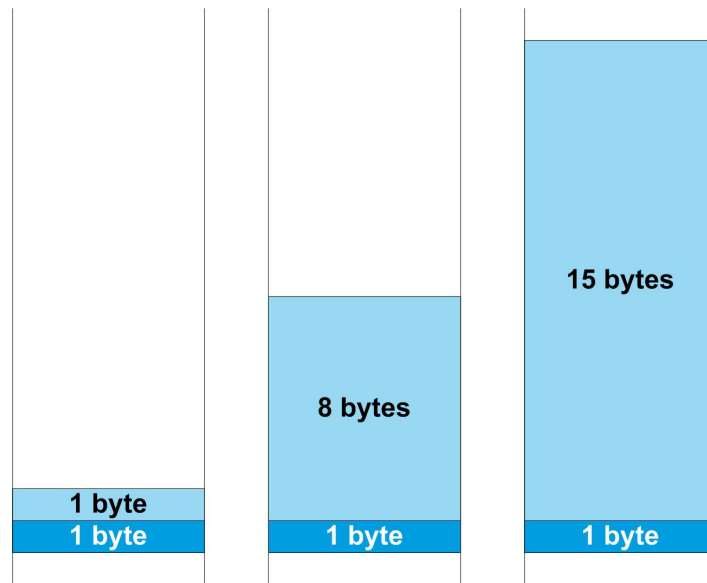


Figure 21.2: Impact of protection region granularity on wasted memory space

need to be allocated. This can be wasteful of memory. Figure 21.2 shows the amount of space that would be wasted when protecting one byte on devices with granularities of 2 bytes, 8 bytes and 16 bytes.

The granularity of your device may require you to aggregate memory sections in order to make the least wasteful use of the memory available.

21.2 AUTOSAR Limitations

The memory protection model defined by the AUTOSAR OS R3.x standard may be too restrictive (or not restrictive enough) for your needs. You may, for example, have multiple third party OS-Applications that you want to share the same protection region, or you may want to assign different levels of protection to different OS-Applications. These use-cases are not supported by AUTOSAR OS R3.x.

To avoid these issues, RTA-OS3.x places the setup of the memory regions and the programming of the MPU in the user domain using a callback mechanism. Each time RTA-OS3.x performs a context switch that requires a switch of the protected regions, it calls the `Os_Cbk_SetMemoryAccess()` callback with the name of the Task or ISR that will be switched into and its owning OS-Application. The callback is responsible for programming the MPU with the required protection settings.

This model provides complete freedom, within the capabilities provided by the hardware, to define whatever protection model is appropriate to your needs. RTA-OS3.x simply applies the chosen scheme at the correct points

in scheduling. Section 21.5 explains how this scheme works and how to get started with sample implementations.

21.3 Defining and Using Memory Regions

You are responsible for defining memory sections and allocating data to them. AUTOSAR already provides a standardized mechanism for this with the memory mapping concept. Memory sections can be defined in application code and can be aggregated and named for linking using MemMap.h. You will then need to use the linker control file to locate the regions accordingly. The same scheme is used by RTA-OS3.x for memory protection.

Recall from Section 3.3.2 that BSW modules can define and allocate memory into the sections using the START_ and END_ macros:

```
#define <Module>_START_SEC_<SectionName>
#include "MemMap.h"
/* Things to place in the section here */
#define <Module>_STOP_SEC_<SectionName>
#include "MemMap.h"
```

The user file MemMap.h defines how these logical sections are aggregated into larger sections (when required) and how those aggregated sections are defined using the compiler's mechanism for placing data into sections:

```
/* Map OS code into the section containing all BSW code */
#ifdef <Module>_START_SEC_<SectionName>
    #undef <Module>_START_SEC_<SectionName>
    #define START_SECTION_<AggregatedSectionName>
#endif
...
/* Name the system section with a compiler primitive */
#ifdef START_SECTION_<AggregatedSectionName>
    #pragma section code "aggregated_section"
#endif
```

The same model can be used by Tasks and ISRs in the application to place their data sections and then aggregate this into regions.

This gives you a huge amount of flexibility in terms of how many sections you define, what they are called and whether you decide to group data together in your Task code or in the MemMap.h file.

21.3.1 Placing Task and OS-Application Data into the same section

In the simple case, assume that all data for all Tasks and their owning OS-Application should be placed into the same memory region. Data needs to be placed into the same section in your application code as shown in Code Example 21.1.


```

#include <Os.h>

/* Place all data in this file into the same section */
#define APP1_START_SEC_SHARED_DATA
#include "MemMap.h"
uint32 sVar;
uint32 aVar;
uint32 bVar;

Task(A){
    ...
}

Task(B){
    ...
}
#define APP1_STOP_SEC_SHARED_DATA
#include "MemMap.h"

```

Code Example 21.1: Placing Task and OS-Application data into the same section

In MemMap.h the 'virtual' section APP1_START_SEC_SHARED_DATA is mapped to a real section name using the mechanism supported by the compiler. Typically this is using compiler vendor specific pragmas. Code Example 21.2 shows how this can be done.

```

#ifndef APP1_START_SEC_SHARED_DATA
    #pragma section data="APP1_data"
#endif

```

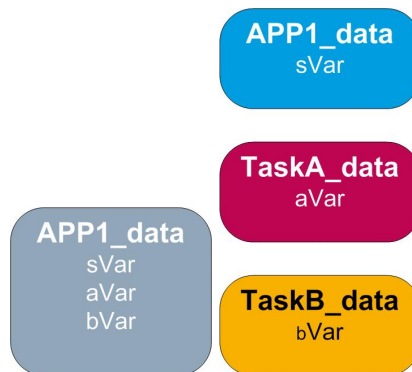
Code Example 21.2: Mapping to locator sections - same section

Figure 21.3(a) shows how all data is placed into the same section.

In this example, there is a single data section, APP1_data which contains the program data. In practice, the compiler may place data into different sections automatically. For example, zero initialized data is often placed in a section called .bss.



You must ensure that the default memory sections used by the compiler/linker are accounted for in the pragmas in MemMap.h. Any variables that are allocated during compilation, but not expressly placed in a named section, will be aggregated automatically and placed in the standard sections.



(a) Data located in the same section (b) Data located in unique sections

Figure 21.3: Placing data into sections

21.3.2 Placing Task and OS-Application Data into unique sections

Now assume that you want to separate each Task's data from the OS-Application's data. Code Example 21.3 shows how to place an OS-Application shared data item and a Task-specific private data into individual sections.

```

#include <Os.h>

/* Place OS-Application shared data in one section */
#define APP1_START_SEC_SHARED_DATA
#include "MemMap.h"
    uint32 aVar;
#define APP1_STOP_SEC_SHARED_DATA
#include "MemMap.h"

/* Place TaskA's data in a different section */
#define TASKA_START_SEC_PRIVATE_DATA
#include "MemMap.h"
    uint32 aVar;
#define TASKA_STOP_SEC_PRIVATE_DATA
#include "MemMap.h"

/* Place TaskB's data in a different section */
#define TASKB_START_SEC_PRIVATE_DATA
#include "MemMap.h"
    uint32 bVar;
#define TASKB_STOP_SEC_PRIVATE_DATA
#include "MemMap.h"

Task(A){
    ...
}

```

```

Task(B){
    ...
}

```

Code Example 21.3: Placing Task and OS-Application data in unique sections

MemMap.h now needs to map each of these three sections as shown in Code Example 21.4.

```

#ifdef APP1_START_SEC_SHARED_DATA
    #pragma section data="APP1_data"
#endif
#ifdef TASKA_START_SEC_PRIVATE_DATA
    #pragma section data="TaskA_data"
#endif
#ifdef TASKB_START_SEC_PRIVATE_DATA
    #pragma section data="TaskB_data"
#endif

```

Code Example 21.4: Mapping to locator sections - unique sections

Figure 21.3(b) shows which data lives in which section.

Merging sections in MemMap.h

After carefully ensuring that all Tasks and application data can be located separately, you may decide that you really did want to protect data at the level of an OS-Application. In this case, all the data from the Tasks and the OS-Application itself can be aggregated into a single linker section in MemMap.h as shown in Code Example 21.5.

```

#ifdef APP1_START_SEC_SHARED_DATA
    #undef APP1_START_SEC_SHARED_DATA
    #define START_SECTION_ALL_NON_TRUSTED_DATA
#endif
#ifdef TASKA_START_SEC_PRIVATE_DATA
    #undef TASKA_START_SEC_PRIVATE_DATA
    #define START_SECTION_ALL_NON_TRUSTED_DATA
#endif
#ifdef TASKB_START_SEC_PRIVATE_DATA
    #undef TASKB_START_SEC_PRIVATE_DATA
    #define START_SECTION_ALL_NON_TRUSTED_DATA
#endif
...
/* Name the system section with a compiler primitive */
#ifdef START_SECTION_ALL_NON_TRUSTED_DATA
    #pragma section data="APP1_data"
#endif

```

Code Example 21.5: Mapping to locator sections - Merged sections

This version of MemMap.h achieves the same result as that shown in Code Example 21.2, but the decision about how to group sections has been made later in the development process.

In this case (if your protection policy changes between different projects) the only change you would need to make is to MemMap.h. Of course, this flexibility is paid for by the additional complexity of marking up all your application's data as shown in Code Example 21.3. Of course, you are free to choose a hybrid approach, for example placing Task data and ISR data in different sections.

21.4 Locating

The regions you define in MemMap.h need to be accessible after linking. Typically this can be done in your linker control file by generating a label to mark each section.

There are primarily two models that are used by linkers to specify a memory region:

Base and Bound uses the base address of the section and the size (bound) of the section. When the bound is added to the base, this gives the end address of the section.

Start and End Address uses the base address of the section and the end address of the section directly.

Both schemes can be supported providing that the start and end addresses of each section can be identified. These addresses must be accessible at runtime but are only known at the point the application is linked. Most linkers allow labels to be defined that are associated with particular addresses. For example, assuming that data is placed into individual sections but needs to be located in the same region of memory, the sections could be aligned like this:

```
label APP1_ram_start;
    section APP1_data;
    section TASKA_data;
    section TASKB_data;
label APP1_ram_end;
```



You should try to minimize the number of memory sections you need to protect at runtime as this will make your application simpler to understand and it will make context switches faster (because there will be less programming of the MPU to do).

The labels define the symbolic names that will need to be accessible to the Os_Cbk_SetMemoryAccess() callback discussed in the following section.

21.5 Interacting with an MPU or MMU

RTA-OS3.x places the programming of the MPU in the user domain. This provides complete flexibility in the protection model. For example, you might want to have higher protection than offered by AUTOSAR OS R3.x (e.g. stopping one OS-Application from reading data owned by another OS-Application) or you might want to have lower protection (e.g. simply prevent erroneous writes to OS-Application data).

When memory protection is configured, an implementation of the `Os_Cbk_SetMemoryAccess()` callback function must be provided. RTA-OS3.x will call this function each time the memory protection regions need to be changed - typically on every context switch.

There are two basic models used by MPUs:

1. non-privileged code is *denied* memory access by default and access rights must be specifically granted.
2. non-privileged code is *granted* memory access by default and access rights must be specifically denied.

The model can impact the amount of reprogramming work that is needed at runtime. For non-trusted OS-Applications, the AUTOSAR OS R3.x model denies write access to all memory except that used by the Tasks and ISRs they own.

If the first model is used, then each of these sections needs to be opened up for use, so you should limit the number of regions that have to be opened. If the second model, then all non-accessible code needs to be closed off for use, so you should limit the number of sections that need to be closed off. In both cases, it is advisable to place all trusted OS-Application data in the same section where possible. This will minimize the number of MPU registers that need to be re-programmed on each context switch and therefore have optimal performance.

```
FUNC(void, OS_APPL_CODE)
Os_Cbk_SetMemoryAccess(Os_UntrustedContextRefType
    ApplicationContext) {
    if (ApplicationContext->Application == a) {
        /* Set memory protection regions that apply for a generally */
        SET_UNTRUSTED_WRITE_RANGE(a_BASE, a_SIZE); /* Example */
    }
}
```

Code Example 21.6: The `Os_Cbk_SetMemoryAccess` callback

21.5.1 Default Implementations

If the target micro controller has an MPU then RTA-OS3.x can provide a sample implementation of `Os_Cbk_SetMemoryAccess()` that implements the mandatory part of the AUTOSAR OS R3.x memory protection scheme:

Read access is unrestricted

Write access is unrestricted for trusted OS-Applications. Tasks and ISRs in non-trusted OS-Applications may only write to:

- their own data sections
- their own stack
- data sections defined by their owning OS-Application

Execute access is unrestricted

The sample `Os_Cbk_SetMemoryAccess()` is generated using:

```
C:\>rtaosgen --samples:[Os_Cbk_SetMemoryAccess] --target:<target>
```

21.6 Interaction with Stack Monitoring

If a stack fault is detected by stack monitoring in Scalability Class 1 or 2, then, by default, the OS calls `ShutdownOS()` with the status `E_OS_STACKFAULT`².

In Scalability Classes 3 and 4, a stack fault will be detected by memory protection as erroneous write access to the stack. In this case, RTA-OS3.x calls the `ProtectionHook()` with the status `E_OS_STACKFAULT`.

Figure 21.4 shows the three types of behavior that can occur when a stack fault is detected.

21.7 Checking Memory Access

Applications can be programmed defensively against memory problems by checking whether they have the necessary permissions to access memory at runtime.

RTA-OS3.x provides two API calls to do this:

1. `CheckTaskMemoryAccess()` checks whether a named Task can access a specified memory range

²In RTA-OS3.x this behavior can be overridden at configuration time so that the `Os_Cbk_StackOverrunHook()` is called instead (which allows you to debug the stack overrun rather than simply be made aware that it has occurred).

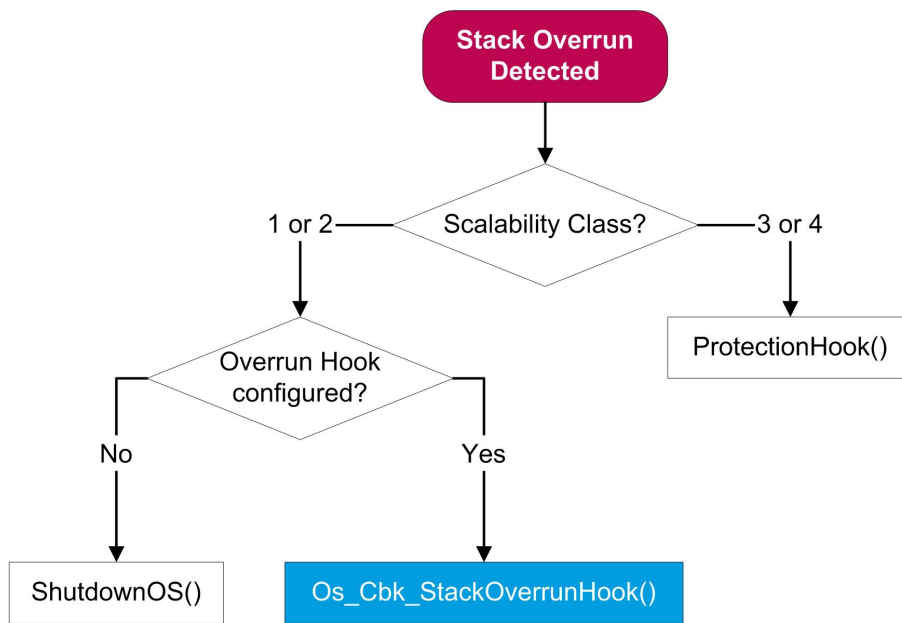


Figure 21.4: Decision tree for stack overrun handling

2. `CheckISRMemoryAccess()` checks whether a named ISR can access a specified memory range

The calls use a base and bound model for checking access. The base defines the start of the memory region to check and the bound defines the number of bytes of memory to check. The access rights are returned by the function. Code Example 21.7 shows how to check whether a Task can access a data structure.

```

uint32 DataValues[100];
...
rights = CheckTaskMemoryAccess(MyTask,&DataValues,
    sizeof(DataValues));
  
```

Code Example 21.7: Checking access to a data structure

21.7.1 Setting Access Permissions

RTA-OS3.x does not know where your code is located or what permissions you have chosen to apply. On each call to `CheckTaskMemoryAccess()` or `CheckISRMemoryAccess()`, RTA-OS3.x calls the `Os_Cbk_CheckMemoryAccess()` callback so that user code can perform the check.

The callback needs to:

- check the permitted memory ranges allowed for the OS-Application (or Task or ISR) depending on what protection scheme you want to enforce.
- Set the permissions accordingly.

RTA-OS3.x provides 4 permission constants that can be used to set the access permissions:

Constant	Definition
OS_ACCESS_EXECUTE	The memory range is executable
OS_ACCESS_READ	The memory range is readable
OS_ACCESS_STACK	The memory range is stack space
OS_ACCESS_WRITE	The memory range is writable

Permissions can be combined using C's bit-wise operators, for example if a section is readable and writeable this can be expressed as:

```
rights = OS_ACCESS_READ & OS_ACCESS_WRITE
```

Similarly, you can also deny permissions using tilde, for example if a section is readable and *not* writeable then this can be expressed as:

```
rights = OS_ACCESS_READ & ~OS_ACCESS_WRITE
```

Code Example 21.8 shows how the `Os_Cbk_CheckMemoryAccess()` callback might be written.

```
FUNC(AccessType, OS_APPL_CODE)
Os_Cbk_CheckMemoryAccess(ApplicationType Application,
                          TaskType TaskID,
                          ISRType ISRID,
                          MemoryStartAddressType Address,
                          MemorySizeType Size) {

    AccessType Access = 0u;

    /* Check for stack space in address range */
    if ((Address >= STACK_BASE)
        && (Address < STACK_BASE + STACK_SIZE))
    || ((Address+Size >= STACK_BASE)
        && (Address < STACK_BASE + STACK_SIZE)) {
        Access |= OS_ACCESS_STACK;
    }

    /* Address range is only writeable if it is not in ROM */
    if ((Address >= ROM_BASE)
        && (Address < ROM_BASE + ROM_SIZE))
    || ((Address+Size >= ROM_BASE)
        && (Address < ROM_BASE + ROM_SIZE)) {
```



```

        Access |= ~OS_ACCESS_WRITE;
    }

    switch (Application) {
        case APP1:
            ...
            Access |= OS_ACCESS_READ | OS_ACCESS_EXECUTE;
            break;
        case APP2:
            ...
            Access |= ~OS_ACCESS_READ | OS_ACCESS_EXECUTE;
            break;
    }
    return Access;
}

```

Code Example 21.8: The Os_Cbk_CheckMemoryAccess callback

21.7.2 Checking Access Permissions

The access permissions that are returned by `CheckTaskMemoryAccess()` and `CheckISRMemoryAccess()` can be used with one of 4 macros to determine access as shown in the following table:

Macro	Definition
<code>OSMEMORY_IS_EXECUTABLE</code>	True if the memory can be executed, false otherwise
<code>OSMEMORY_IS_READABLE</code>	True if the memory can be read, false otherwise
<code>OSMEMORY_IS_STACKSPACE</code>	True if the memory is stack space, false otherwise
<code>OSMEMORY_IS_WRITETABLE</code>	True if the memory can be written, false otherwise

All of the macros will only evaluate to true only if the full range of locations satisfies the access type. For example, if one byte of `DataValues` was not readable then `OSMEMORY_IS_READABLE(rights)` would return false.

21.8 Summary

- Memory protection requires the placement of data into named sections.
- Applications can use the AUTOSAR memory mapping concept to place data into sections.
 - Any number of sections can be defined.
 - The sections can follow any naming convention you require.

- You can aggregate sections in application code, in `MemMap.h` or a combination of both.
- Locating can further aggregate different memory areas into a single block of data to protect.
- The locator needs to export labels that you can use to program the memory ranges and access types for the MPU.
- RTA-OS3.x uses the callback `Os_Cbk_SetMemoryAccess()` to program the MPU on each context switch.

22 Contacting ETAS

22.1 Technical Support

Technical support is available to all users with a valid support contract. If you do not have a valid support contract, please contact your regional sales office (see Section 22.2.2).

The best way to get technical support is by email. Any problems or questions about the use of the product should be sent to:

`rta.hotline.uk@etas.com`

If you prefer to discuss your problem with the technical support team, you call the support hotline on:

+44 (0)1904 562624.

The hotline is available during normal office hours (0900-1730 GMT/BST).

In either case, it is helpful if you can provide technical support with the following information:

- your support contract number;
- your .xml and/or .rtaos configuration files;
- the command line which caused the error;
- the version of the ETAS tools you are using;
- the version of the compiler tool chain you are using;
- the error message you received (if any); and
- the file Diagnostic.dmp if it was generated.

22.2 General Enquiries

22.2.1 ETAS Global Headquarters

ETAS GmbH

Borsigstrasse 14
70469 Stuttgart
Germany

Phone: +49 711 89661-0
Fax: +49 711 89661-300
WWW: www.etas.com

22.2.2 ETAS Local Sales & Support Offices

Contact details for your local sales office and local technical support team (where available) can be found on the ETAS web site:

ETAS subsidiaries www.etas.com/en/contact.php
ETAS technical support www.etas.com/en/hotlines.php

Index

A

Access Rights, [308](#)
Alarms, [142](#)
 Absolute, [147](#)
 Action on expiry, [143](#)
 Activating Tasks, [143](#)
 Auto-starting, [151](#)
 Callbacks, [144](#)
 Canceling, [152](#)
 Cyclic, [148](#), [151](#)
 Incrementing Counters, [145](#)
 Periodic, *see* Cyclic
 Relative, [147](#), [150](#)
 Setting Events, [144](#)
 Single-shot, [147](#), [150](#)
API Protection, *see* Service Protection
Application Modes, [217](#)
Applications, *see* OS-Applications
Assistants, [36](#)
AUTOSAR, [20](#)
 Include file dependencies, [38](#)
 Operating System, [20](#)
 Scalability Class, [20](#)
AUTOSAR includes
 Compiler.h, [38](#)
 Compiler_Cfg.h, [38](#)
 MemMap.h, [40](#)
 Os_Compiler_Cfg.h, [38](#)
 Os_MemMap.h, [40](#)
 Platform_Types.h, [38](#)
 Std_Types.h, [38](#)
AUTOSAR OS includes
 Os.h, [43](#)
 Os_Cfg.h, [43](#)
 Os_MemMap.h, [43](#)

C

C Startup Code, [212](#)
Compilation, [45](#)
Compiler, [36](#)
Configuration Files, [29](#)
 Project Files, [30](#)
 XML, [29](#)

Conformance Classes, [56](#)
Context switch, [73](#)
Counter Attributes
 Accessing at runtime, [137](#)
 MAXALLOWEDVALUE, [129](#)
 MINCYCLE, [129](#)
 TICKSPERBASE, [129](#)
Counter Driver, [186](#)
Counters, [128](#)
 Cascading, [145](#)
 Free running timers, [139](#)
 Getting the value of, [138](#)
 Hardware, [128](#)
 Hardware Driver, [134](#)
 Schedule Tables and Alarms on,
 [158](#)
 Software, [128](#)
 Software Driver, [130](#)
 Ticks, [128](#)
CPU Clock rate, *see* Instruction Rate
Critical Section, [103](#)

D

Deadline, [285](#)
Deadline Monotonic, [49](#)
Deadlock
 Freedom from, [105](#)
Debugging
 API Usage, [225](#)
 ORTI, [251](#)
 RTA-TRACE, [255](#)
 Stack Monitoring, [233](#)
 Time Monitoring, [242](#)
Development process, [24](#)
Deviation, [176](#)

E

ECU Link
 Debugger, [270](#)
 Serial, [271](#)
 Asynchronous Dump, [272](#)
 Byte-wise, [273](#)
 Driver Callbacks, [273](#)

- Interrupt Driven, 275
 - Polling, 274
- Error Codes, 225
- Error Handling, 225
- Events, 119
 - Clearing, 125
 - Multiple Waits, 122
 - Setting, 124
 - Waiting On, 120
- Execution Budget, 289
- Expiry Point
 - Processing, 167
 - Runtime Adjustment, 180
- Extended Status, 225
- Extended Tasks
 - Risk of deadlock, 123
 - Simulation using Basic Tasks, 126

F

- Fixed Priority, 49
- Free running timer, 139

G

- Generated files, 42

H

- Hooks
 - Error, 226
 - PostTask, 81
 - PreTask, 81
 - Shutdown, 222
 - Startup, 217

I

- Idle Mechanism, 80
 - Limitations, 80
- Imprecise Computation, 247
- Instruction Cycle Rate, 248
- Instruction Rate, 242
- Internal Resources, 110
 - Shared with interrupts, 110
 - Stack Saving with, 113
- Interrupt Source
 - Disabling, 283
- Interrupts, 44, 88
 - Category 1, 89, 95
 - Category 2, 89, 96
 - Compiler Directives, 95
 - Default Interrupt, 99
 - Enabling and disabling, 98
 - Multi-level, 88
 - Nested, 88
 - Priority, 89
 - Register Sets, 99
 - Single-level, 88
- ISR, 96

L

- Library, 36
 - Name of, 43
- Linked Resources, 108
- Locating, 324

M

- Memory
 - Checking access permissions, 326
- Memory Protection, 317
 - MPU Usage, 325
 - Section Definition, 320
- MISRA, 34
- MissingReferences.xml, 31
- MPU
 - Programming for protection, 325
- Mutual Exclusion, 103

O

- Optimization
 - Fast Task Termination, 79
 - Omit Schedule() API, 78
 - Wait Event Stack, 67
- Optimizations
 - Customized Scalability Classes, 278
 - Stack Reduction, 112
- ORTI, 251
- OS-Application, 300
- OS-Applications, 300
 - Access Rights, 308
 - Accessing Applications, 303

- Error Hook, [313](#)
- Restart Task, [310](#)
- Safety Risks, [307](#)
- Shutdown Hook, [313](#)
- Startup Hook, [312](#)
- Termination, [309](#)
- OS-level, [92](#)
- OSEK, [17](#)
 - Operating System, [18](#)

P

- Priority Ceiling Protocol, [104](#)
- Priority Inversion, [104](#)

R

- Rate Monotonic, [49](#)
- Register Sets
 - Saving in ISRs, [99](#)
 - Saving in Tasks, [83](#)
- Reports, [34](#)
- RES_SCHEDULER, [114](#)
- Reset, [211](#)
- Resources, [103](#)
 - Ceiling Priority, [104](#)
 - Internal, [110](#)
 - Linked, [108](#)
 - Nesting locks, [108](#)
 - Race Conditions, [117](#)
 - Sharing with Interrupts, [105](#)
- Response Time, [285](#)
- Restart Task, [310](#)
- Restarting, [223](#)
- RTA-TRACE, [255](#)
 - Burst Mode, [257](#)
 - Categories, [264](#)
 - Classes, [258](#)
 - Configuration, [256](#)
 - ECU Links, [269](#)
 - Enumerations, [267](#)
 - Filters, [259](#)
 - Format Strings, [266](#)
 - Free-Running Mode, [257](#)
 - Instrumentation, [255](#)
 - Intervals, [262](#)
 - Task Tracepoints, [262](#)

- Tracepoints, [261](#)
 - Triggering Mode, [257](#)
- RTA-TRACE Configuration, [28](#)
- rtaoscfg, [24](#)
 - Builder, [42](#)
- rtaosgen, [41](#)
 - Invoking from rtaoscfg, [29](#)

S

- Sample Code, [43](#)
- Samples, [43](#)
- Schedulability Analysis, [286](#)
- Schedule Table
 - Deviation, [176](#)
 - Explicit Synchronization, [173](#)
 - Implicit Synchronization, [171](#)
 - Stopping, [167](#)
- Schedule Tables, [156](#)
 - Absolute Start, [163](#)
 - Attributes, [158](#)
 - Disable Synchronization, [183](#)
 - Expiry Points, [156](#), [160](#)
 - Relative Start, [165](#)
 - State, [184](#)
 - Switching, [169](#)
 - Synchronization, [160](#), [180](#)
 - Explicit, [160](#)
 - Implicit, [160](#)
 - None, [160](#)
 - Synchronized Start, [166](#)
- Scheduler, [49](#)
- Scheduling
 - Cooperative, [51](#), [77](#)
 - Non-Preemptive, [50](#)
 - Preemptive, [49](#)
- Scheduling Policy, [49](#)
- Semaphore, *see* Resources
- Service Protection, [299](#)
- Shutdown, [222](#)
- Shutdown Hook, [222](#)
- Single-Stack, [61](#)
 - Extended Tasks, [62](#)
- Stack, [61](#)
 - Allocation, [67](#)

- Default Allocation, [70](#)
 - Mandatory Figures, [64](#)
 - Measurement of, [238](#)
 - Optimization, [67](#)
 - Per Task and ISR Figures, [66](#)
 - Reducing Size, [113](#), [114](#)
 - Stack Management
 - Overruns, [70](#)
 - Stack Monitoring
 - Interaction with Memory Protection, [326](#)
 - Stack Resource Protocol, [104](#)
 - Standard Resources
 - Stack Saving with, [114](#)
 - Standard Status, [225](#)
 - Starting RTA-OS3.x, [44](#)
 - StartOS, [216](#)
 - Startup
 - Activating Tasks, [60](#)
 - Alarms, [151](#), [220](#)
 - Schedule Tables, [221](#)
 - Tasks, [219](#)
 - Startup Hook, [217](#)
 - Static Interface
 - Software Counters, [132](#)
 - Status
 - Extended, [225](#)
 - Standard, [225](#)
 - Stopwatch, [243](#)
 - Scaling, [244](#)
 - Uncertainty, [243](#)
 - Stopwatch Speed, [248](#)
 - SystemCounter, [138](#)
 - SystemTimer, *see* SystemCounter
- T**
- Tasks, [44](#), [49](#)
 - Activation, [54](#)
 - Basic, [51](#)
 - Entry Function, [72](#)
 - Extended, [52](#)
 - Fast Termination, [79](#)
 - Maximum supported, [57](#)
 - Optimization, [57](#), [60](#), [62](#)
 - Queuing Activations, [55](#)
 - Register Sets, [83](#)
 - Sharing Priorities, [54](#)
 - States, [52](#)
 - Synchronization, [52](#)
 - Termination, [54](#), [78](#)
 - Termination
 - Forced, [282](#)
 - Tick/Time Conversion, [139](#)
 - Time Frame, [292](#)
 - Time Measurement
 - Arbitrary Code, [246](#)
 - ISR, [244](#)
 - Tasks, [244](#)
 - Time Monitoring, [242](#), [248](#)
 - Budgets, [249](#)
 - Resetting Budgets, [250](#)
 - Time-base, [139](#)
 - Timing Protection, [285](#)
 - Active Detection, [292](#)
 - Blocking, [295](#)
 - Configure, [294](#)
 - Configuring Timing Protection, [294](#)
 - Execution Budget, [289](#)
 - Lock Budgets, [295](#)
 - Time Frame, [292](#)
 - Toolchain, [36](#)
 - Trusted Functions, [314](#)
- U**
- Uncertainty, [243](#)
 - User-level, [92](#)
- V**
- Vector Table
 - Generation, [94](#)
 - Writing by hand, [94](#)