Safety-critical Control in Mixed Criticality Embedded Systems

Master Thesis Royal Institute of Technology Stockholm, Sweden

Emil Hjelm emilhje@kth.se

March 13, 2017



Master Thesis MMK2017:Z MDAZZZ

Safety-critical Control in Mixed Criticality Embedded Systems

Emil Hjelm

Approved: Examiner:		Supervisor:		
(datum) Martin Törngren		Bengt Eriksson		
	Uppdragsgivare:	Kontaktperson:		
	Alten	Detlef Scholle		

Abstract

Modern automotive systems contain a large number of Electronic Control Units, each controlling a specific system of a specific criticality level. To increase computational efficiency it is desired to combine multiple applications into fewer ECUs, this leads to mixed criticality embedded systems. The assurance of safety critical applications not being affected by non-critical applications on the same system is crucial.



Examensarbete MMK2017:Z MDAZZZ

Säkerhetskritisk kontroll i blandkritiska inbyggda system

Emil Hjelm

Godkänt: Examinator: Handledare:		Handledare:
(datum)	tum) Martin Törngren Bengt Eriksson	
	Uppdragsgivare:	Kontaktperson:
	Alten	Detlef Scholle

Sammanfattning

Denna del kommer att inehålla en sammanfattning av arbetet på svenska.

Preface

Credit where credit is due.

Emil Hjelm Stockholm

Contents

Pı	eface	e iii
\mathbf{A}	bbre	viations viii
1	Intr	oduction 1
	1.1	Background
		1.1.1 Definition of safety-critical systems
		1.1.2 Different levels of criticality
		1.1.3 EMC ² development board $\dots \dots 3$
		1.1.4 Platooning
	1.2	Problem statement
	1.3	Purpose
	1.4	Goals
		1.4.1 Team goal
		1.4.2 Individual goal
	1.5	Scope
	1.6	Research design
	1.7	Ethical considerations
2	Stat	te of the art
	2.1	Mixed criticality systems
		2.1.1 Economical benefits of MCS
		2.1.2 Sharing processor
		2.1.3 Different criticality on different processors 9
		2.1.4 Sharing memory
	2.2	Current system
		2.2.1 TrustZone
	2.3	Standards
		2.3.1 IEC 61508
		2.3.2 ISO 26262
		2.3.3 AUTOSAR

3	System design	16
4	Implementation	17
5	Results	18
6	Discussion 6.1 Information retrieval	19
7	7.1 MCS using virtualization	20 20
	7.4 Economical benefits for pursuing MCS	<i>2</i> U

List of Figures

2.1	Percentage of schedulable tasks. [5]	Ĝ
2.2	Flowchart of the boot sequence of the CPU. [24]	10
2.3	System overview of the MCS in place. [24]	11
2.4	AUTOSAR. [3]	15

List of Tables

2.1	ASIL as a function of severity, probability and controllability	13
2.2	ASIL cost heuristics	13

Abbreviations

Abbreviation Description ECU Electronic Control Unit MCS Mixed Criticality System EMC^2 Embedded Multi-Core systems for Mixed Criticality applications in dynamic and changeable real-time environments RTOS Real-Time Operating System GPOS General Purpose Operating System FPGA Field Programmable Gate Array SILSafety Integrity Level ASIL Automotive Safety Integrity Level Development Assurance Level DALVMMVirtual Machine Monitor EMC^2DP EMC² Development Platform Rate Monotonic RMDM Deadline Monotonic FPFixed Priority EDF Earliest Deadline First AUTOSAR AUTomotive Open System ARchitecture

Introduction

This chapter will introduce the subject of mixed criticality embedded systems and the EU project "EMC²" to the reader.

1.1 Background

Today, modern automotive systems contain around 70-100 Electric Control Units (ECU)s [16]. Each ECU controls a subsystem of a specific criticality level such as safety-critical anti-lock brake system, or non-critical entertainment systems [20]. Having the ECUs isolated ensures that the numerous critical and non-critical applications do not interfere with each other, thus it is a simple task to certify an individual ECU. However, this approach leads to an inefficient use of system resources and expensive system implementation [8]. In order to lower the cost of the collective system and increase system efficiency (utilization), applications of different criticality levels can be integrated into a single multicore platform, leading to a Mixed Criticality System (MCS). However, this approach increases system complexity, and hinders the certification of safety-critical systems [24]. In order to facilitate the design, test, and certification of such systems, spatial and temporal partitioning can be used in the architecture of the system as described by [24].

Protecting the integrity of a component from the faults of another is desired in all systems hosting multiple applications. However, it is of higher significance if the different applications have different criticality levels. Without such protection all components on the same system would need to be engineered to the standards of the highest criticality level, potentially massively increasing development costs [8].

The EU project "Embedded Multi-Core systems for Mixed Criticality applications in dynamic and changeable real-time environments" (EMC²) was founded in order to "find solutions for dynamic adaptability in open systems, provide handling of mixed criticality applications under real-time conditions, scalability and utmost flexibility, full scale deployment and management of integrated tool chains, through the entire lifecycle" [20].

1.1.1 Definition of safety-critical systems

The term "safety-critical system" has many definitions, most quite similar. Most definitions relate to systems with the potential to harm humans if the system malfunctions. According to [11] it is defined as "A system in which any failure or design error has the potential to lead to loss of life." Further, [9] defines safety-critical systems as "A computer, electronic or electromechanical system whose failure may cause injury or death to human beings." A Wikipedia article, [22], defines a safety-critical system (or "life-critical system") as a system whose failure or malfunction may result in one (or more) of the following outcomes:

- death or serious injury to people
- loss or severe damage to equipment/property
- environmental harm

In this thesis, a safety-critical system will be defined as "a system whose failure may cause injury or death to human beings."

1.1.2 Different levels of criticality

Different names of levels of criticality are typically Safety Integrity Level (SIL), Automotive Safety Integrity Level (ASIL) and Development Assurance Level (DAL). The IEC 61508 standard [13] defines four different levels and the ISO 26262 standard [14] and the DO-178C standard [10] define five different levels each. These levels range from low or no hazard up to life-threatening or fatal in the event of a malfunction requiring the highest level of assurance that the dependent safety goals are sufficient and have been achieved.

In this report the number of criticality levels will be restricted to two: "safety-critical" and "non-critical". This is due to the constraints presented in section 1.1.3 and 1.5.

1.1.3 EMC² development board

As a part of the EMC² project, Alten has developed a system for handling applications of mixed criticality on the same piece of hardware. The MCS developed at Alten is implemented on a Xilinx Zynq-7000 [23]. The development board is called EMC² Development Platform, or EMC²DP. It employs two operating systems to handle applications of different criticality. A General Purpose Operating System (GPOS) for non-critical applications and a Real-Time Operating System (RTOS) for safety-critical applications. A Virtual Machine Monitor (VMM) is used to alternate between the two.

Peripherals connected to the board are separated between safety-critical and non-critical via ARM TrustZone [1].

The board also has a Field Programmable Gate Array (FPGA).

For more detailed information, see the report by [24].

1.1.4 Platooning

"The platooning concept can be defined as a collection of vehicles that travel together, actively coordinated in formation. Some expected advantages of platooning include increased fuel and traffic efficiency, safety and driver comfort" [7].

1.2 Problem statement

An ideal MCS ensures partitioning between different criticality levels while still sharing resources efficiently.

The MCS developed at Alten (EMC²DP) 1.1.3 switches Operative System (OS) to enable partitioning between safety-critical and non-critical applications, which takes about 2 μs . This mode switch introduces additional deadlines which makes processor scheduling more difficult.

To evaluate the performance of the system, a distance keeping control algorithm for platooning will be implemented on it. A demonstrator will be constructed in the form of a RC car capable of following a vehicle in front of it at a specified distance. If the lead car exceeds a predefined maximum speed or deviates from the road, the following car should not exceed the maximum speed. The performance of the embedded controller and the control

algorithm will be measured during heavy non-critical computational load, and without any non-critical load altogether.

It should be verified that no matter the computational load and eventual crashes of the Linux based non-critical system, the distance keeping algorithm on the RTOS never crashes. It should also be investigated at how high frequencies the control algorithm can operate while still maintaining functionality on the GPOS.

This problem leads to the research question:

• How well can a safety-critical control system perform when implemented on a mixed criticality system using virtualization?

alternatively:

• "How, in a disciplined way, to reconcile the conflicting requirements of partitioning for safety assurance and sharing for efficient resource usage?" [8]

alternatively:

• Is virtualization an efficient approach when trying to reconcile the conflicting requirements of partitioning for safety assurance and sharing for efficient resource usage when implementing a safety-critical control system?

1.3 Purpose

Reducing the amount of computers in automotive systems would have many effects. Manufacturing costs would decrease and with fewer physical components maintenance costs would also decrease. However, the system complexity would increase and thereby increasing time and cost to design the system.

SafeCOP (Safe Cooperating Cyber-Physical Systems) is an European project that targets cyberphysical systems-of-systems whose safe cooperation relies on wireless communication [15]. SafeCOPs Use Case 3 (UC3) regarding "Vehicle control loss warning" together with the EMC2 goals tie well in with the problem statement and use case described in 1.2.

1.4 Goals

In this project there are both team goals and individual goals that do not always necessarily align with each other.

1.4.1 Team goal

The team consists of five master thesis students. The students areas of work are: control theory and system modeling, data aggregation, safety-critical communication in MCS, lane detection and finally safety-critical control in MCS. Together the team will build a vehicle capable of following a vehicle ahead of it while keeping inside road markers.

1.4.2 Individual goal

Verify quantitatively the performance of safety-critical distance keeping controller, see 1.6. Solve the problems described in 1.2.

1.5 Scope

The work of this thesis and the implementation on the demonstrator will build upon the work of Youssef Zaki [24].

The embedded computer is constrained to the Xilinx Zynq-7000 SoC ¹. The thesis is produced at Alten AB.

1.6 Research design

The plan is to make an confirmatory investigation using quantitative data/operations with a deductive approach. This is a quantitative research method where data is gathered during experiments and from simulations of the environment [12].

The position of the demonstrator will be read by a separate sensor of the same type as the one on the demonstrator. The performance of the control system and the embedded controller will be measured and compared with the same system without any non-critical computational load. This will also be done for a simulation of the system. The measures regarding control system performance will consist of

¹https://www.xilinx.com/products/silicon-devices/soc/zynq-7000.html

- Response time
- Overshoot
- Settling time

Data points for the performance of the embedded controller will be extracted from the RTOS, and the measures will consist of

- Missed deadlines
- CPU utilization

1.7 Ethical considerations

When designing a MCS it is crucial to ensure that errors made by a lower criticality application cannot propagate to higher criticality applications. This could have catastrophic consequences. Because of this the requirement of partitioning must have higher priority than the need of sharing.

State of the art

This chapter will go through relevant articles and already known knowledge on the subject of mixed criticality systems. It will also explain the EMC2DP.

2.1 Mixed criticality systems

A MCS is achieved by letting applications of different criticality share resources. These resources could be the CPU, memory, peripherals, input/output ports etc. The most explored area is sharing the CPU between multiple criticality levels [8]. The benefit of combining previously distributed systems is higher resource efficiency, which leads to economical benefits.

2.1.1 Economical benefits of MCS

Potential benefits with pursuing MCS as opposed to distributed systems are reduced physical space required, reduced weight, reduced heat generation, reduced power consumption and reduced production costs [8]. This would all ultimately lead to economical benefits.

Potential downsides are increased complexity which could lead to higher system design costs. Building applications on the same platform to share resources could require engineering teams to work more closely together, potentially leading to administrative difficulties and costs. This needs to be investigated and could vary from industry to industry. To combat the potential downsides, the EMC² project aims at creating platforms for easier development of MCS.

The EMC² project lists several goals [21]:

- Reduce the cost of the system design by 15%
- Reduce the effort and time required for re-validation and re-certification of systems after making changes by 15%
- Manage a complexity increase of 25% with 10% effort reduction
- Achieve cross-sectorial reusability of Embedded Systems devices and architecture platforms that will be developed using the ARTEMIS JU results.

2.1.2 Sharing processor

To deal with many different tasks needing processor time, different schedulers can be used to appropriately distribute processor time among the tasks.

Conventional scheduling

Fixed priority Deadline monotonic Rate monotonic Earliest deadline first

Mixed criticality scheduling

The area of sharing the processor in MCS was first explored by Steve Vestal [19] in 2007. His paper showed that neither Rate Monotonic (RM) nor Deadline Monotonic (DM) priority assignment was optimal for MCS; however Audsley's optimal priority assignment algorithm [2] was found to be applicable.

In 2008 Baruah and Vestal [6] showed that EDF (Earliest Deadline First) does not dominate FP when criticality levels are introduced, and that there are feasible systems that cannot be scheduled by EDF.

1. (audsley) [2] 2. EDF-VDL 3. SMC 4. AMC-max 5. AMC-rtb [5]

"This figure plots the percentage of task sets generated that were deemed schedulable for a system of 20 tasks, with on average 50% of those tasks having high criticality and each task having a high criticality execution time that is twice its low criticality execution time. The compared approaches are (from least effective to most effective): CrMPO which assigned priorities in criticality order, SMC-NO (static mixed criticality with no run-time monitoring) which is Vestal's original approach, SMC which is an adaptation of Vestal's approach in which LO-criticality tasks are monitored at run-time and are prevented from executing for more than C(LO), and AMC-rtb and AMC-max which are the two methods introduced in the previous paragraph

(AMC for adaptive mixed criticality). In the graph the UB-H&L line bounds the maximum possible number of schedulable task sets. It serves to illustrate the quality of the AMC-max approach." [8]

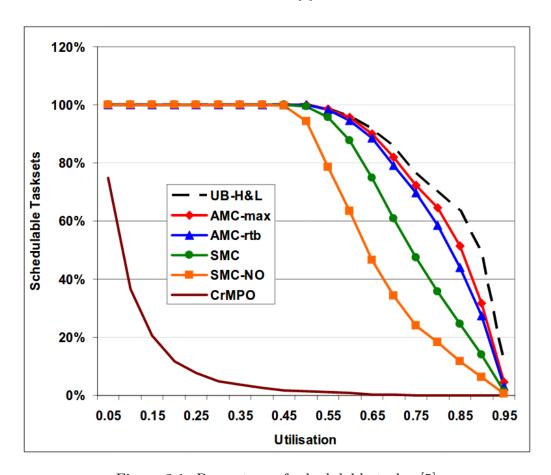


Figure 2.1: Percentage of schedulable tasks. [5]

For a more complete review of work done on MCSs with a shared processor, see the paper by Alan Burns [8].

2.1.3 Different criticality on different processors

2.1.4 Sharing memory

2.2 Current system

The EMC2DP uses two operating systems, a RTOS for safety-critical applications and a GPOS for non-critical applications. A Virtual Machine

Monitor (VMM) or "Hypervisor" is used to alternate between safety-critical RTOS and non-critical GPOS. The RTOS is TOPPERS FMP kernel [17], and the GPOS is a custom modified Linux distribution. The VMM used is SafeG [18], also developed by TOPPERS. It switches processor state via a hardware switch. See figure 2.2.

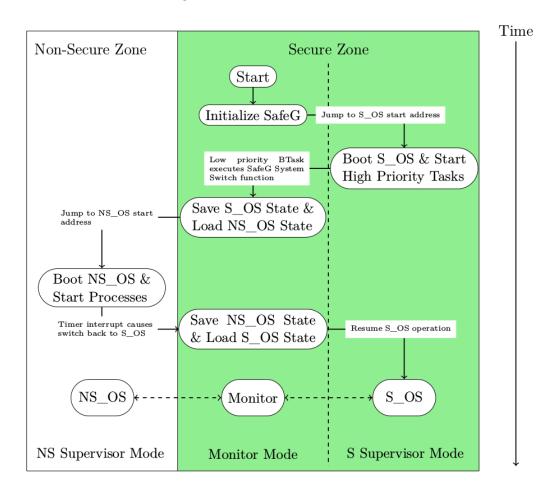


Figure 2.2: Flowchart of the boot sequence of the CPU. [24]

A Field Programmable Gate Array (FPGA) as interface between processor and board. Peripherals are separated as secure and non-secure using ARM TrustZone [1], see section 2.2.1.

An overview of the system can be seen in Figure 2.3.

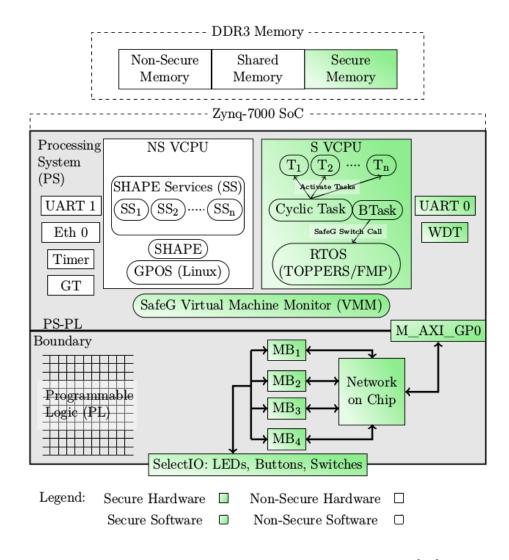


Figure 2.3: System overview of the MCS in place. [24]

2.2.1 TrustZone

"TrustZone is a security extension available in modern ARM processors that creates a security infrastructure designers can use to protect critical system assets. This infrastructure is achieved by enabling the partition of system components, both hardware and software, into either a Secure and a Normal world (or zone). Resources that are marked as normal are not permitted to access Secure zone components. This mechanism is enforced by the AMBA3 (Advanced Microcontroller Bus Architecture) AXI (Advanced eXtensible Interface) bus system. It contains an extra control signal for each of the read

and write channels (Non-secure or NS bits) that dictate the access rights Non-secure bus masters to the Secure slaves. Each processor with an enabled TrustZone security extension can be partitioned into a Normal and a Secure virtual CPU. The virtual processors execute in a time-multiplexed fashion, and use the "Monitor Mode" state to create an efficiently switching mechanism between Normal and Secure zones. The NS-bit, bit[0] of the Secure Configuration Register (SCR) in the System Control Coprocessor (CP15), controls the activation of secure state of the processor. Whenever the NS-bit set high, the processor state immediately switches to the Normal world. However, if the processor is in Monitor Mode, it remains in the Secure world regardless of the state of the SCR NS-bit. The processor state can nter monitor mode by either issuing a special instruction, SMC (Secure Monitor Call), in software or by hardware exception mechanisms such as IRQ, FIQ, external Data Abort, or external Prefetch Abort. In general, the software running in monitor mode serves the purpose of saving the state of the current world, and loading the state the other." [24]

2.3 Standards

Safety practices are becoming more regulated as industries adopt a standardized set of practices for designing and testing products.

2.3.1 IEC 61508

IEC 61508 [13] is intended to be a basic functional safety standard for electrical and electronic systems applicable to all kinds of industry. It defines four different safety integrity levels, SIL 1 being the least dependable up to SIL 4 which is the most dependable level.

2.3.2 ISO 26262

ISO 26262 [14] addresses the needs for an automotive-specific international standard that focuses on safety critical components. ISO 26262 is a derivative of IEC 61508.

ASILs

ISO 26262 describes five different levels relating to hazard and risk. Ranked from lowest (no) hazard to highest hazard, these levels are: QM, A, B, C and

D. A function is assigned an ASIL depending on the severity if the function fails, the probability that the function fails and the controllability of the function, see table 2.1.

Severity	Probability	Controllability			
Severity		C1	C2	С3	
	E1	QM	QM	QM	
S1	E2	QM	QM	QM	
	E3	QM	QM	Α	
	E4	QM	A	В	
	E1	QM	QM	QM	
S2	E2	QM	QM	A	
52	E3	QM	A	В	
	E4	A	В	С	
	E1	QM	QM	A	
S3	E2	QM	A	В	
	E3	A	В	С	
	E4	В	С	D	

Table 2.1: ASIL as a function of severity, probability and controllability.

The various integrity levels can be translated into integers (ASIL QM = 0; A = 1; B = 2; C = 3 and D = 4). If a hazard requires several components to fail, the added ASIL of these components is used to determine if there is an violation, assuming the components faults are statistically independent of each other. For example, a safety level ASIL B can be met by two independent components which each individually only meet ASIL A (and thus effectively A + A = B). [4]

The different ASILs relate to cost, see table 2.2.

Cost Heuristic	$\mathbf{Q}\mathbf{M}$	A	В	\mathbf{C}	D
Linear	0	10	20	30	40
Logarithmic	0	10	100	1000	10000
Experimental-I [4]	0	10	20	40	50
Experimental-II [4]	0	20	30	45	55

Table 2.2: ASIL cost heuristics.

Freedom from interference

In ISO 26262, Part 1, Definition 1.49, freedom from interference is defined as: Absence of cascading failures between two or more elements that could lead to the violation of a safety requirement.

element = "system or part of a system including components, hardware, software, hardware parts, and software units" (ISO 26262, Part 1, Definition 1.32)

cascading failure = "failure of an element of an item causing another element or elements of the same item to fail" (ISO 26262, Part 1, Definition 1.13)

2.3.3 AUTOSAR

"AUTOSAR (AUTomotive Open System ARchitecture) is a international development partnership of automotive interested parties founded in 2003. It pursues the objective of creating and establishing an open and standardized software architecture for automotive electronic control units (ECUs) excluding infotainment. Goals include the scalability to different vehicle and platform variants, transferability of software, the consideration of availability and safety requirements, a collaboration between various partners, sustainable utilization of natural resources, maintainability throughout the whole "Product Life Cycle"." [3]

The AUTOSAR Architecture distinguishes on the highest abstraction level between three software layers: Application, Runtime Environment (RTE) and Basic Software (BSW) which run on a Microcontroller. [3] See figure 2.4.

- The application software layer is mostly hardware independent.
- The RTE represents the full interface for applications.
- The BSW is divided in three major layers and Complex Drivers: Services, ECU Abstraction and Microcontroller Abstraction. Services are divided furthermore into functional groups representing the infrastructure for System, Memory and Communication Services.

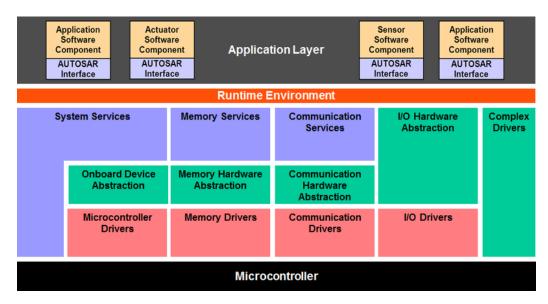


Figure 2.4: AUTOSAR. [3]

System design

This chapter will derive the design of the controller to be implemented.

Implementation

This chapter will describe the implementation of the control system in the demonstrator.

Results

This chapter will present results from the demonstrator to the reader.

Discussion

Discussion about the results produced by the thesis.

6.1 Information retrieval

Printing to SYSLOG takes time and affects the system that's being monitored.

Future work

This chapter will contain thoughts and ideas for future work building on this thesis or in the area of MCS in general.

7.1 MCS using virtualization

Facilitate for more than two different criticality levels. Examine different scheduling methods.

7.2 MCS using other means of partitioning

Examine limitations for other configurations of MCS, for example different CPUs for different criticality levels.

7.3 Amount of criticality levels

Research should be done to investigate how many different levels of criticality, n, to have on MCSs in different industries. In the automotive for example, n should be between 1 and 5 since ISO26262 defines 5 different ASILs. If the applications are spread uniformly across all criticality levels it might be of higher interest to have n closer to 5. Similarly, if the applications are heavily concentrated on a certain criticality level n probably should be closer to 2.

7.4 Economical benefits for pursuing MCS

It is not clear how much the potential economical benefit would be from pursuing MCS. The economical impacts of MCS might be different in different industries. It must be calculated more exactly how large the potential benefits would be to gauge the need for pursuing MCS.

Bibliography

- [1] ARM Ltd. ARM TrustZone, February 2017. https://www.arm.com/products/security-on-arm/trustzone.
- [2] N C Audsley. On priority assignment in fixed priority scheduling. *Information Processing Letters*, 79(1):39–44, 5 2001.
- [3] AUTOSAR. AUTOSAR Homepage, February 2017. http://www.autosar.org/.
- [4] Luís Silva Azevedo, David Parker, Yiannis Papadopoulos, Martin Walker, Ioannis Sorokos, and Rui Esteves Araújo. Exploring the Impact of Different Cost Heuristics in the Allocation of Safety Integrity Levels, pages 70–81. Springer International Publishing, Cham, 2014.
- [5] S. K. Baruah, A. Burns, and R. I. Davis. Response-time analysis for mixed criticality systems. In 2011 IEEE 32nd Real-Time Systems Symposium, pages 34–43, Nov 2011.
- [6] Sanjoy Baruah and Steve Vestal. Schedulability analysis of sporadic tasks with multiple criticality specifications, 2008.
- [7] Carl Bergenhem, Henrik Pettersson, Erik Coelingh, Cristofer Englund, Steven Shladover, and Sadayuki Tsugawa. Overview of platooning systems. In *Proceedings of the 19th ITS World Congress*, Vienna, Austria, October 2012.
- [8] Alan Burns and Robert I. Davis. Mixed criticality systems a review. Available online: https://www-users.cs.york.ac.uk/burns/review.pdf, July 2016.
- [9] Dictionary.com. safety-critical system, January 2017. http://www.dictionary.com/browse/safety-critical-system.

- [10] Software considerations in airborne systems and equipment certification. Standard, Radio Technical Commission for Aeronautics, Washington, DC, USA, December 2011.
- [11] Encyclopedia.com. safety-critical system, January 2017. http://www.encyclopedia.com/computing/dictionaries-thesauruses-pictures-and-press-releases/safety-critical-system.
- [12] Anne Håkansson. Portal of research methods and methodologies for research projects and degree projects, July 2013.
- [13] Functional safety of electrical/electronic/programmable electronic safety-related systems parts 1 to 7. Standard, International Electrotechnical Commission, Geneva, CH, April 2010.
- [14] Road vehicles functional safety part 9: Automotive safety integrity level (asil)-oriented and safety-oriented analyses. Standard, International Organization for Standardization, Geneva, CH, November 2011.
- [15] SafeCOP. SafeCOP part B, October 2016.
- [16] H. Thompson. Mixed criticality systems, February 2012.
- [17] TOPPERS Project, Inc. Introduction to the TOPPERS/FMP kernel, February 2017. https://www.toppers.jp/en/fmp-kernel.html.
- [18] TOPPERS Project, Inc. TOPPERS SafeG, February 2017. https://www.toppers.jp/en/safeg.html.
- [19] Burns Vestal. Preemptive scheduling of multi-criticality systems with varying degrees of execution time assurance, 2007.
- [20] W. Weber, A. Hoess, F. Oppenheimer, B. Koppenhöefer, B. Vissers, and B. Nordmoen. EMC2 a Platform Project on Embedded Microcontrollers in Applications of Mobility, Industry and the Internet of Things. In 2015 Euromicro Conference on Digital System Design, pages 125–130, August 2015.
- [21] Werner Weber. A Platform Project on Embedded Microcontrollers in Applications of Mobility, Industry and the Internet of Things, Mars 2015. https://artemis-ia.eu/publication/download/1131.pdf.
- [22] Wikipedia.com. Life-critical system, January 2017. https://en.wikipedia.org/wiki/Life-critical_system.

- [23] Xilinx Inc. Zynq-7000 All Programmable SoC, February 2017. https://www.xilinx.com/products/silicon-devices/soc/zynq-7000.html.
- [24] Youssef Zaki. An embedded multi-core platform for mixed-criticality systems: Study and analysis of virtualization techniques. Master's thesis, KTH, School of Information and Communication Technology (ICT), 2016.