Clereco

Cross-Layer Early Reliability Evaluation for the Computing cOntinuum

SEVENTH FRAMEWORK PROGRAMME

Project Number: FP7-611404

D2.3 – Definition of operation modes for future systems

Authors

G. Rafiq, T. Løkstad, K.-J. Alme, F. Reichenbach, N. Aymerich, R. Martinez, S. Ozdemir, A. Grasset, D. Gizopoulos, N. Foutris

Version 1.2 – 02/05/2014

Lead contractor: ABB

Contact person:

Gulzaib Rafiq
ABB Corporate Research
Bergerveien 12 1375 Billingstad Norway

Tel. +47 22872000
E-mail: gulzaib.rafiq@no.abb.com

Work package: WP2

Affected tasks: T2.2

Nature of deliverable[1]	[✓ R]	P	D	O
Dissemination level[2]	PU	PP	RE	[✓ CO]

[1] R: Report, P: Prototype, D: Demonstrator, O: Other

[2] **PU**: public, **PP:** Restricted to other programme participants (including the commission services), **RE** Restricted to a group specified by the consortium (including the Commission services), **CO** Confidential, only for members of the consortium (Including the Commission services)

COPYRIGHT

© COPYRIGHT CLERECO Consortium consisting of:

- Politecnico di Torino (Italy) – Short name: POLITO
- National and Kapodistrian University of Athens (Greece) - Short name: UoA
- Centre National de la Recherche Scientifique - Laboratoire d'Informatique, de Robotique et de Microélectronique de Montpellier (France) - Short name: CNRS
- Intel Corporation Iberia S.A. (Spain) - Short name: INTEL
- Thales SA (France) - Short name: THALES
- Yogitech s.p.a. (Italy) - Short name: YOGITECH
- ABB (Norway) - Short name: ABB

CONFIDENTIALITY NOTE

THIS DOCUMENT MAY NOT BE COPIED, REPRODUCED, OR MODIFIED IN WHOLE OR IN PART FOR ANY PURPOSE WITHOUT WRITTEN PERMISSION FROM THE CLERECO CONSORTIUM. IN ADDITION TO SUCH WRITTEN PERMISSION TO COPY, REPRODUCE, OR MODIFY THIS DOCUMENT IN WHOLE OR PART, AN ACKNOWLEDGMENT OF THE AUTHORS OF THE DOCUMENT AND ALL APPLICABLE PORTIONS OF THE COPYRIGHT NOTICE MUST BE CLEARLY REFERENCED

ALL RIGHTS RESERVED.

INDEX

Scope of this document

This document is an outcome of the task T2.2, “definition of operation modes”, elaborated in the description of work (DoW) of CLERECO project under the work package 2 (WP2). This document aims to list typical operational modes of the Embedded Systems (ES) as well as General Purpose (GP) and the High Performance Computing (HPC) systems. The classification of the operational modes is carried out according to the typical application scenarios/environments of the ES and GP/HPC systems, as viewed by various industrial segments associated to these domains. The definition of operational mode includes defining the range of various environmental parameters, e.g., temperature, pressure, vibration, electromagnetic noise, dust, etc. Thus, each operational mode defines a set of specific environmental parameters relevant to a typical application scenario, both for the ES and GP/HPC worlds. It must be noted that this definition/classification of operation modes is valid only within the context of this deliverable (or WP2). So, in other WPs/deliverables, an operation mode may refer to the states in which a system (ES or GP/HPC) is operating, e.g., boot, standby, failure, safety, shutdown, etc.

This document contains confidential information provided by the industrial partners of this project; therefore, the targeted audience for this deliverable is personnel contributing to various WPs of CLERECO project, specifically WP3, WP4 and WP5.

1. Background

Information technology is at the core of our society and it relies completely on the design of electronic information processing systems. Today's computing is a true continuum that ranges from smartphones to mission-critical datacenter machines and from desktops to automobiles. On aggregate, these computing devices represent a total addressable market approaching a billion processors a year, which is expected to explode to more than two billion per year before 2020.

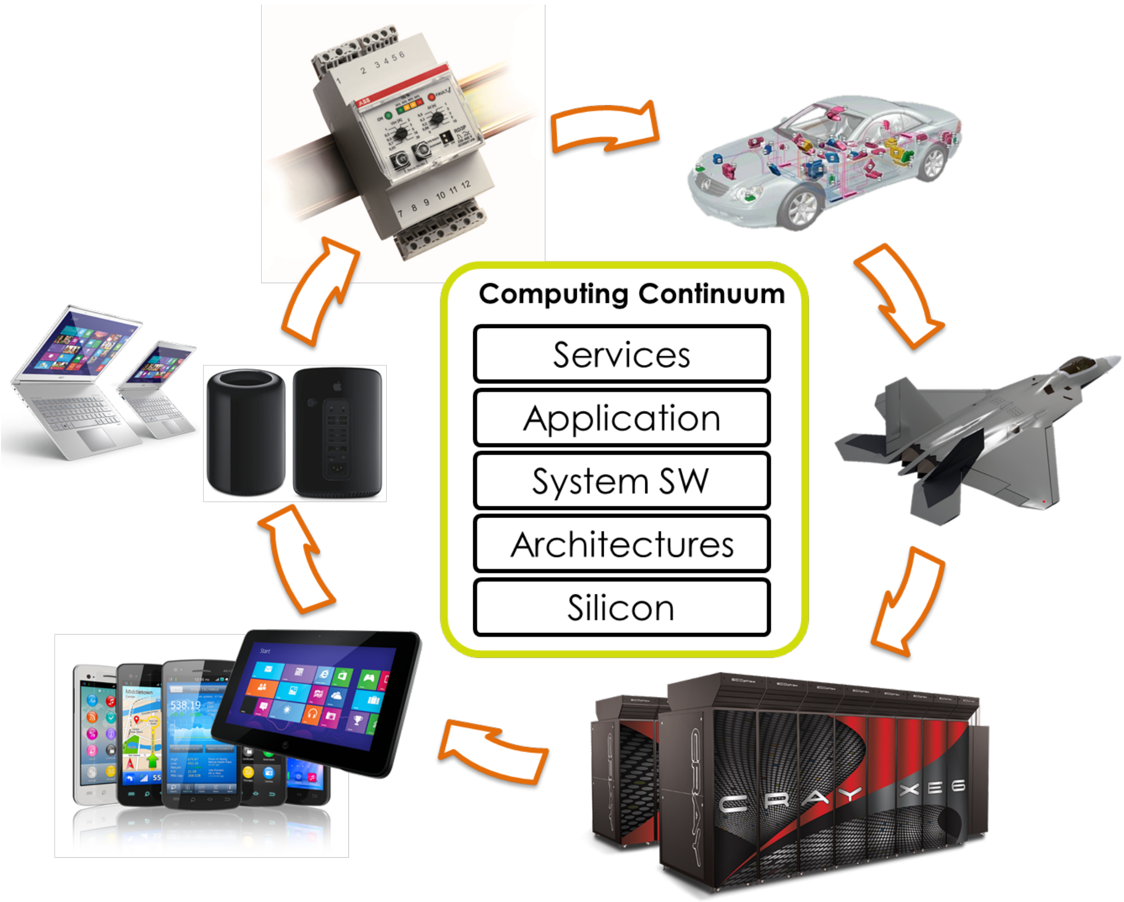


Figure 1: The computing continuum.

Computing systems are operated in a wide range of different environments that may influence their behavior and induce errors in their operation, such as offshore oil platforms with high levels of vibration with humidity and corrosion, offices with air conditioning system, and factory floors with dust, power plants with EMC sources and voltage/current distortion. For example, temperature levels can range from -20° C to +40° C, or the altitude can go from zero to several thousand meters with a much higher cosmic radiation. All these variations in environmental factors may certainly impact the components reliability and lead computing systems to fail.

Systems with high reliability are crucial to fulfill the high requirements being set in today's industry. In several cases, systems for safety critical infrastructures like power plants, energy grids, aircrafts, or satellites, typically require an expected system lifetime of 20 years and longer. Early and unexpected wear out of these components can lead to severe consequences including human safety. End customers have to operate systems subjected to very strict regulations and directives while they are under tough cost pressure.

In the case of high performance computing, the impact of system failures is mainly economic. In today's supercomputers, the failure of a single node may cause an application to crash or even require rebooting the whole system. Therefore, the availability and throughput of the system can be significantly degraded. In particular, Mean-Time Between Failures (MTBF) is important information for system operators, because it impacts their maintenance intervals and consequently their maintenance costs. Keeping MTBF high makes end customers operational cost efficient and competitive, while retaining the high degree of reliability. To keep MTBF high, an increasing amount of resources are devoted on resiliency mechanisms in order to detect, diagnose, recover from, repair/correct, contain, and report errors while system is still running.

The term reliability has briefly been defined in the DoW of CLERECO. Moreover, in the literature, there also exist several slightly varying definitions of the term reliability. This document is based upon the definition from ISO/IEC 2382-14 [18] stating the following for the completeness:

- Reliability: "*The ability of a functional unit to perform a required function under given condition for a given time interval*".
- Functional unit: "*An entity of hardware or software, or both, capable of accomplishing a specified purpose*".
- Maintainability: "*The ability of a functional unit, under given conditions of use, to be retained in, or restored to, a state in which it can perform a required function when maintenance is performed under given conditions and using stated procedures and resources*".
- Availability: "*The ability of a functional unit to be in a state to perform a required function under given conditions at a given instant of time or over a given time interval, assuming that the required external resources are provided*".

It has to be noted that in CLERECO we consider the hardware as being the only contributor as a source of errors. No software errors are considered.

Different engineering techniques are used in traditional reliability engineering, such as Reliability Hazard analysis, Failure mode and effects analysis (FMEA), Fault tree analysis (FTA), etc. Reliability plays a key role in cost-effectiveness of systems with focus on:

- costs of failure caused by system downtime,
- costs of spares, repair equipment, personnel, and
- costs of warranty claims.

In order to achieve the required reliability and availability of the product (or system) a comprehensive requirement specification needs to be developed, which consists of several categories including, but not limited to: engineering, environment, interfaces, safety, performance, customer, and timing.

One of the most important requirement specification categories for the assessment of reliability/availability of any computing system is the environmental factors. The environmental requirements will have strong impact on the design and the manufacturing process of the product. For example a product that will be installed in a control room will have less stringent requirements on vibration sustainability than a product being used on an off-shore oil platform. Normally industrial products need to be certified for various standards before being used in industrial environments. For instance, it would be unthinkable to use a standard office PC as a controller for critical closed loops in a safety system, because it was not designed for such an environment. An unexpected reboot or hanging would be disastrous. Such a device will probably fail in a much shorter time than required.

A very important parameter to be considered (and it can vary for different applications) is the t_{on}/t_{off}, i.e., the time an equipment is expected to be ON and OFF. This time is crucial both for thermal/power considerations but especially for soft-error computations. For example, automotive (in which a reset of the device may happen in average each hour) is quite different from certain industrial applications in which equipment could be ON for days or weeks. Since most HPC as well as industrial systems follow a 24/7 ON time, we will elaborate more on this in Section 2.3.

Finally there is a need to define the term operation(al) mode, which is used differently throughout literature. According to [1], a hardware/software "*may have various operational modes, for example running or standby*". Operational modes in an industrial system could be start-up, up-date, shut down, etc. However, in this project, specifically within the context of this deliverable, the operational mode refers to a set of environmental factors in which a device or system may be operating. For example, a typical operational mode for ES (discussed in more detail in the next section) is a CONTROLLED operational mode, referring to the control room kind of environment, where the operating conditions are carefully maintained. Similarly, based on the application scenarios and location of the installed equipment, one can define various other operational modes, e.g., based on the severity of the operational environment.

For the CLERECO purpose, knowing a precise definition of the operating mode of the system under analysis will help providing more accurate reliability estimation, and eventually, reaching the reliability requirements earlier in the design refinement stage. In other words, leveraging the operating modes information appropriately enables savings in product Time-to-Market and reliability costs given a particular reliability target.

2. Reliability and the Computing Continuum

Out of all possible computing systems that exist nowadays, we can broadly split them into two general categories according to the different modalities they have, namely the General Purpose computing systems, and the Embedded computing systems. In this first subsection 2.1, we focus on the general purpose systems, which in turn can be split into High Performance Computing for more professional uses and Mobile Computing for personal usage. In the second and third subsections 2.2, and 2.3, we present the major concerns related to reliability in Embedded Systems for Industrial, Safety-Critical and Mission-Critical applications.

2.1. Reliability in General Purpose and High Performance Computing Systems

2.1.1. Reliable High Performance Computing

HPC systems represent the higher end of the computing continuum plane comprising a wide range of systems from supercomputers to commercial servers and high-end workstations. Since HPC systems are essentially evaluated in terms of their performance and total cost of ownership; the HPC market evolves quickly and is not conservative with respect to new technologies to allow faster clock speeds and reduced cost per device. Also, because of this rapid technology change, the useful life cycle of HPC systems is much shorter than other computing segments such as the Embedded Systems. Take for example the TOP500 supercomputers list [2], almost every year a new top first supercomputer appears providing the highest performance in the world, and only 8 years after that, the top 500 supercomputers have already overcome that performance, see Figure 2. In fact, despite a supercomputer being able to work reliably for more than 8 years, the HPC market competition is so high that it pushes supercomputer owners to continuously apply upgrades and improvements to their systems. As Figure 3 shows, around 400 supercomputers are replaced every year in the TOP500 supercomputer list.

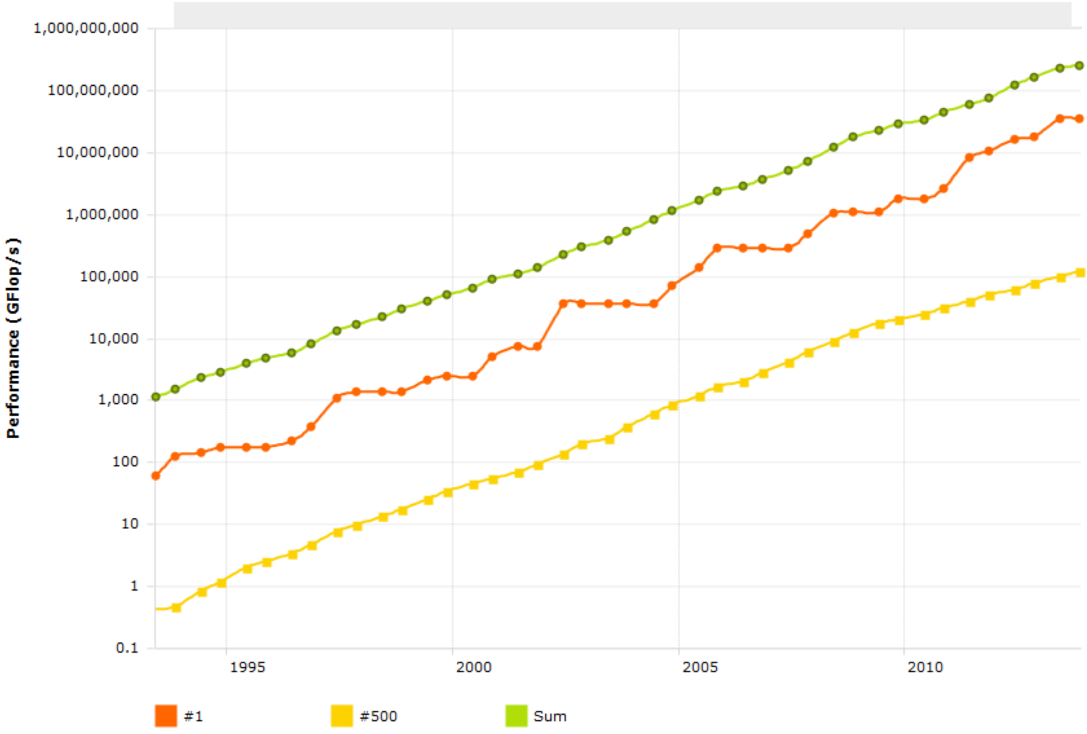


Figure 2: Performance development of HPC systems in the TOP500 (updated every 6 months) [2].

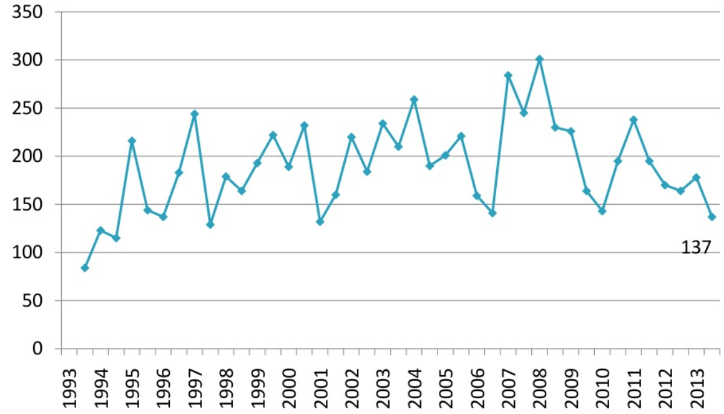


Figure 3: Replacement rate of HPC systems in the TOP500 (updated every 6 months) [2].

In the following, there are some important trends detected by the International Data Corporation (IDC) in the market of HPC servers [3]:

- Global HPC economy is growing (7% growth forecasted over the next five years).
- Major challenges: power, cooling, real estate, system management, and software hurdles.
- The worldwide Petascale Race is in full speed.
- Big data and accelerators are trendy new technologies.

As a result, the main design trends in processors and accelerators for HPC include [4]:

- Hybrid multicore: large and small cores, and accelerators.
- Many-core: larger number of small cores on a chip and specific hardware (graphic operations).

Moreover, the number of nodes in HPC systems is also increasing resulting in an exponential growth in the number of computational cores as shown in Figure 4. The main requirement of HPC systems is essentially throughput, which is related to the amount of work that can be done in a given amount of time with a given cost. This metric is directly determined by performance, but also availability. If a HPC system has very high performance but processes are failing every few minutes, it becomes ineffective. Due to the large number of components in HPC systems and the circuit technology trends towards ever smaller devices (more prone to random errors), even small rates of failures per device can lead to high number of system crashes.

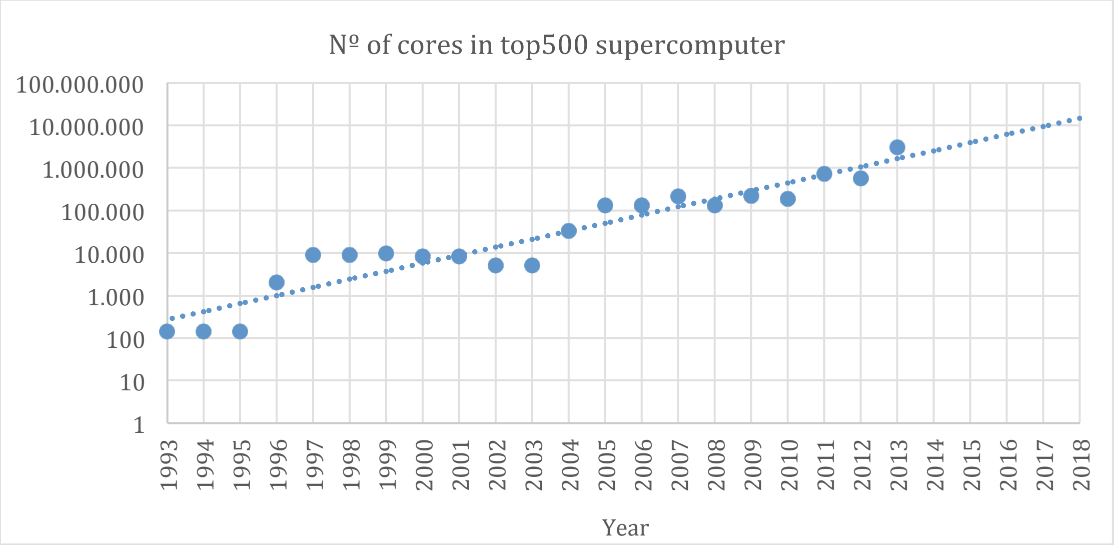


Figure 4: Evolution of number of cores in world's top 1st supercomputer (data from [2]).

Therefore, in the world of HPC, the main purpose of system reliability engineering is not particularly focused on enlarging the system life cycle as much as possible or avoiding catastrophic consequences on the users. Performance and cost prevail over long system lifetime and safety analysis benefits from the fact that there is no direct relation between a system failure and any human risk. The main requirements of HPC systems are essentially reliability and availability in order to provide a cost-effective way to perform highly computational and/or data intensive tasks. Aligned with this, we can find today a lot of research work providing design directions that seek to optimize the Total-Cost-of-Ownership (TCO) of HPC systems [19]. A number of tradeoffs generated by different server configurations, performance variability, MTTF, and ambient temperature are explored in the literature. Some of the most relevant characteristics are summarized in the following list:

1. Servers with different computing performance and power consumption merit exploration to minimize TCO and the environmental impact.
2. Performance variability is desirable if it comes with a drastic cost reduction.
3. Shorter processor MTTF is beneficial if it comes with a moderate processor cost reduction.
4. Increasing by few degrees the ambient datacenter temperature reduces the environmental impact with a minor increase in the TCO.
5. Higher cost for a 3D-stacked processor with shorter MTTF and higher power consumption can be preferred, over a conventional 2D processor, if it offers a moderate performance increase.

Note that failures in a HPC system can be produced by hardware (computation cores, power supply, switches, disks, etc.) and software (operating system, client daemon, file system, etc.) sources that affect one or more computation nodes. Thus, it is becoming increasingly important the development of fault-tolerant techniques to sustain the high reliability and availability requirements of the ever improving HPC segment. With this aim, cross-layer resilient systems take a holistic approach and try to distribute the responsibility for tolerating errors across different layers of the system stack [5]. This approach has the objective of providing the required reliability levels in the most cost-effective way.

Some of the possible fault-tolerant techniques that can be applied at different layers of the system stack are listed below:

- Technology and circuit level: silicon insulator, radiation-hardened cells, clock frequency guard banding, modular redundancy, multiplexing, etc.
- Processor microarchitecture level: error detection mechanisms, error correcting codes, reconfiguration, cycle-by-cycle lock-stepping, redundant multithreading, etc.
- Software level: exception handling, redundant code, etc.
- Whole HPC system level: dual power supplies, multiple gateways to the network or storage, mirrored boot disks, check-pointing, process migration, etc.

Big HPC systems reliability requirements are not usually expressed in terms of maximum failure rates of processors, like in other segments, but are related to the whole system. Moreover, they can be expressed in terms of typical metrics like MTBF, but also in other ways like indicating the amount of increased time execution assuming some dynamic recovery mechanism like check-pointing [6].

2.1.2. Reliable Mobile Computing

The landscape of the computing continuum also includes a wide variety of handheld devices such as cell phones, smart phones, laptops, netbooks and tablets. All these devices already prevail in the computing systems marketplace, thus steering major design and architectural decisions of microprocessor as well as memory vendors. Over the last decade, the importance of mobile computing devices has increased significantly, due to an exponential growth in market size. Figure 5 shows that within only 7 years after their introduction, the number of smartphone users reached more than 1/5 of the global population; while a more recent addition to the mobile computing family, handheld tablets, are following in the same direction. With the strong competition in the mobile computing segment, these numbers are expected to continue this trend in the near future.

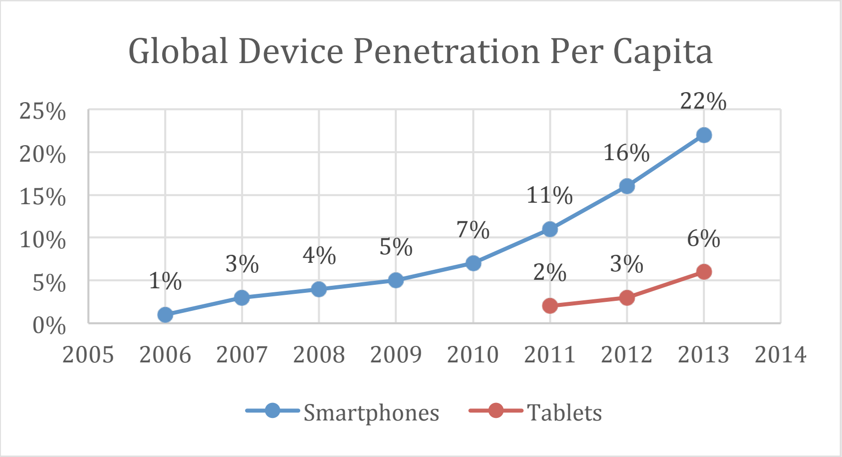


Figure 5: Global penetration of smartphones and tablets per capita [22].

Reliability evaluation and protection provisions for mobile/handheld systems should consider the following factors:

- Mobile/Handheld systems have a low criticality operation, i.e. configure a safety machine or send the position of a maintenance engineer to the control room for immediate evacuation actions when needed. Therefore, reliability provisions should be dedicated to the critical ones.
- Mobile/Handheld systems are battery-operated and thus the extra power/energy for reliability must be carefully adjusted so that it does not severely affects the power/energy behavior of the system. For example, periodic self-tests/diagnostics should not be too frequently activated.
- Mobile/Handheld systems design is size- and weight-sensitive. Therefore, reliability protection mechanisms should have a minimum impact on the system size (both on-chip and off-chip).
- Mobile/Handheld systems have a much shorter lifecycle (in the range 2-3 years) than HPC systems or embedded systems operating in safety-critical application domains.

From the broader reliability point of view, mobile/handheld systems are much less reliability-demanding than the majority of systems in the computing continuum; however, their multi-faceted operation (computing, communication, user interfaces, etc.) requires non-negligible support for reliability, because with this hardware faults will affect:

- The user experience (slow response times, multiple system restarts, lower quality of communication).
- The computing abilities (which although lower than desktop or HPC systems are non-negligible for business laptops, etc.)

2.2. Reliable ES for Industrial Applications

The majority of devices used in the industrial market are based on Embedded Systems like PLC's, IO's, actuators, and sensors. The industrial embedded market is conservative with respect to accepting emerging technologies, which have not been thoroughly investigated for industrial used cases. One of the reasons is due to the long life cycle of a product, often in the range between 20 and 30 years. Beside the necessity of providing reliability, in addition the system must be available – ideally 24 hours every day (often referred to as 24/7). This requires a robust and reliable design. The "Proven in use" argument is much more accepted than theoretical estimations or testing in the industrial sector[3]. Often electronic components have to be "pre"-purchased in high quantities, because their production will be stopped before the industrial product is off the market. In the following, there are some important trends in the industrial embedded evolution:

- Connected everywhere with security, authentication, openness, must handle security and safety in combination.
- Increased hardware complexity (the SoC design paradigm has allowed larger circuits through larger design productivity) using multi-core environments requiring more performance with fan-less devices and adapted ecosystems (debugger, RTOS, compilers, IDE, programming methodologies, code generators, etc.) .
- Cope with increasing software complexity – Smart / Intelligent Devices with multiple features into a single device (M2M, control info between devices, camera, etc.).

[3] *This will change in IEC 61508 3rd edition for which we are currently working. The IEC 61508 will follow the ISO 26262 for which the statistical estimations based on detailed computation is accepted and widely used.*

- Wireless with focus on low cost, low power, short-range operation and availability.
- Improving Lifecycle culture with reuse of legacy code (proven in use).

Embedded building blocks with well-defined layered interfaces will give the application developer a higher abstraction level increasing the potential for reuse and increased productivity (domain specific). A generic layered design architecture for ES is shown in Figure 6.

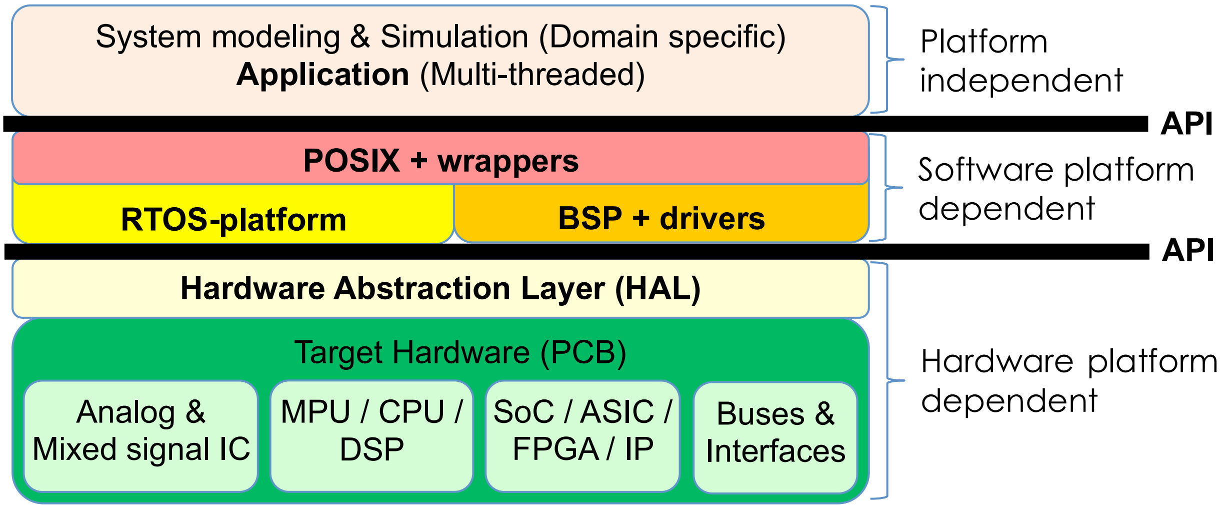


Figure 6: Embedded system architecture – a layered view.

In Figure 6 the final application is what the customers see and what is to be supported in the end. The POSIX standard (OS) with wrapper[4] API makes the Application software platform independent. This is important to consider even when development of own driver is required. Should the driver call the OS directly or should it be done through the API. For performance issues it is important that all API are as thin as possible.

The HAL API hides all hardware complexity from the software, making it hardware independent. If another hardware platform is selected, only HAL is changed. For multicore the HAL can be seen as a Hardware Virtual Machine interface.

In this view, different reliability characteristic is applicable for different layers. The following platform recommendations are applicable for reliable ES:

- Use product-specific standards and (if not available) use general standards giving qualitative (fault behavior and tolerance) and quantitative (PFD, PFH, reliability, availability) requirements.
- Re-use well-proven and robust components and tools.
- Consider different redundant architectures.
- Follow well-proven development processes with required computing support from methods and tools.
- Enable traceability between requirements and all other parts in the chosen development model, so that changes can be made properly and impact analysis can be executed.

Any product developed for industrial applications goes through different stages/phases, defined overall as the product lifecycle. The design, development, and verification phase are driven by market requirements directly (customer needs), from standards (needed to fulfill di-

[4] *Wrapper is an API towards own drivers and real-time adaptations.*

rectives), and from specifications for the environmental factors the product will be operated in. Table 1 shows the different phases and considerations.

Table 1: Product lifecycle overview.

Lifecycle phase	Effect / Considerations
Requirement/Design	Standards, market, environment
Development	Method and tool, processes and procedures, backward compatibility
Verification and Validation	Test and performance, tools, formality, error handling, traceability
Support / Maintenance	Release and version handling, trouble report, change request, monitoring, logging
Termination	Product termination and replacement

In [7] it is stated that products placed on the European market have to be CE marked. The CE mark and the document Declaration of Conformity (DOC) verify that the product complies with applicable EU Directives. If there is a requirement for specific installation measures an installation guide is issued for the product. Directives typically applicable to industrial products are the EMC-directive [16] and the Low Voltage Directive (LVD) [17].

The protection and control equipment supplied by an industrial vendor, shows in the technical documentation compliance with standards and directives. The products are routinely tested to validate the conformity to such standards and directives. Next is a list of the main standards typically applicable to protection and control equipment.

Table 2: Main standards for protection and control equipment.

Classification	Standard	Description
General	IEC 60255	Electrical relays
General	IEC 60038	IEC Standard voltages
General	IEC 60068	Environmental testing
General	IEC 60664	Insulation co-ordination for equipment within low voltage systems
General	ANSI C37.90-1995	IEEE Standard for Relays and Relay Systems Associated with Electric Power Apparatus
CE requirements	EN 50081-2	Electromagnetic compatibility - Generic emission standard - Industrial environment
CE requirements	EN 50082-2	Electromagnetic compatibility - Generic immunity standard - Industrial environment
CE requirements	EN 50178	Electronic equipment for use in power installations
CE requirements	EN 60255-6	Electrical relays - Measuring relays and protection equipment – Safety

2.3. Reliable ES for Safety-Critical and Mission-Critical Applications

Embedded systems in the field of aerospace, space, and security applications represent another extreme case of the computing continuum if we consider the fact that they combine high reliability requirements and harsh environmental constraints. Indeed, these systems have stringent dependability requirements as they are generally mission-critical or safety-critical (i.e. their failure can lead to the failure of the mission or to the injury or death of persons or environment). The decreasing reliability of hardware platforms is thus of the utmost importance for the design of these systems. Continuing to design dependable embedded systems on top of less reliable semiconductor technologies will involve addressing some specific aspects:

- **Long lifetimes:** Critical embedded systems are designed for long service life ranging from around 15 years in the case of a satellite (without possibility to replace any defective component) to 25 years in the case of an avionic system. Due to accelerated aging of the devices in advanced technology nodes, early wear-out effects could occur during the service life of the system in the future.
- **Operational reliability:** Embedded systems used in mission-critical or safety-critical applications have highly reliability requirements (with failure rates as low as 10^{-9} FIT). Furthermore, they often operate in harsh environments (in term of radiations, temperature ranges, temperature cycles, mechanical constraints…). In this context, achieving the required level of reliability becomes a time- and resources consuming task because of the greater sensitivity of devices to Single-Event Effects and the increasing device variability.
- **Predictability:** Most critical embedded systems are hard real-time systems, where the violation of a timing constraint can lead to the failure of the whole system. So, timing is as important as the functional correctness (while power consumption for reliability provisions is in general more flexible apart from battery-operated systems). Providing guarantees on program execution times is thus mandatory. However, determining the Worst-Case Execution Time (WCET) is now extremely difficult with the introduction of aggressive architectural techniques and the emergence of multi-core processors. Moreover, the impact of faults on the timing of the application and on the temporal isolation between tasks is a concern for critical applications.
- **Certification and qualification:** Safety is the primary concern in avionics and transportation systems. Design of such systems has to comply with multiple safety standards (e.g. DO-178B for SW development, DO-254 for hardware design in case of avionics, EN 50128/9 and IEC 61508 in case of transportation) and complex certification procedures. The qualification of the components or of the technologies used for these systems is essential to ensure that the system meets the expected criteria. The relentless technology scaling will dictate a revision of the qualification procedures and of the failure models as we move into the nanoscale regime.

As a result of all these constraints, the domain of critical embedded systems is globally more conservative than the mainstream consumer electronic domain. Although critical embedded systems have in common high reliability requirements, significant differences can be observed from one application domain to another. Different application requirements and environmental constraints result in different selection of technologies and components.

For instance, computing systems are embedded in satellites for the satellite management (altitude and orbit control, thermal control, etc.) as well as for the local payload processing (beamforming, radar, GPS …). Figure 7 shows an example of processing unit embedded in the Poseidon-2 satellite.

Figure 7: Example of processing unit (on-board the Poseidon-2 satellite).

In the space domain, the need for higher communication satellite capacity as well as high-end on-board radar processing pushes towards higher on-board processing power. Given the huge costs of a satellite launch, this evolution is performed within severe constraints of mass and volume. The power is also severely constrained as the solar panels and the embedded batteries of the satellite are the only available sources of energy. Providing the expected computing power, while meeting the tight power budgets, requires specialized processing architectures. So, ASIC and FPGA components are extensively used in on-board processing equipment. This equipment is designed to support the extreme conditions of the space environment. Indeed, on-board satellite equipment needs to operate autonomously with high levels of availability. Components are therefore hardened against radiation effects. Radiation-hardening is generally achieved through specific semiconductor processes (Radiation-Hardening-By-Process) or specific standard-cells library (Radiation-Hardening-By-Design). Due to these specific constraints, the space domain is generally conservative with respect to new technologies. Architecture-level or system-level mitigation techniques have also been investigated. Such alternatives could then enable the leverage of commercial ASIC technologies and libraries.

In the avionics domain, COTS (Commercial Off-The-Shelf) processors have been favoured, and using them is now a commonplace. These high-performance processors enable to integrate an increasing number of avionics functions (see Figure 8) in a reduced number of computing units.

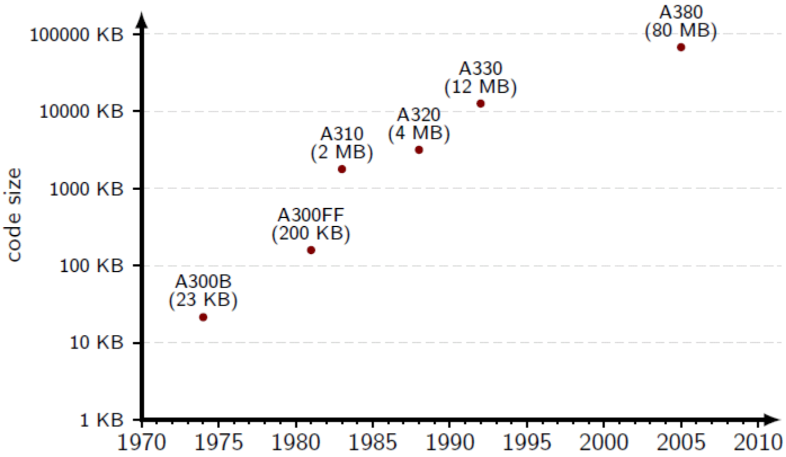


Figure 8: Code size for Airbus aircrafts [20].

Such integration is driven by the need to reduce the size and weight of on-board equipment and cabling (and so the CO2 emissions), and the dissipated power (and the burden of the cooling infrastructure). Using generic processors is also an effective way to reduce the number of computing unit types and so to improve the maintainability and serviceability of the equipment. The current trend of avionics systems is to execute multiple avionics applications on top of a common Integrated Modular Avionics (IMA) platform. A strict time and space partitioning between these applications is ensured by the platform to guarantee the safety of the system. In the presence of faults, maintaining time and space partitioning is essential and fault isolation has to be guaranteed between avionic partitions. Figure 9 shows for example the structure of the PikeOS operating system with its different partitions.

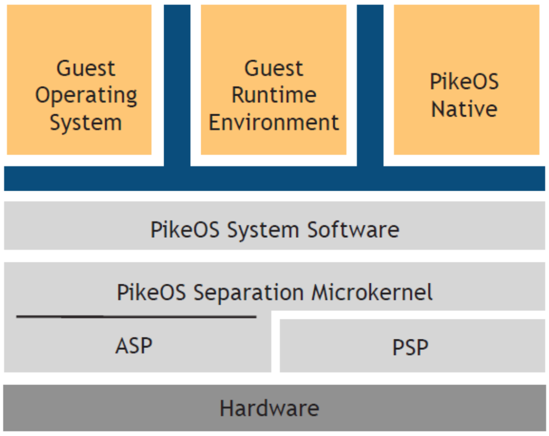


Figure 9: Example of time and space partitioning in PikeOS operating system [8].

To pursue this trend in the future, higher performance processors are required. But, processors implemented in future advanced technology nodes could be used only if they satisfy the high and strict reliability requirements of critical embedded systems. During the last decades, the evolution of avionics systems has indeed contributed to increase the aviation safety. For instance, Figure 10 shows the global rate of accidents involving passenger fatalities per 10 million flights, for scheduled commercial air transport operations.

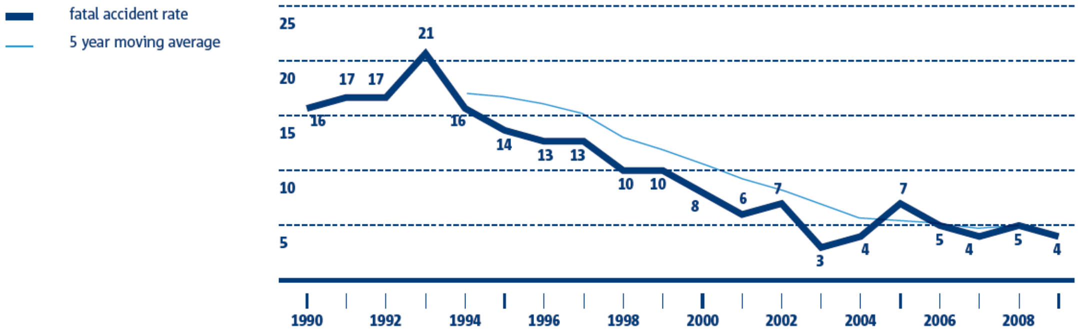


Figure 10: Global rate of accidents involving passenger fatalities per 10 million flights [21].

2.4. Summary table

The following table summarizes the main concerns of each computing segment with respect to reliability engineering and classifies their impact on different target metrics of a computing system design. Using a three level classification, we provide a quick overview of what are the major concerns per segment. This classification facilitates refining the definition of operating modes by making sure that all the factors that influence areas with higher impact from the reliability perspective are taken into account. For the purpose of CLERECO, it is important to have general guidelines for distributing design effort and resources to high impact areas.

Table 3: Summary table of major reliability concerns per computing segment.

	Functionality	Performance	Power / Energy	Safety	Total-Cost-of-Ownership (TCO)	Reliability Design Effort	Availability
HPC	Medium	High	Medium	Low	High	Medium	High
Mobile	Medium	Medium	High	Low	High	Low	Medium
Industry ES	Medium	Medium	Low	Medium	Medium	High	High
Safety critical ES	High	Medium	Low	High	Medium	High	High

3. Common Operational modes of GP and Embedded Computing Systems

As mentioned earlier, the operational mode of a system is defined with respect to the external factors in a given environment and for any application scenario. In this chapter we will define typical operational modes for GP and ES. Later we will map one or more typical application scenarios for each sector (HPC, Mobile, Industrial, and Mission Critical) into each of the operation modes. The Operational Modes (OM) can be classified according to the following three categories (or levels):

- **OM1: Controlled** – is a stable and comfortable environment, even for human beings, typically office environment. It normally includes cooling systems that control temperature and humidity. Other environmental factors, such as vibration and EMI, are also limited by the design of working space. It also includes typical outdoor environments for handheld devices where humans feel comfortable, which implicitly limits the ranges of the environmental factors.
- **OM2: Uncontrolled** – is much more exposed to environmental conditions and sudden changes. This is the case of factory and substation sites where there is no mechanism to prevent from high changes in temperature, humidity, vibration, etc.
- **OM3: Harsh** – is where typically sensors and actuators are placed with rough environmental conditions. Airplanes, helicopters, oil rigs and other harsh environments are subject to extreme environmental conditions that need particular treatment from the design perspective. Parameters defining operational modes of a system are mainly environmental factors influencing the operation of HPC/ES systems. It must be noted that there also exist special/extreme condition operational modes with extreme characteristics, such as nuclear, marine / oil platforms, trains, airplane / spacecraft, etc., where environmental requirements are much more strict and are often stated in standards. These kinds of modes are typical for safety and mission critical ES applications.

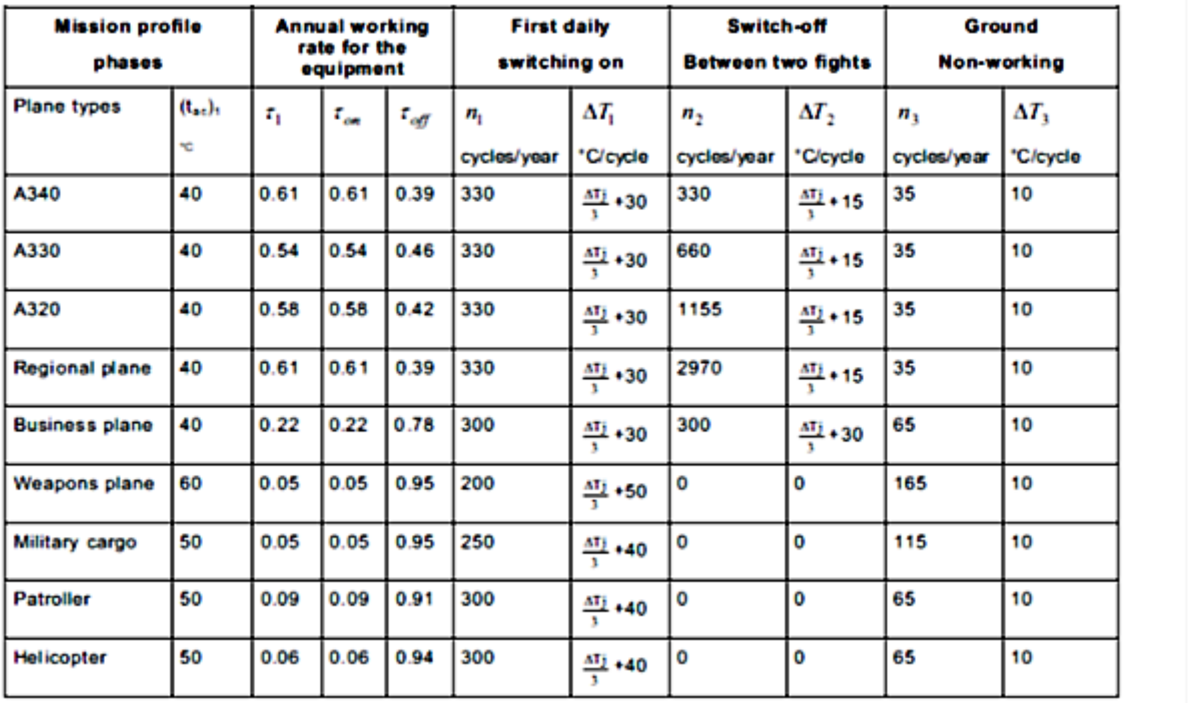
Within the context of this report, the environmental factors are the external factors that define the operational modes of a computing system. The table next shows the different environmental factors and their effect/considerations on reliability.

Table 4: Environmental factors.

Parameter	Effect / Considerations	Standards
Temperature cycling	Low-power, fan less, load balancing, etc. Office environment is typically in the range +10 to +43 °C, where in-field devices might range from -40 to +85 °C.	IEC 60068-1 (General) IEC 60068-2-1 (Environmental testing) IEC 60068-3-1 (Background information) IEC 62380 (see note below the table)
Humidity / dust / acid / salt	Sealed, fast degradation, heat sensitive, external cool fins, often solved by specially designed cabinets. Typically IP 65 for field devices, where 6 (dust-tight) is protection against foreign objects and 5 (low pressure water jets protected) is protection against water.	IEC 60068-2-2 IEC 60068-2-30 IEC 60068-2-52
Vibration / shock / pressure / gravity	Glued components, shock absorber, SW supervision (components, cables, etc.), often solved by specially designed cabinets. Different standards is applicable (IEC 60068-2, MIL-STD-810G/883, ISO 16750) considering velocity, acceleration and displacement with max test and different waveforms.	IEC 60255-21 (industrial applications; for measuring relays and protection equipment) IEC 61373 (traction application) IEC 60068-2-6 IEC 60068-2-27
Explosion (EXN)	Shall not set of an explosion when surrounded by specified flammable gases or dust. Special considerations with respect to contacts and short circuits (no sparks).	IEC/EN 60079
EMC / EMI	Attention on communication error, cabinet shielding required. Electromagnetic Compatibility, EMC, Directive 2004/108/EC (updated July 2007). EMI stands for interference.	EN 55011 (Gr.1 / Class B) EN 50140 EN 50141 EN 50082-2
Radiation	This will vary depending on where the equipment is used like power plant, aircraft or space, and the effect is closely related to the process technology used. Reduced by ECC protection with scrubbing / restart continuous diagnostic, radiation hardened hardware, redundancy.	IEC 62396 JEDEC (e.g. JESD-89 Test Method for Alpha Source Soft Error Rate)
Electrical stability / Insulation	Sudden power loss, voltage peaks, dielectric, etc. Need for UPS with surge protection.	EN 61000-4 IEC 60947-1 (8.3.3.4.1, item 3) IEC 255-4

Altitude	Standard IEC 60068-2-13 only defines how to test, but it is no derating effect given for temperature range reduction.	IEC 60068-2-13

The IEC 62380 is also a useful standard defining “mission profiles”: for example Figure 11 describes the mission profiles for military and civil avionics.



Mission profile phases		Annual working rate for the equipment			First daily switching on		Switch-off Between two fights		Ground Non-working	
Plane types	$(t_{ac})_i$ °C	τ_1	τ_{on}	τ_{off}	n_1 cycles/year	ΔT_1 °C/cycle	n_2 cycles/year	ΔT_2 °C/cycle	n_3 cycles/year	ΔT_3 °C/cycle
A340	40	0.61	0.61	0.39	330	$\frac{\Delta T_j}{3}+30$	330	$\frac{\Delta T_j}{3}+15$	35	10
A330	40	0.54	0.54	0.46	330	$\frac{\Delta T_j}{3}+30$	660	$\frac{\Delta T_j}{3}+15$	35	10
A320	40	0.58	0.58	0.42	330	$\frac{\Delta T_j}{3}+30$	1155	$\frac{\Delta T_j}{3}+15$	35	10
Regional plane	40	0.61	0.61	0.39	330	$\frac{\Delta T_j}{3}+30$	2970	$\frac{\Delta T_j}{3}+15$	35	10
Business plane	40	0.22	0.22	0.78	300	$\frac{\Delta T_j}{3}+30$	300	$\frac{\Delta T_j}{3}+30$	65	10
Weapons plane	60	0.05	0.05	0.95	200	$\frac{\Delta T_j}{3}+50$	0	0	165	10
Military cargo	50	0.05	0.05	0.95	250	$\frac{\Delta T_j}{3}+40$	0	0	115	10
Patroller	50	0.09	0.09	0.91	300	$\frac{\Delta T_j}{3}+40$	0	0	65	10
Helicopter	50	0.06	0.06	0.94	300	$\frac{\Delta T_j}{3}+40$	0	0	65	10

Figure 11: Mission profiles for military and civil avionics.

Depending on the application scenario, a system can operate in one or many of the above mentioned environments. For example a plane has to fulfill very different requirements for each of the flight phases (rolling to the start runway, take off, flying, landing, rolling again), where each phase is represented by specific environmental parameters affecting the modes of the plane. In the industry we usually distinguish between storage and operation due to different environmental conditions.

Table 5 below shows the summary of environmental factors listed in the description of different operation modes and shows their impact on various error types that will be taken into consideration in CLERECO. The exact impact and the degree of the impact of each factor depends on the particular technology used in manufacturing the components that are subjected to the particular operation mode as well as the upper and lower limits on environmental factors set by the operation mode. However, it can still be seen that all these environmental factors have a direct impact on one or more error types, which makes it crucial to quantify the range of values these parameters can take under different OMs and usages for an accurate reliability evaluation. An in-depth description of failure mechanisms and the impact of OM environmental factors on them will be presented in deliverable “**D2.2.1 – Characterization of failure mechanisms for future systems**” of the CLERECO project.

Table 5: Impact of OM environmental factors on different error types.

Parameter	**Impact on Transient Errors**	**Impact on Intermittent Errors**	**Impact on Permanent Errors**
Temperature / Temperature cycling	Increase in transient failures with higher temperatures due to higher energetic particles and increased leakage (e.g. soft errors in DRAMs)	Increase in intermittent failures due to device degradation/healing effects (e.g. NBTI effects) and thermal stress (e.g. contact short/open due to thermal expansion/contractions)	Increase in permanent failures due to device degradation effects. (e.g. electromigration effects) and thermal stress (e.g. wear out effects)
Humidity / dust / acid / salt	-	-	Increase in permanent failures due to corrosion/shorting on contacts
Vibration / shock / pressure / gravity / explosion	-	May cause intermittent failures depending on the strength of the effect	Increase in permanent failures due to mechanical stress and contact/solder breaks
EMC / EMI / Radiation / Altitude	Increased soft errors with to increased interference (e.g. IR effects, magnetic storage technologies)	May cause intermittent failures for unshielded components that last throughout the exposure period (e.g. solar EMP)	Oxide failure or metal melt due to ESD; power surges due to HEMP and HPM; damage on circuits due to an electrostatic discharge Increase in permanent failures due to device degradation effects (Total Ionizing Dose) and destructive effects (Single-Event Latch-Up)

4. Application Scenarios per Sector

4.1. Application scenarios in GP and HPC Systems

Apart from very few exceptions, all the GP/HPC systems work in controlled environments, and thus, under the CONTROLLED operating mode (OM1). The usual working place is a control room with a stable environment and minor fluctuations. Still, non-controlled human-friendly outdoor environments are also considered here for mobile applications. Despite having stable operating conditions compared to other computing applications, we can still distinguish four different groups within OM1 GP/HPC systems. This sub-classification, which is based on the level of control exercised over the environmental conditions, is in concordance with the convention used by the ASHRAE in its work "Thermal Guidelines for Data Processing Environments" [9]:

- **Group 1:** Data centers with tight control environmental parameters (dew point, temperature, and relative humidity) and mission critical operations. E.g., enterprise servers and storage products (generally with raised floor).

- **Group 2:** Information technology space or lab environment with some control of environmental parameters (dew point, temperature, and relative humidity). E.g., volume servers, storage products, personal computers, and workstations.
- **Group 3:** General purpose computing systems in typical office environments with minimal control on temperature. E.g., personal computers, and workstations.
- **Group 4:** Mobile systems that operate in non-controlled environments (e.g. outdoors) but implicitly controlled conditions because manufacturers do not guarantee proper functioning beyond certain environmental ranges. E.g., laptops, tablets, and smartphones.

Table 6 summarizes the specific environmental conditions present in each CONTROLLED application scenario.

Table 6: Environmental conditions for OM1 (CONTROLLED) [9].

	Group 1 Tight control	**Group 2 Some control**	**Group 3 Minimal control**	**Group 4 Implicit control**
Environmental factor	**Range**	**Range**	**Range**	**Range**
Temperature	15—32 °C	10—35 °C	5—35 °C	0—35 °C
Rate of temperature change	20 °C/h	20 °C/h	-	-
Relative Humidity	20—80 %	20—80 %	8—80 %	0—90 %
Dew point	< 17 °C	< 21 °C	< 28 °C	-
Altitude	0—3050 m	0—3050 m	0—3050 m	0—3050 m
Sand and dust	-	-	-	-
Vibration	-	-	-	-
Mechanical shock	-	-	-	-
Electromagnetic sources	-	-	-	-

In the following, we review some of the major characteristics of a control room for high performance computation. Thanks to the fast technology scaling trends, computing systems have experienced a huge improvement in its capabilities (performance and computation power) per square millimeter. However, tied to the same trend comes an important increase in the power density that must be dissipated in order to keep the system operating reliably. Thus, the cooling system of HPC systems has become a major challenge nowadays. Some of the considerations that are being analyzed for the problem of temperature dynamics in data centers are listed below:

1. Cooling systems with different airflows directions.
2. Airflow delivery from overhead or from under a raised floor.
3. Ceiling heights to eliminate "heat traps" or hot air stratification.
4. Height of raised-floors to avid hot spots.
5. Distribution of computer equipment in the room.
6. Air distribution configurations with respect to cooling effectiveness.
7. Redistribution of computing workload among the HPC nodes.

Figure 12 depicts a basic floor plan of a data center based on the combination of cooled and uncooled aisles [10]. Evidently, the final configuration of the data center highly depends on the type of use and the size of the system. In the next sub-chapters we provide some examples to better understand the environmental conditions associated to each group of CONTROLLED application scenarios.

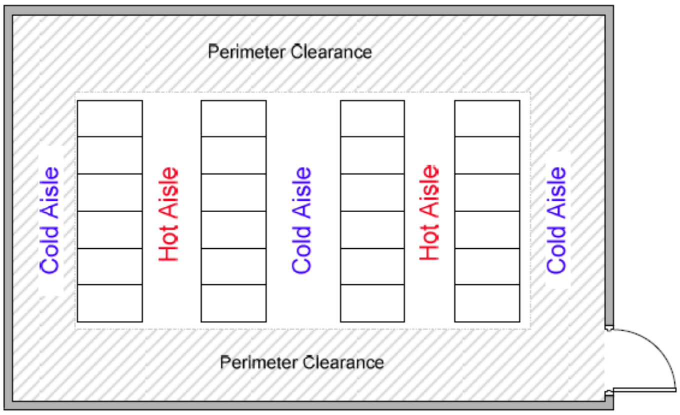


Figure 12: Basic hot-aisle/cold-aisle data center equipment layout.

HPC systems are required for computing intensive tasks and the level of success in the results that can be achieved is highly dependent on the available computing power. The main applications normally assigned to high performance computing systems can be classified according to the following list [3]:

- **Government labs:** Basic and applied research.
- **University/academic**: Basic and applied research.
- **Bio-sciences and the human genome**: Drug discovery, disease detection/prevention.
- **Computer aided engineering (CAE)**: Automotive design and testing, transportation, structural, mechanical design.
- **Defense and energy**: Nuclear stewardship, basic and applied research.
- **Electronic design and automation (EDA)**: Electronic component design and verification.
- **Geosciences and geo-engineering**: Oil and gas exploration and reservoir modeling.
- **Digital content creation (DCC) and distribution**: Computer aided graphics in film and media.
- **Weather forecasting**: Near term and climate/earth modeling.
- **Economics/financial**: Wall Street risk analysis, portfolio management, automated trading.
- **Chemical engineering**: Process and molecular design.
- **Mechanical design and drafting**: 2D and 3D design and verification, mechanical modeling.

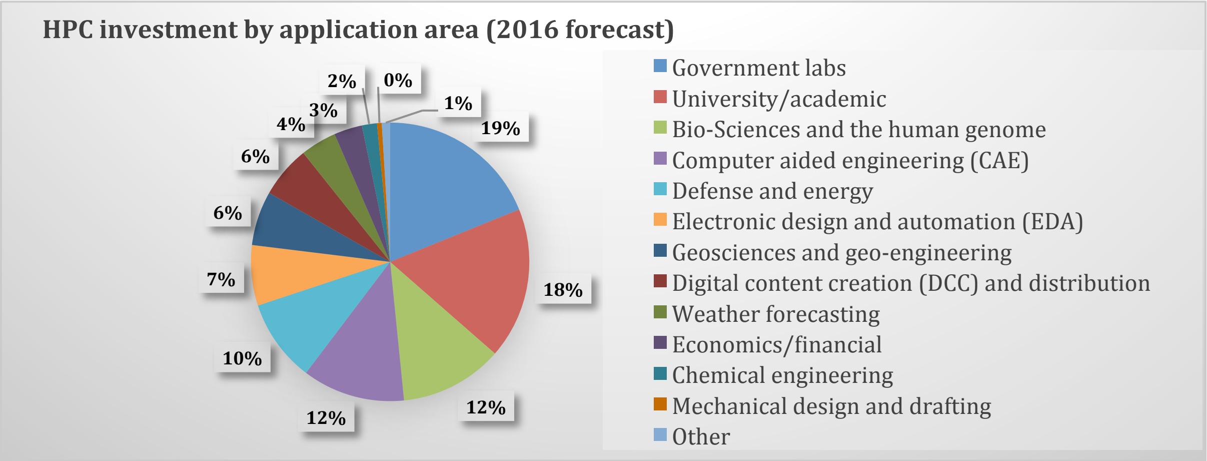


Figure 13: Forecasted investment in HPC systems by application area in 2016 (data from [3]).

There is a widespread range of application fields in the world of HPC systems as observed in Figure 13. In order to clearly identify different application scenarios and the associated uses, we review different applications of the Intel microprocessor Xeon E5.

4.1.1. Tight Environmental Control

This first group of controlled environments corresponds to data centers or supercomputers with tightly controlled environments using ultimate cooling system designs that take precise control over the dew point, temperature, and relative humidity. The associated floor plan is carefully analyzed and optimized considering sizes, heights of racks and ceiling. This kind of data center generally uses raised floor to enhance the cooling efficiency. The main uses for this group are the university or academic research, bio-sciences, defense and energy, geosciences, weather forecasting, and chemical engineering. A good example for this group is the current top supercomputer in the world called Tianhe-2 that is based on the Intel Xeon E5 microprocessor.

Figure 14: Application example: The Tianhe-2 supercomputer [2].

Tianhe-2 is the current top supercomputer in the world. It was developed by China’s National University of Defense Technology and has a performance of 33.86 petaflop/s on the Linpack benchmark. It consists of 32,000 Intel Ivy Bridge Xeon E5-2692 12C sockets and 48,000 Xeon Phi 31S1P boards. Other relevant characteristics of Tianhe-2 are summarized in Table 7.

Table 7: Tianhe-2 characteristics

Theoretical peak performance: 54.9 Pflop/s Peak performance (HPL benchmark): 30.65 Pflop/s (62.3% efficiency) Configuration: 125 racks x 4 frames x 16 boards x 2 nodes x (2 Intel Ivy Bridge sockets [12 cores] + 3 Intel Xeon Phi 31S1P boards [57 cores]) Total of 3,120,000 cores Operating System: Kylin Linux, Speed 2.200 GHz	Memory per Node (GB) 64, Total Memory: 1PB Global Shared parallel storage system: 12.4 PB The footprint is 720 square meters Cabinets: 125compute+13communication+24storage(total162) The peak power consumption of processors, memory, and interconnect network under load for the system is at 17.8 MWs The peak power consumption adding the cooling system is 24 MWs

This world's fastest supercomputer is used to solve high complexity problems, perform difficult physical simulations, analyze big amounts of data, and run government security applications. It uses a cooling system based on a closed-coupled chilled water-cooling system with a customized liquid water-cooling unit.

4.1.2. Some environmental control

This group corresponds to data centers with controlled environments using cooling system. The associated floor plan is not carefully studied and optimized. It normally uses standard designs that minimize the cost of implementation. This group mostly corresponds to big companies and government/institutions and the main uses are government labs, computer aided engineering (CAE), electronic design and automation (EDA), digital content creation (DCC) and distribution, economics/financial, and mechanical design and drafting. A good example for this group is the Cloud Blade server NX5440 that is based on the Intel Xeon E5 microprocessor.

Figure 15: Application example: Cloud Blade server NX5440 [11].

Cloud Blade server NX5440 is a high-performance rack server system developed by Inspur. It consists of 20 blades with 2 Xeon E5 processors operating at 2.13GHz, 8 DIMM slots of up to 256GB. It requires specific environment conditions to operate: temperature between 10°C and 35°C and relative humidity between 35% and 80%.

4.1.3. Minimal environmental control

This third group corresponds to personal computers, workstations and laptops data centers with minimal control over the temperature, generally based on a single fan. There are many different possible uses for this group but if we limit them to the ones typical of HPC systems we have computer aided engineering (CAE), electronic design and automation (EDA), digital content creation (DCC) and distribution, economics/financial, and mechanical design and drafting. A good example for this group is the Apple Mac Pro computer that is based on the Intel Xeon E5 microprocessor.

Figure 16: Application example: Apple Mac Pro [12].

Mac Pro is a general-purpose desktop computer with the Intel Xeon E5 processor. It has 12 cores and four-channel memory controller providing 60GB/s of memory bandwidth. Other relevant characteristics of Mac Pro are summarized in Table 8.

Table 8: Mac Pro characteristics

Speed 2.7GHz	Storage (flash): 1TB
Memory: 64GB	Dual AMD FirePro D700 graphics processor

It requires specific environment conditions to operate: temperature between 10°C and 35°C and relative humidity between 5% and 95%.

4.1.4. Implicit environmental control

The last OM1 group corresponds to Mobile Computing systems where the control over the environmental factors is implicit to the manufacturer design specifications and the practical usage conditions. In this case, most of the applications focus on personal usage such as gaming, social networking, and web browsing. For example, according to Flurry Analytics the average time spent on iOS and Android connected devices is split in the following way [23]: 32% gaming, 28% social networking, 18% productivity and other utility applications, 14% web browsing, and 8% video entertainment. Typical examples for Mobile computing system are the two bestselling mobiles of 2013: the iPhone 5s and the Samsung Galaxy S4.

Figure 17: Application example: iPhone 5s and the Samsung Galaxy S4 [24].

In this group the user has minimal control over the environmental conditions. Therefore, the device should be operational under all normal conditions. The operating requirements of a typical smart phone or tablet are summarized in Table 9.

Table 9: Operating limits of mobile products from Apple (iPhone, iPad, and MacBook Pro) [12].

Operating ambient temperature:	**0° to 35° C**
Storage temperature:	–25° to 45° C
Relative humidity:	0% to 95% noncondensing
Operating altitude:	tested up to 3,000 meter
Maximum storage altitude:	4,500 meter
Maximum shipping altitude:	10,500 meter

4.2. Application scenarios in the Industrial sector

Embedded systems with hardware and software are executing functionality according to the "mode of the device". For example, during startup or configuration the system is not operational in the sense that it cannot perform full run-time functionality. For simplification, within the scope of this report, the following clarifications are made for definition of Operation mode:

- Devices are deployed and fully functional (in run mode)
- Reliability parameters will be specific for the application area they are defined

In order to define parameters, which characterize an operational mode of a system, we need to first identify a relevant system with typical application scenarios. In the regime of embedded systems, there exist a multitude of systems operating in various environments. An example from the ABB product portfolio covering most of the aspects of embedded system is an AC 800PEC controller, which has a wide range of applications in commercial and industrial environments.

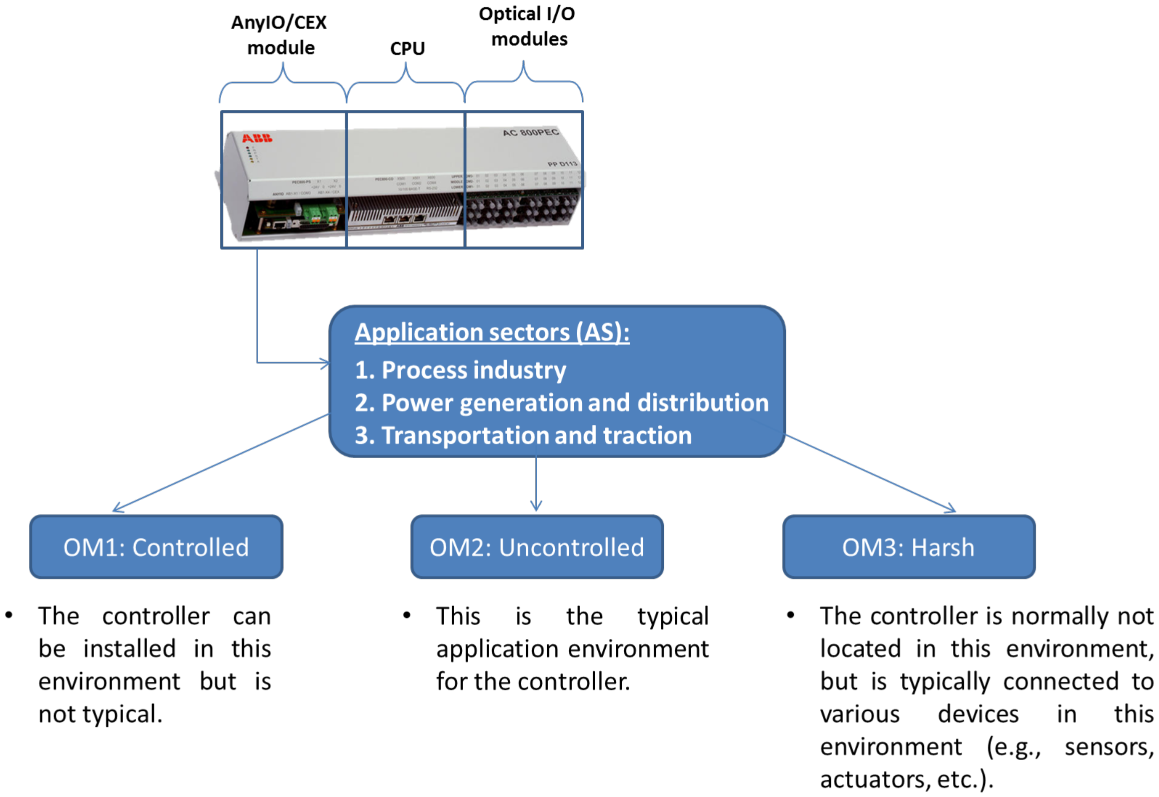
ABB
AC 800PEC
PP D113
PEC800-PS X1 X2
+24V 0 +24V 0
ANYIO AB1-X1 / COM3 AB1-X4 / CEX
PEC800-CO X500 X501 X600
COM1 COM2 COM4
10/100 BASE-T RS-232
UPPER OM3- 01 02 03 04 05 06 07 08 09 10 11 12
MIDDLE OM2- 01 02 03 04 05 06 07 08 09 10 11 12
LOWER OM1- 01 02 03 04 05 06 07 08 09 10 11 12

Figure 18: AC 800PEC Controller from ABB.

The AC 800PEC is a high performance process control system and belongs to the Control IT product line from ABB. It is the optimum solution for combining high-speed control requirements of power electronics applications with low-speed process control tasks usually carried out by separate PLC units. The AC 800PEC controller incorporates equipment that meets the most challenging – and also contradictory – requirements in process control. It includes a wide range of I/O modules to cover all power electronics control requirements. The different I/O modules can be connected with the AC 800PEC controller to cover most automation requirements. The following application scenarios (AS) are covered by the AC 800PEC Controller:

- AS-1: Process industry
- AS-2: Power generation and distribution
- AS-3: Transportation and traction

The above application scenarios have different operational criteria depending on environmental view. The application scenarios for AC 800PEC are illustrated in the next figure.



Application sectors (AS):

1. Process industry
2. Power generation and distribution
3. Transportation and traction

OM1: Controlled

- The controller can be installed in this environment but is not typical.

OM2: Uncontrolled

- This is the typical application environment for the controller.

OM3: Harsh

- The controller is normally not located in this environment, but is typically connected to various devices in this environment (e.g., sensors, actuators, etc.).

Figure 19: Application Sectors with environments for the AC 800PEC controller.

The classification of the application environments (or operational modes) shown in Figure 19 is based on the operational limitations typically common for one particular environment in all application sectors. As shown the controller can be installed in a Control Room or Factory floor / Substation environment and in special cases also out in the field. It is the worst operational conditions that will define the design criteria. This is always a balance between development cost and market price, where market price is the critical part. For AC 800PEC most of the requirements come from OM2. The in-field environments (OM3) are the core sources of information flow and may represent the power generation site for AS-2 with sensors and actuators, and the traction and engine system deployed on the transportation services AS-3, such as in ships and trains. In this last example the control room is directly connected to the in-field environment as shown in Figure 20.

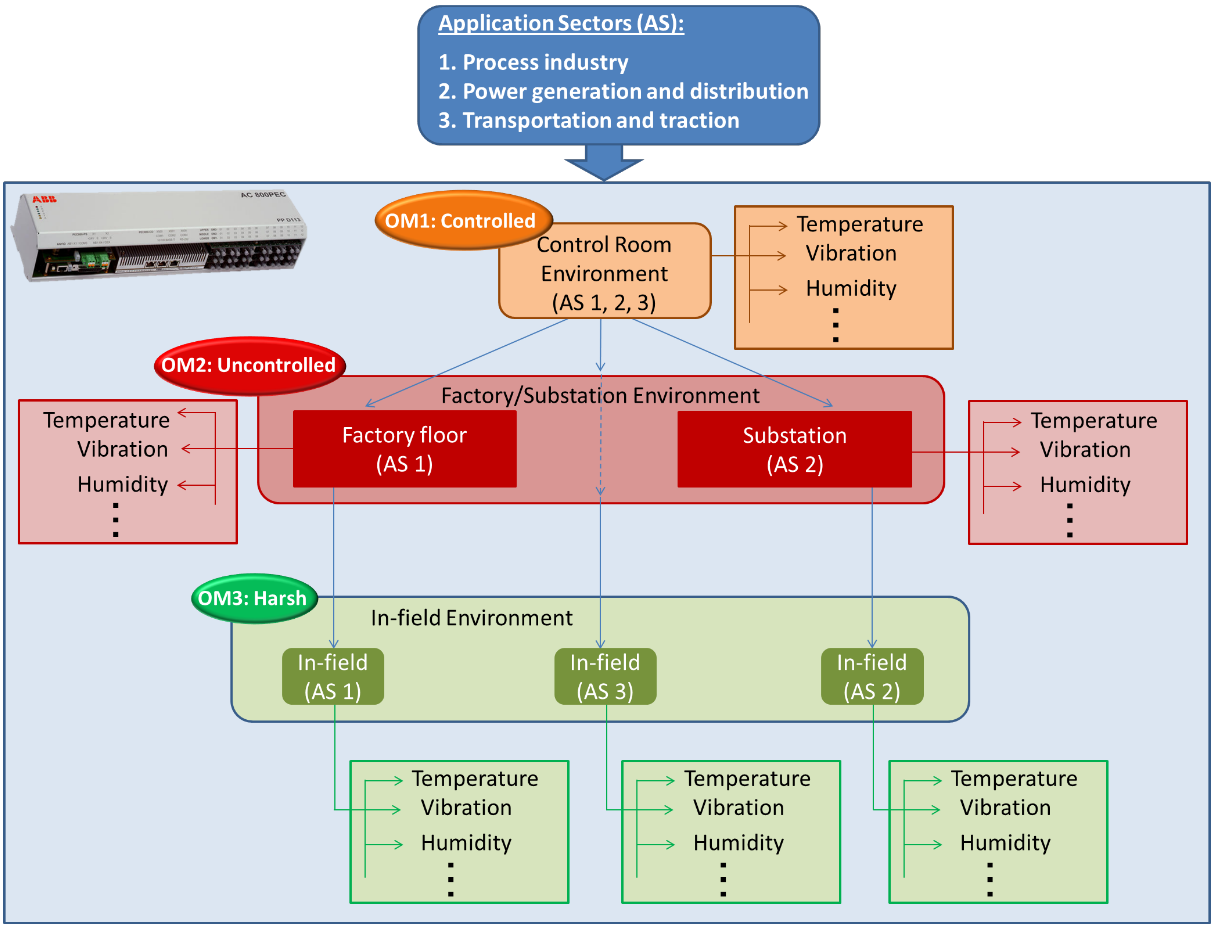


Figure 20: The environmental factors associated to the application scenarios of the AC 800PEC.

When OM3 (in-field harsh environment) requirements are applicable this is normally solved by using special cabinets (water proof, elastic suspension, etc.) or typical require the substations to be heated for protecting the AC 800PEC.

4.2.1. CONTROLLED application scenario

A CONTROLLED application scenario in industry is a control room. This is a stable environment with minor fluctuations in environmental conditions. Main focus is servers and PC workstations, but also embedded devices like switches and routers can be found here.

Table 10: List of environmental factors in Control room.

Environmental factor	Characterization	Range	Typical remedies / solutions
Temperature cycling	Stable, fan cooling required	0 to 40°C	Avoid wear out by renewing
Humidity / dust / acid / salt	Stable, regular maintenance, IP-class	No condensation (<80% humidity)	Cleaning regularly
Vibration / shock / pressure / gravity	Stable, turn off before move	-	Avoid accidents

Explosion (EXN)	Normally invalid	Non	.
EMC / EMI	Low, equipment dependent	Normal	Shielding (if applicable)
Radiation	Normally low, but altitude dependent	Normal	Shielding (if applicable)
Electrical stability / Insulation	Stable, normally with UPS and surge protection	Normal	UPS will assure shutdown
Altitude	Same height expected over lifetime, but can be at max.	Less than 2000m	-

4.2.2. UNCONTROLLED application scenario

UNCONTROLLED applications scenarios in industry are a Factory floor or a Substation. These are much more exposed to environmental conditions and sudden changes, but typically in places where human can perform work. This environment is typically industrial computers for harsh environment with a rock-solid industrial-grade performance, preferably without fan and disk.

Table 11: List of environmental factors in Factory floor / Substation.

Environmental factor	Characterization	Range	Typical remedies / solutions
Temperature cycling	Fluctuate, with/without fan (typically not installed)	Operation: -10°C to 55°C (Automotive Manufacturer) Storage: -25°C to 70°C	Overheat need heat sink
Humidity / dust / acid / salt	Burst	5% ~ 95%, non-condensing; IP65	-
Vibration / shock / pressure / gravity	Moving components	Shock: Half-sine shock test 5G/11ms, 3 shocks per axis; Operation Vibration: MIL-STD-810F 514.5 C-1	Welding, contacts, etc.
Explosion (EXN)	No risk assumed	no	-
EMC / EMI	Many components/devices in the vicinity	CE/FCC compliance	Extra EMC shielded if needed
Radiation	No, but in future products more requirements expected	-	-
Electrical stability / Insulation	Critical components will probably have UPS	Low	Power backup

4.2.3. HARSH application scenario

HARSH application scenarios are typically exposed to natural environments with wind, vibration, water or gases. Devices that are commissioned in such environments are in-field devices, typically sensors and actuators placed in rough environmental conditions, often not a place for human being. A sensor could be placed on rotating machinery or down a drilling hole and is typically connected via wire or wireless to the controller.

Table 12: List of environmental factors in In-field devices.

Environmental factor	Characterization	Range	Typical remedies / solutions
Temperature cycling	Fluctuates highly	Operating: -40° to +85°C Storage: -40 to +25°C	Overheat need heat sink
Humidity / dust / acid / salt	Burst	High	IP-class: IP66 (dust-tight and resistant to powerful water jetting)
Vibration / shock / pressure / gravity	Fast moving components, high frequent motors	High	Glue/additional mounting possible
Explosion (EXN)	Zone 0, Ex ia IIC T4	-40°C/+85°C	-
EMC / EMI	Large number of components/devices in the vicinity	High	EMC shielded
Radiation	Altitude dependent. Same height ex-pected over lifetime, but can be at max.	Less than 2000m	-
Electrical stability / Insulation	May/may not have UPS/power back up	High	-

4.3. Application scenarios in the Safety-critical and Mission Critical Sectors

Computers have been progressively introduced in aircrafts to support Fly-By-Wire systems and to assist the pilots. Now, they are everywhere in an aircraft and have multiple functions like Flight Control System, engine control, braking and steering, cockpit display system, Auto-Pilot, Flight Management System, radio-communications, drinking water control, cabin light, maintenance, etc. They are used in all categories of airplane (civil aircraft, regional airliner, and business jets), helicopters and UAV (Unmanned Aerial Vehicles). Figure 21 shows examples of avionics controllers used in an Airbus A340 airplane.

Figure 21: Airbus A340 avionic computers (P. Darphin ©THALES).

Avionics systems are based on dedicated computing platforms as well as on generic computing platforms hosting multiple applications. COTS boards can be used for the less critical functions. Even if avionic equipment mounted in the cockpit of an airplane operates in a controlled temperature and pressurized location, operating conditions are however not as stable as in a ground controlled room. It is therefore important for embedded systems operating in changing environment to precisely take account of all the evolutions of the environmental parameters and of the stress applied to the equipment.

Operating conditions may vary with the outside environment, the use of the equipment but will also depend on the type of equipment and its location in the aircraft. To estimate the reliability of such embedded systems, the life profile of the system with the evolution of operating conditions have thus to be defined. For this purpose, typical profiles of missions are generally based on specifications and in-field measurements. To note that typical environmental conditions are considered for reliability prediction. These operating conditions do not correspond to the extreme operating conditions that the equipment could encounter in field of operations and that are of interest for qualification but not relevant for reliability prediction.

Three different applications scenario, which have significantly different life profile, have been selected and are described below. First application scenario corresponds to avionic equipment in a medium haul civil aircraft while the second scenario corresponds to equipment in a helicopter. The third application scenario in the space domain is the example of a satellite in an earth orbit. The first two descriptions of the life profile have been done in agreement with the guidelines of the FIDES methodology [13], which has been elaborated for electronic reliability prediction in the aerospace domain. These descriptions of mission profiles are notably based on decomposition into phases.

4.3.1. Avionic computer mounted in a medium haul civil aircraft

The following scenario corresponds to the case of an avionic computer mounted in the cockpit of a medium haul civil aircraft. It is assumed in the life profile that such airplane operates 350 days per year, and the remaining days are used for the daily maintenance or the airplane is waiting on standby. Furthermore, a medium haul civil aircraft performs in average 3 flights per day of operation with 2 intermediate stopovers. It is considered here that the equipment is not powered off during parking and is only switched off during non-operating days. Indeed, this scenario corresponds to the worst-case situation with regards to the reliability.

The "typical" life profile of the equipment can be represented by different phases:

- **“Ground-Dormant” phase:** The average duration of this phase is assumed to be 6.2 hours including the daily stops, the first embarking and the last disembarking of the plane. With 1 cycle per day, the total duration of the phase is thus 2170 hours per year (350 x 6.2).
- **Operating phase on the ground during stopovers:** The average duration of stopovers is assumed to be 2 hours. With 2 stopovers per day (between 3 flights), the total duration of the phase is 1400 hours per year (350 x 2 x 2).
- **“Ground-Taxiing” phase**: Before the flight, this phase begins at the moment at which the aircraft leaves the boarding zone and ends at the moment at which the aircraft is ready to take-off. After the flight, this phase begins at the moment at which the airplane lands and ends when the airplane reaches the boarding zone. The average taxiing duration is assumed to be 18 minutes. With a taxiing phase before and after each flight, the total duration of the phase is 630 hours per year (350 x 3 x 2 x 0.30).
- **Flight phase:** The total duration of the average flight is assumed to be 4 hours (including 3 hours of "Stable flight" and 1 hour of “Climb/Descent”) and it is assumed that that the airplane performs 3 flights per day. So, the total duration of the phase is 4200 hours per year (350 x 3 x 4).
- **“Ground-off” phase:** This phase corresponds to the 15 days per year where the equipment is switched off (airplane in daily maintenance or in stand-by). The total duration of the phase is 360 hours per year (in other words 24 x 15).

Temperature profile and temperature cycling

Even if avionic systems are located in temperature controlled zones, their operating temperature is quite important and fluctuates during the day. The ambient temperature is assumed to be 15°C when the equipment is off. When the avionic computers are switched on, the internal temperature inside the equipment rises to 40°C due to the heat dissipation inside the equipment and inside the zone of the equipment. This temperature remains constant during the group operations and during the flight. However, the temperature regulation is less efficient during the stopovers (aircraft doors being opened) and as a result the temperature increases to 55°C during these phases.

This temperature profile is illustrated by the following diagram:

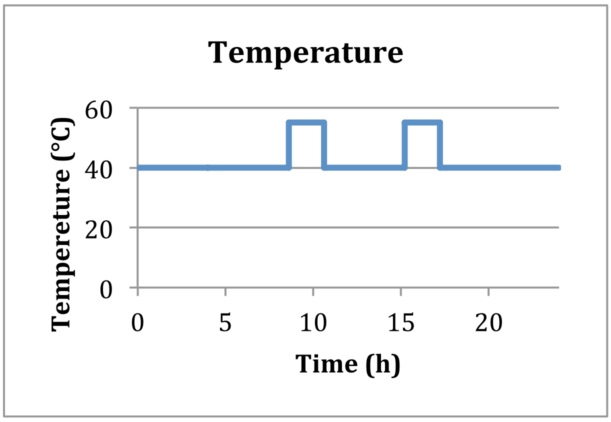


Figure 22: Temperature profile.

Humidity and chemical profile

When the airplane is in parking or in daily maintenance with electronic equipment switched off, the average humidity is of about 70%. When electronic equipment is ON, the relative humidity decreases due to the internal temperature rise. If the aircraft doors are opened (Ground - Dormant, Ground - Operation Stopover), the relative humidity is then estimated at 30%. In flight or if the aircraft doors are closed on the ground (Ground - Taxiing) the ambient air is very dry and the relative humidity is of only 10% due to the ventilation system.

In avionic controllers operating in a closed environment the level of pollution is globally low. The equipment is only subject to the outside environmental pollution on the ground.

Vibrations

Avionic systems are of course subject to vibrations. The vibration stress is very low during ground operations at the exception of the parking phase during which it is assumed to be 5 Grms. During the flight phase, the average vibration stress is estimated to be 0.6 Grms.

Atmospheric radiation environment

Ionizing particle fluxes increasing with altitude, the atmospheric radiation environment can be a significant source of problems for avionic systems [14]. As atmospheric neutrons have been identified as the main cause of single event effects (SEE) at elevated altitudes, only the neutron flux is considered for reliability prediction. The evolutions of the radiation environment with the altitude of the airplane, its geographical position and with the location of equipment inside the airplane are not taken into account. The worst-case operating conditions are considered for reliability evaluation in the avionic domain. It is thus assumed that we have a neutron flux of 8600 $n.cm^{-2}.h^{-1}$ (neutron flux at 40 000 feet).

Pressure

Avionic systems are generally installed in a pressurized location. So, the pressure inside the equipment is (almost) stable and no pressure stress is applied to the equipment.

Electromagnetic environment

Lightning and sources that radiate radio-frequencies (VHF, weather radar, etc.) are the most severe electromagnetic threats in an airplane. Avionics systems have thus to be adequately protected against the effects of lightning and radio frequency energy. Requirements and test procedures regarding the electromagnetic environment and power supply are covered in the standard RTCA/DO-160G.

Table 13 gives the identified environmental factors for the different phases of the typical life profile of an avionic computer in a medium haul civil aircraft.

Table 13: List of environmental factors for an avionic controller.

	Phase title	Ground-Off	Ground - Dormant	Ground - Taxiing	Flight	Ground - Operation Stopover
	Calendar time (hours)	360	2170	630	4200	1400
Temperature and Humidity	On/Off	Off	On	On	On	On
	Ambient temperature (°C)	15	40	40	40	55
	Relative humidity (%)	70	30	10	10	30
Temperature cycling	ΔT (°C)	10	-	-	-	15
	Number of cycles (/year)	15	350	2100	1050	700
	Cycle duration (hours)	24	6.2	0.30	4.00	2.00
	Maximum temperature during cycling (°C)	20	-	-	-	55
Mechanical	Random vibrations (Grms)	-	0.05	5	0.6	0.05
Chemical	Saline pollution	Low	Low	Low	Low	Low
	Environmental pollution	Moderate	Moderate	Low	Low	Moderate
	Application pollution	Low	Moderate	Moderate	Moderate	Moderate
	Protection level	Non hermetic	Non hermetic	Non hermetic	Non hermetic	Non hermetic

4.3.2. Avionics computer on-board a helicopter

The second application scenario that has been selected in the domain of critical ES is the one of an avionic computer in the cockpit of a VIP helicopter. It is estimated that a VIP helicopter is in service for 250 days per year and that it is in stand-by or in daily maintenance during the remaining days (115 days per year). A VIP helicopter typically operates at a daily rate of 2 flights per day of operation (an outbound and return flight).

The mission profile is composed of 3 phases:

- **Off phase:** Electronic equipment is switched off each time the helicopter is in parking.
- **Ground-on phase:** The operation phase on the ground includes the test of the electronic systems, the preparation for take-off and the time to stop systems. The average duration of this phase is 30 minutes. Electronic systems and the engine are ON during this phase.
- **Flight phase:** The average duration of a flight is 1 hour.

Temperature profile and temperature cycling

The operating temperature of avionic equipment is higher in a helicopter than in an airplane due to less efficient ventilation. During the parking phase, the average ambient temperature is assumed to be 15°C varying between 10°C at night and 20°C during the day. But when the electronic systems are switched on, the temperature inside the equipment rises to 75°C. The temperature during the flight is assumed to be constant and equal to the temperature on the ground with the electronic equipment powered on (+75°C).

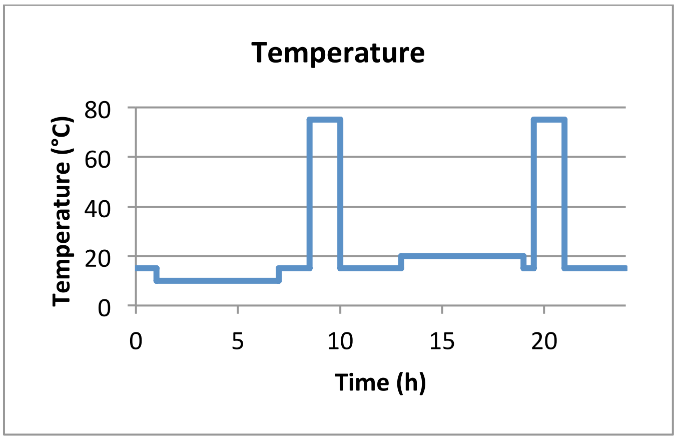


Figure 23: Temperature profile.

Humidity and chemical profile

The levels of humidity in a helicopter and in an airplane are quite similar. When the equipment is off, it is assumed that the average relative humidity is 70% for the considered scenario. When the equipment is switched on, a drying of ambient air is caused by the heat dissipation from equipment and the humidity drops to 20%.

The level of pollution is also globally moderate.

Vibrations

Vibrations are generated by the engine on the ground. The vibration amplitude is of about 0.5 Grms on the ground. During the flight, the vibration amplitude increases to 6 Grms. The vibration stress is indeed higher in a helicopter than in an airplane and happens mainly during the flight.

Atmospheric radiation environment

The radiation environment is less severe for a helicopter than in an airplane. An average helicopter has indeed a ceiling altitude of 10 000 feet. The neutron flux at this altitude is about 400 n.cm^{-2}.h^{-1}.

Pressure

The cabin of a helicopter is not pressurized (at the difference of an airplane) but the flight altitude is quite low. So, avionic equipment in helicopter is not subject to specific pressure stress.

Electromagnetic environment

The same rules apply for a helicopter and an airplane.

Table 14 lists the environmental factors of a typical life profile for an avionic computer onboard a helicopter.

Table 14: List of environmental factors for a computer on board a helicopter.

	Phase title	Off	Ground-on	Flight
	Calendar time (hours)	8010	250	500
Temperature and Humidity	On/Off	Off	On	On
	Ambient temperature (°C)	20	75	75
	Relative humidity (%)	70	20	20
Temperature cycling	ΔT (°C)	10	55	-
	Number of cycles (/year)	365	500	500
	Cycle duration (hours)	21	0,5	1
	Maximum temperature during cycling (°C)	20	75	-
Mechanical	Random vibrations (Grms)	0	0.5	6
Chemical	Saline pollution	Low	Low	Low
	Environmental pollution	Moderate	Moderate	Moderate
	Application pollution	Moderate	Moderate	Moderate
	Protection level	Non hermetic	Non hermetic	Non hermetic

4.3.3. Space application scenario

Spacecrafts are used for different missions: earth observation and meteorology, communication and navigation, astronomy, human exploration and space stations. According to mission types, different orbits are typically used: HEO for science, GEO for telecommunications, LEO for earth observations, MEO for global navigation. As electronic systems used in satellites have to operate continuously in a harsh environment, it is mandatory to take account of the space environment when designing such systems. Indeed, these systems are directly exposed to solar wind, Van Allen radiation belts and cosmic rays.

Temperature profile and temperature cycling

The temperature profile of a satellite is strongly dependant on its orbit even if a thermal control of the satellite allows to limit the excursion of the operating temperature and to reduce thermal cycling. The operating temperature is typically in the range [-10°C, +60°C]. The thermal cycling period is of 90 minutes in the case of a satellite on a LEO orbit and of 24h in the case of a satellite on a GEO orbit.

Vibration and shock

A spacecraft is subject to vibrations and shocks during the launch and orbit corrections. Electronic systems have to be appropriately designed to support them and are tested accordingly.

Radiation environment

The radiation environment of satellites is particularly severe and also depends on the orbit. A satellite on a LEO orbit is typically exposed to a Total Ionizing Dose (TID) value of 0.1 krad per year, while a satellite on a GEO orbit is typically exposed to TID value of 10 krad per year. SEU rates are predicted using the CREME model [15] according to the characteristics of the mission and take account of shielding of electronic units.

Pressure

Satellites have to operate in ultra-high vacuum. The pressure at a geosynchronous orbit is in the order of 10^{-9} Pa.

Electromagnetic environment

The plasma environments can cause spacecraft charging that can provoke EMI and ESD in electronic units. The problem is also mitigated at design and qualification.

5. Conclusions/Summary

The information in this document gives an insight into the typical operational environments future HPC/ES are expected to operate in. The information in this document will be useful when studying the impact of the different sources of failure in the final reliability assessment.

A key observation is that HPC and Industrial ES follow standards to conform directives like EC. Thus, they do not only collect specific customer requirements related to the environmental factors. Customers in these markets rather use the standards to determine whether their application scenarios fit or not. If they do not fit, customers look for special devices like ruggedized devices or hardened products.

This report has also shown that for most of the application scenarios a standard set of parameters are sufficient. In most cases this will be temperature, vibration, humidity, dust and EMC. Thus, it is crucial that these factors are taken into account within the CLERECO project.

We have tried to map application scenarios onto Operation Modes. For HPC with mainly server farms and Industrial ES with relatively location fixed devices this worked very well. However, Mission Critical ES cannot be easily classified statically, because they move between Operation Modes like a flight with a plane or a helicopter.

This report has also revealed that the different domains share several similarities like:

- Security and Safety requirements
- Meshed networks – everything is connected via wireless or Ethernet
- High availability – a special need for redundancy
- Easy to maintain and repair
- Standard solutions preferred

This document will be input for both, optimal selection of future technologies and also the right set of design changes for HW and SW in order to achieve high reliability for the systems. So, this document will serve as a key input to WP3, WP4 and WP5 of CLERECO project.

6. Acronyms and Definitions

6.1. Table of Acronyms

Abbreviation	Term
ABB	Company, one of the industrial partners

API	Application programming interface
AS	Application scenarios
ASIC	Application-specific integrated circuit
BSP	Board Support Package
CAE	Computer aided engineering
CE	The CE marking or formerly EC mark, is a mandatory conformity marking for certain products sold within the European Economic Area (EEA) since 1985
CLERECO	Cross-Layer Early Reliability Evaluation for the Computing cOntinuum
COTS	Commercial Off-The-Shelf
DCC	Digital content creation
DO	Produced by Radio Technical Commission for Aeronautics, Inc. (RTCA), the FAA's Advisory Circular AC20-115B establishes DO standards as the accepted means of certifying all new aviation software.
DOC	Declaration of Conformity
DoW	Description of Work
DRAM	Dynamic RAM
DSP	Digital Signal Processor
EC	
EDA	Electronic design and automation
EMC	Electromagnetic Compatibility
EMI	Electromagnetic Interference
EMP	Electromagnetic pulse
EN	European Norm
ES	Embedded Systems
ESD	Electro-static discharge
FIT	Failure in time
FMEA	Failure mode and effects analysis

FPGA	Field programmable gate array
FTA	Fault tree analysis
GP	General Purpose
GPS	Global Positioning System
HAL	Hardware abstraction layer
HEMP	High-altitude electromagnetic pulse
HPC	High Performance Computing
HPM	High power microwave
IC	Integrated circuit
IDC	the International Data Corporation
IDE	Integrated development environment
IEC	International Electrotechnical Commission
IMA	Integrated Modular Avionics
IO	Input Output
IP	Intellectual property
ISO	International Organization for Standardization
LVD	Low Voltage Directive
M2M	Machine to machine
MCU	Microcontroller unit
MPU	Microprocessor unit
MTBF	Mean Time Between Failure
MTTF	Mean Time To Failure
OM	Operational Modes
OS	Operating system
PC	Personal Computer
PCB	Printed circuit board

PFD	Probability of dangerous Failure on Demand
PFH	Probability of dangerous Failure per Hour
PLC	Programmable Logic Controller
RTOS	Real-time Operating System
SOC	System on Chip
TCO	Total-Cost-of-Ownership
TID	Total Ionizing Dose
UAV	Unmanned Aerial Vehicles
UPS	Protection system
WCET	Worst-Case Execution Time
WP	Work package

6.2. List of Definitions

The following definitions are specific within the context of reliability of computing systems.

- **Operation mode:** Set of computing system conditions (external and internal) that relevant for reliability assessment.
- **Environmental factors:** Set of computing system external aspects that may influence reliability (e.g. temperature and altitude).
- **Environmental conditions:** Specific values or ranges of environmental factors (e.g. temperature between -10°C and 20°C).
- **Internal execution modes:** Set of internal computing system aspects that may influence reliability (e.g. booting, running, and maintenance).
- **Computing segments:** Main classification of computing systems according to the end use (e.g. Embedded Systems (ES), cell phones, laptops, and High Performance Computing (HPC)).
- **Application sector (AS):** Category of computing system usages per computing segment (e.g. process industry, power generation, and transport for ES)
- **Application scenario:** Usage scenario for a computing system (e.g. spacecraft computing, military applications, medical equipment, and supercomputing system) that determines a set of environmental conditions and typical software stack. May be categorized inside a computing segment and probably an application sector.

7. References

[1] Arnljot Høyland, Marvin Rausand “System Reliability Theory: Models, Statistical Methods, and Applications” ISBN 0-417-47133-X

[2] TOP500 Supercomputer Sites, website: http://www.top500.org/

[3] Earl Joseph, IDC, HPC Trends, Big Data and the Emerging Market for High Performance Data Analysis (HPDA), Salishan, April 2013

[4] François Robin, GENCI and CEA, HPC Trends: Opportunities and Challenges, PRACE Industrial Seminar, Amsterdam, September 3, 2008

[5] Nicholas P. Carter, Helia Naeimi and Donald S. Gardner, Design techniques for cross-layer resilience, DATE, 2010

[6] Jon Stearley, Defining and measuring supercomputer Reliability, Availability, and Serviceability (RAS), In Proceedings of the Linux Clusters Institute Conference, 2005

[7] ABB document, Conformity with standards for protection and control equipment, 1MRK 500 019-BEN, April 2002

[8] S. Tverdyshev, “PikeOS: Multi-Core RTOS for IMA”, International Scientific and Practical Conference «Integrated Modular Avionics. State and Prospects of Development», Moscow, October 29-30, 2012.

[9] ASHRAE TC 9.9, Thermal Guidelines for Data Processing Environments – Expanded Data Center Classes and Usage Guidance, 2011

[10] Neil Rasmussen, and Wendy Torell, Data Center Projects: Establishing a Floor Plan, White Paper, 2007

[11] Inspur Manufacturer and Provider, website: http://www.inspurglobal.com/product-cloud_blade_server_nx5440-58.html

[12] Apple Inc., website: http://store.apple.com/es/buy-mac/mac-pro

[13] FIDES Group, “FIDES guide 2009 - Reliability Methodology for Electronic Systems”, Edition A, September 2010.

[14] E. Normand, “Single-event effects in avionics”, IEEE Transactions on Nuclear Science, vol.43, no.2, pp. 461-474, 1996.

[15] A.J. Tylka et al., “CREME96: A Revision of the Cosmic Ray Effects on Micro-Electronics Code,” IEEE Transactions on Nuclear Science, vol. 44, no. 6, pp. 2150-2160, Dec 1997.

[16] EMC-directive 89/336/EC (valid from 20th July 2007: EMC-directive 2004/108/EC)

[17] Low Voltage Directive: LVD 73/23/EEC with update to LVD 2006/95/EC

[18] ISO/IEC 2382-14:1997 Information technology - Vocabulary - Part 14: Reliability, maintainability and availability

[19] Damien Hardy, Marios Kleanthous, Isidoros Sideris, Ali G Saidi, Emre Ozer, and Yiannakis Sazeides, An analytical framework for estimating TCO and exploring data center design space, ISPASS 2013.

[20] G. Edelin. Embedded Systems at THALES: the Artemis challenges for an industrial group. In ARTIST Summer School in Europe, 2009.

[21] EASA. Annual Safety Review 2009, website: http://www.easa.europa.eu/

[22] John Heggestuen, Business Insider; One In Every 5 People In The World Own A Smartphone, One In Every 17 Own A Tablet, website: http://www.businessinsider.com/smartphone-and-tablet-penetration-2013-10, December 2013.

[23] Danyl Bosomworth, Smart Insights, Mobile Marketing Statistics 2014, website: http://www.smartinsights.com/mobile-marketing/mobile-marketing-analytics/mobile-marketing-statistics/, March 2014.

[24] Nick Olanoff, The Comparison Between Samsung Galaxy S4 and iPhone 5, website: http://www.technologyace.com/mobile/comparison-samsung-galaxy-s4-iphone-5/, July 2013.