

Chapter 20

Controls Technologies

J. Serrano, G. Daniluk, E. Gousiou and C. Roderick

CERN, BE Department, Genève 23, CH-1211, Switzerland

HL-LHC will pose new challenges on the accelerator control system. Although the overall architecture will be preserved and most of the currently deployed equipment will continue its operation, three areas have been identified for renovation in response to the new requirements: logging system, new hardware platform in the distributed I/O tier and radiation-tolerant fieldbus.

1. Overview

For the commissioning and subsequent operation of the HL-LHC, some physical elements, in particular front and back-end CPUs and storage will have been upgraded due to obsolescence. However, the overall control system strategy and architecture is sufficient for the HL-LHC needs and will not change in its conceptual structure.

Nevertheless, three areas have been identified as needing to be addressed so that the control system can respond to the new challenges: logging system, new hardware platform in the distributed I/O tier, and radiation-tolerant fieldbus.

During HL-LHC operation there will be an increase of radiation in some areas which will require re-designs and relocation of electronics. There are also new magnets which will raise the need for more diagnostics data in different subsystems. Higher data rates will also be needed during the commissioning of HL-LHC as equipment groups will need to fine-tune their systems and will

This is an open access article published by World Scientific Publishing Company. It is distributed under the terms of the Creative Commons Attribution 4.0 (CC BY) License.

therefore require more diagnostics. In order to assure correct functionality up to the end of the HL-LHC operation period with ultimate performance, it is important to be conservative regarding the design choices and to share proven solutions as much as possible. This approach assures that proven solutions persist and that all efforts can be concentrated on making a few designs very robust instead of spreading efforts into a large number of sub-optimal designs.

2. Data Logging

The CERN Accelerator Logging Service (CALs) was designed in 2001, has been in production since 2003 and stores data from all of CERN's accelerator infrastructure and beam observation devices. Initially expecting 1TB / year, the Oracle-based system scaled to cope with 2.5TB / day coming from >2.3 million signals. It serves more than 1000 users making an average of 5 million extraction requests per day. CALs is considered as being mission-critical and the go-to service when investigating problems with equipment or unexpected beam behavior.

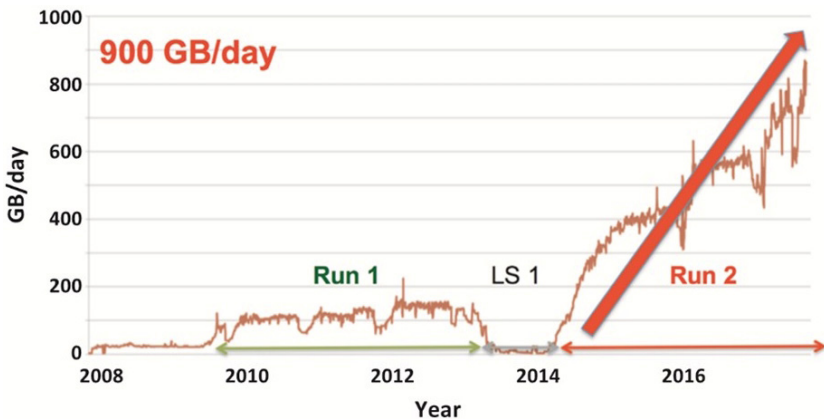


Fig. 1. Logging Service daily storage evolution in GB / day.

The CALs system has scaled well in terms of ensuring long term storage of acquired data and providing linear response times for data extractions. However, with basic accelerator operation reaching a high level of maturity, attention has turned to more complex analyses such as studying beam effects

over longer periods of time. CALS has increasingly been subjected to extraction of much larger datasets over longer periods of time to support advanced data analytics. It is in this domain, during LHC Run 2, that the CALS system quickly started to show its limits. In 2016, the NXCALS project was launched with the aim of replacing CALS from LHC Run 3 onwards. The idea is to gain operational experience with NXCALS during several years and then have time to adapt further as needed, during LS3 while still ahead of High-Luminosity LHC commissioning.

2.1. NXCALS Architecture and Technologies

In recent years, the so-called “Big Data” technology landscape has evolved significantly to support large-scale data logging and analysis, opening up new possibilities to perform efficient analysis of large data sets.

The NXCALS system is based on a microservices architecture as shown in Figure 2. The aim of this is to be able to easily upgrade or replace different aspects of the system in the future as necessary, without being forced to put in place a completely new system. From a technology perspective, NXCALS is based on in-house developments combined with open-source software such as Hadoop (HDFS and HBase) [1], Kafka [2], Spark [3], and Jupyter notebooks [4].

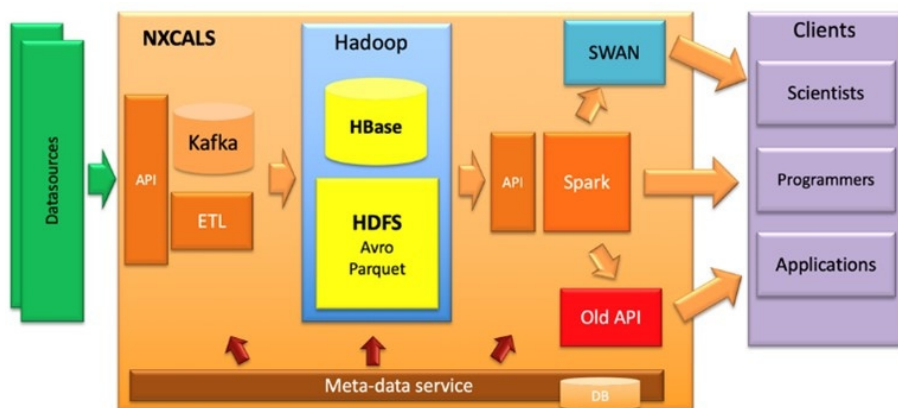


Fig. 2. NXCALS architecture.

Regarding ingestion, data is sent to the system from data acquisition processes called “Datasources”, to Apache Kafka via the NXCALS data ingestion API. Kafka is a highly reliable, high-throughput, low-latency platform for handling real-time data feeds. In NXCALS, data is stored on Kafka until it has been transferred into Hadoop by an in-house developed ETL (Extract-Transform-Load) process.

The data is stored in the Hadoop layer which is comprised of two main parts: HBase which serves as a low-latency repository from which data of the last 36 hours can be extracted by users and HDFS which serves as the long-term storage of data, in highly compressed Apache Parquet files.

A client API based on Apache Spark, with NXCALS-specific extensions, allows users of NXCALS to extract data and/or perform advanced data analysis. Spark is an analytics engine for performing large-scale distributed data processing on computing clusters. SWAN (Service for Web based ANalysis) is a CERN platform to perform interactive data analysis from the Web using Apache Jupyter notebooks.

In order to properly manage the overall coherency of the system, an in-house developed service is employed to manage the meta-data.

Finally, the core technologies used in NXCALS are based on the concept of “horizontal scalability”, which essentially means the ability to increase performance by adding more resources to the underlying infrastructure. From this perspective, the NXCALS system has the potential to adapt to the required performance needs of the future, provided sufficient resources can be financed and that sufficient physical hosting capacity is available.

3. Distributed I/O Tier Modular Kit

The HL-LHC will place challenging demands on data acquisition to/from the accelerator components which need to be controlled and diagnosed, such as the new Nb₃Sn magnets. The need for larger amounts of diagnostics information will result in a requirement for more throughput in the lower layers of the control system and will therefore affect the electronics in this tier and the communication links used to send the information up the controls stack. The current custom electronics-based controls architecture has front-end computer systems (VME or PICMG 1.3) with a large variety of reusable electronic cards to control accelerator components by sending and receiving

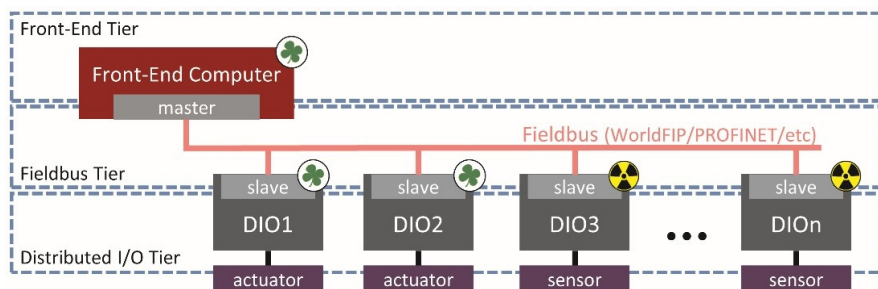


Fig. 3. Three lowest hardware layers of a typical control system.

data and carrying out calculations in real-time. In the LHC, these front-end computers typically drive a fieldbus which connects to Input/Output (I/O) modules sitting close to the accelerator, as shown in Fig. 3.

Historically, in the Front-End and Fieldbus tiers there has been a lot of sharing and reuse of design effort between equipment groups, unlike the lower Distributed I/O Tier (DI/OT) where we find many custom-made modules in different form factors.

For HL-LHC, the proposal is to extend the sharing model of the Front-End and Fieldbus to the DI/OT layer. The electronics in this layer are designed to perform early data processing and transmit to/from actuators and sensors attached to accelerator components. These I/O modules are connected to a fewer number of high-performance front-end computers which further process the data and perform the necessary calculations. By collaborating with equipment groups and providing a centralized service in the DI/OT layer, we will ensure a uniform level of quality and increase of the overall availability of electronics deployed in this tier, including those subject to radiation.

3.1. DI/OT hardware kit

In the frame of the HL-LHC project (Work Package 18) a generic and modular hardware kit (Fig. 4) is being developed in close collaboration with equipment groups, allowing different applications to benefit from a common infrastructure. The kit targets both radiation-exposed and radiation-free areas. It will consist of a 3U crate conforming to the CompactPCI Serial (CPCI-S) standard, one radiation-tolerant System Board (crate controller board for radiation-exposed systems), one non-radiation-tolerant System Board (crate controller

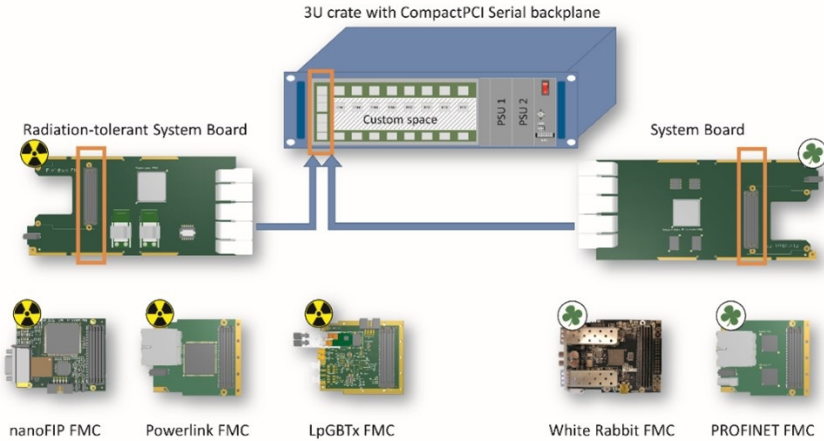




Fig. 4. Distributed I/O Tier modular and reusable hardware kit consisting of radiation-tolerant  and non-radiation-tolerant  modules.

board for radiation-free systems) and a set of interchangeable fieldbus communication mezzanines. An instance of the DI/OT crate will consist of one System Board hosting a single fieldbus mezzanine and the remaining crate slots will be filled with application-specific boards. Fieldbus mezzanines implement various communication technologies (*WorldFIP*, *Powerlink*, *LpGBTx*, *White Rabbit*, *Profinet*) and ensure control and data exchange with the Front-End tier. The System Board acts as the crate controller and interfaces with the application-specific Peripheral Boards designed by the equipment groups, plugged into the other slots of the 3U crate.

Each System Board can be programmed with desired early data processing algorithms as it features a Field Programmable Gate Array (FPGA) for the application-specific logic. A small fraction of the FPGA resources is dedicated to implement general crate monitoring services (e.g. temperatures, voltage levels, current levels, fan speeds).

Basing the kit on industrial standards (to benefit from the work already done by a large community) and designing the individual modules with experts from equipment groups and external companies (to benefit from the review of many developers) are the key principles of the project contributing to the increase in machine availability. On top of that, dedicated resources for reliability studies of all the components of the kit will provide a clear reliability assessment.

The modularity of this kit caters to different needs in equipment groups. Survey, WIC (Warm Interlocks) and PIC (Powering Interlocks) will use the full radiation-tolerant kit, including the crate, the System Board and the WorldFIP communication mezzanine. They will design their own add-in boards (Peripheral Boards) in 3U Europa card format to interface with their sensors and actuators. Other groups have their own designs for a system board and will plug one of the DI/OT communication mezzanines in it. The non-radiation-tolerant variant will be used in the Full Remote Alignment System (FRAS).

3.1.1. *DI/OT crate*

Among the various industrial standards for modular electronics, CompactPCI Serial [5] was selected as a base for the DI/OT 3U crate. The standard features a robust connector targeting transportation applications and a fully passive backplane, which makes it suitable for systems in radiation-exposed areas. Although it specifies complex protocols like PCI-Express, USB and SATA for inter-board communication, the passive backplane enables the use of the basic physical infrastructure of CompactPCI Serial without following further prescriptions on protocols. A simple communication technology (such as high-speed SPI) will be used instead, with support for automatic discovery of hardware modules. This is much more suitable for electronics in radiation-exposed areas, where the complexity of a system must be reduced as much as possible.

The fact that the DI/OT crate complies with the CPCI-S specification, enables designers to use off-the-shelf crates for lab prototyping as well as standardizes the voltages, connectors and monitoring interfaces inside the crate. However, the crates currently available on the market are not suitable “as-is” for wide deployments in HL-LHC mainly due to dimension and cost limitations. To overcome these, an open hardware crate and CPCI-S backplane are being designed in the frame of the project. The crate design will use a standard 3U sub-rack mechanical kit (available from all major crate manufacturers). It will also allow hosting boards that are longer (220mm) and wider (6 Horizontal Pitch) comparing to most common 160mm x 4 Horizontal Pitch CPCI-S boards.

To further increase the availability of DI/OT systems, the crate will be equipped with dual modular redundancy power supplies in load sharing configuration. For radiation-free areas, these will be off-the-shelf CPCI-S power supplies. However, for radiation-exposed applications, such a straightforward approach cannot be applied. Regular, switch-mode power supplies are known to fail in radiation due to both single event effects (SEE) and total ionizing dose (TID). Therefore, the vast majority of radiation-tolerant electronics currently deployed at CERN is equipped with linear power supplies. Those are less complex and thus more resilient to radiation-related effects. However, linear power supplies suffer from poor efficiency, large heat dissipation and a form factor dictated by the bulky 50Hz transformer for output powers in the order of 100W (required for DI/OT).

The first, by design radiation-tolerant, 100W switch-mode AC/DC power supply is being developed for the DI/OT hardware kit. It will be mechanically compliant with the CPCI-S standard and capable of providing 100W on +12V DC and 10W on the +5V DC power rail.

4. Radiation-tolerant Fieldbus

For the HL-LHC it has been foreseen to anticipate higher rates of data extraction from the machine by the equipment groups. Currently the only radiation-tolerant fieldbus for the accelerator is WorldFIP [6]; while operating reliably since the first LHC start-up, its bandwidth is limited to 2.5Mbps. The complexity of the Nb₃Sn magnets increases by a factor of 10 the amount of post-mortem data that needs to be transmitted for the QPS system for example, this would make the current solution based on WorldFIP sub-optimal.

An industrial solution, based on 100Mbps Ethernet, providing μ s-level synchronization and supporting up to 50 slaves per segment is proposed. After a market review including leading Industrial Ethernet technologies such as Profinet, EtherNet/IP and EtherCAT, it was decided to design a radiation-tolerant implementation of Ethernet Powerlink [7]. It is the simplest of the Industrial Ethernet protocols which makes it feasible to implement a slave-node in a radiation-tolerant FPGA. Moreover, Powerlink features an open-source implementation of its stack which gives us direct access to reliable source code.

A mature option for making radiation-tolerant digital designs, is using flash-based FPGAs. Critical applications in the accelerator sector are making use of FPGA families like ProASIC3 and Smartfusion2 which have proven to be reliable for doses of a few hundred Gy. The radiation-tolerant WorldFIP slave-node for example features a ProASIC3 FPGA. The FPGA configuration is stored in flash-cells which are immune to Single Event Upsets (SEU) in the LHC environment. The pure logic is protected from SEUs by applying triple-modular-redundancy of the flip-flops, followed by voting. A more complex approach to radiation-tolerant digital designs that is being evaluated at the time of writing is instantiating a soft-core processor inside the flash-based FPGA. While replacing a complex HDL design with a real-time processor running software is a wide-spread technique outside of radiation, it has not yet been established under radiation. Fig. 5 shows the most likely scenario for the implementation of a radiation – tolerant Powerlink stack inside a flash-based FPGA. A RISC-V core is triplicated and runs software stored in ECC-protected (Error-Correction Code) memory. The data for the program resides in a separate ECC-protected memory.

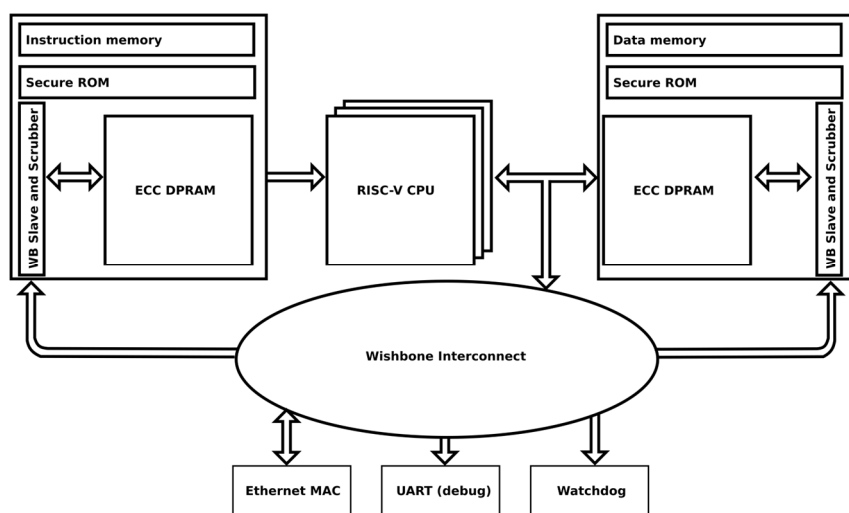


Fig. 5. Simplified block diagram of the rad-tol Ethernet Powerlink implementation.

The main challenge in our context is the validation of the RISC-V implementation in a typical flash-based FPGA and to reduce the size of the open-source Powerlink stack, originally developed for desktop systems with no memory limitations, so as to run in the amount of memory available in such FPGAs.

At the time of writing, an alternative to the flash-based FPGAs is being evaluated at CERN, the new family of low-cost and rad-hard-by-design SRAM-based FPGAs launched by the European company nanoXplore. Being rad-hard-by-design these FPGAs are tolerant to MGy doses and are immune to SEU in the configuration and the logic. They would not require triplication/voting techniques and offer memories with embedded ECC protection; this would simplify the implementation of the Powerlink stack on a RISC-V processor.

For the master side of the fieldbus, which is always outside of radiation, Powerlink being a standard, it offers commercial off-the shelf solutions, including PCIe add-in boards hosted in Linux PCs and also bus masters in Programmable Logic Controllers (PLCs). This illustrates a common theme in this work package: using industry standards as far as possible to benefit from a set of verified solutions and customizing them only as needed.

Table 1.

Aspect	LHC (2018)	HL-LHC
Logging database data rate	CALS performance limited by current architecture to 2.5 TB/day.	NXCALS horizontally scalable architecture, where performance can be increased by adding more database servers.
Resources optimization and cost reduction	Each group independently develops custom, application-specific solutions.	Common effort between groups to design a modular radiation-tolerant DI/OT ecosystem with high reliability and availability.
Efficient radiation-tolerant power supplies	Mostly linear power supplies with expensive 50Hz transformers, low efficiency (~50%) and large dimensions.	Switching mode power supply with high efficiency of at least 80%, lower cost and smaller form factor.
Radiation-tolerant fieldbus data rate	2.5 Mbps	100 Mbps

Table 2. HL-LHC systems improved by WP18.

Application	LHC (2019)	HL-LHC
Alignment and internal metrology (WP15)	Little remote diagnostics, regular personnel interventions in the tunnel.	Resources optimization and cost reduction: By-design radiation-tolerant control electronics with full diagnostics and remote alignment corrections. Efficient radiation-tolerant power supply
Warm Interlocks Controller (WP7)	Control based on commercial off-the-shelf industrial modules, not radiation-tolerant by design and not available anymore.	Resources optimization and cost reduction: By design radiation-tolerant control electronics. Efficient radiation-tolerant power supply
Powering Interlocks Controller (WP7)	Control based on legacy custom electronics and commercial off-the-shelf modules, not radiation-tolerant by design and not available anymore.	Resources optimization and cost reduction: By design radiation-tolerant control electronics. Efficient radiation-tolerant power supply
Beam Loss Monitors / Beam Position Monitors (WP13)	No link redundancy.	Resources optimization and cost reduction: Easy integration of a redundant supervision link.

5. Summary

The new Controls Technologies will provide a hardware ecosystem and improved services to the equipment groups. Table 1 highlights the main assets of the new technologies while Table 2, how the equipment groups will be benefitting from them.

References

1. <https://hadoop.apache.org/>.
2. <https://kafka.apache.org/>.
3. <https://spark.apache.org/>.
4. <https://jupyter.org/>.
5. PICMG CompactPCI Serial (CPCI-S.0) Rev 2.0.
6. <https://www.ohwr.org/projects/cern-fip/wiki/WorldFIP>.
7. <https://www.ethernet-powerlink.org>.