Development of High-Availability ATCA/PCIe Data Acquisition Instrumentation

Miguel Correia, Jorge Sousa, António J. N. Batista, Álvaro Combo, Bruno Santos, António P. Rodrigues, Paulo F. Carvalho, Bernardo B. Carvalho, Carlos M. B. A. Correia, and Bruno Gonçalves

***Abstract*—Latest Fusion energy experiments envisage a quasi-continuous operation regime. In consequence, the largest experimental devices, currently in development, specify high-availability (HA) requirements for the whole plant infrastructure. Highly available systems operate seamlessly in the case of component failure, ensuring safety of equipment, people, environment and investment. Control and Data Acquisition (C&DA) systems for Fusion diagnostics are considered mission-critical and require high degrees of availability. IPFN developed a C&DA system for fast control of advanced Fusion devices, targeting HA. The hardware platform is based on in-house developed Advanced Telecommunication Computing Architecture (ATCA) instrumentation modules–digitizing and data switch blades using PCI Express (PCIe) over the ATCA backplane Fabric Channel Interface, connecting to an external PCIe host computer. At the hardware management level, the system architecture takes advantage of ATCA's HA characteristics, such as its redundant hardware components, redundant backplane topologies and Field Replaceable Unit (FRU) "Hot Swap". At the software level, PCIe supports "Hot Plug" graceful device insertion and removal. The implementation of PCIe Hot Plug for the ATCA form-factor is one of the major tasks involved and a solution for such implementation was also developed. The paper describes how IPFN C&DA system can be setup to take advantage of both ATCA and PCIe features to perform with the desired degree of availability, by implementing fail-over mechanisms based on the use of redundancy, thus being suitable for advanced C&DA systems in Fusion.**

***Index Terms*—ATCA, high availability, hot plug, hot swap, PCIe, redundancy.**

I. Introduction

NEXT generation of Fusion devices envisages plasma scenarios to the scale of a future power plant. The mission is to demonstrate not only the feasibility of Fusion power but also to ensure that there will not be a negative impact on equipment, people, environment or investment. Diagnostics are an essential part of a Nuclear Fusion reactor's operation, providing an extensive set of measurements of plasma behaviour. Diagnostics for machine protection and plasma control are mission-critical and will consequently face stringent reliability and availability demands, extensive to the associated control and data acquisition (C&DA) systems [1]. Availability studies for the next generation devices have been performed since the first years of their conception. A first approximation was to target Fusion reactors to the same inherent availability level of a Fission reactor (70% to 80%) [2]. A more recent RAMI study for ITER indicates a figure of 60% [3]. However, the breakdown of this value in ITER's multiple functionalities, tasks and subsystems reveals a figure of 99% for C&DA systems, setting new challenges to the involved instrumentation architecture towards high availability (HA). The present work describes the development of a modular C&DA system, built on the Advanced Telecommunication Computing Architecture (ATCA) standard [4], using PCI Express (PCIe) over the ATCA backplane Fabric Channel Interface, and connecting with an external PCIe host computer. The system architecture takes advantage of both ATCA and PCIe resources for HA, namely ATCA Hot Swap and the PCIe Hot Plug mechanisms, which allied to the use of redundancy contribute to improve system reliability and availability [5].

Manuscript received June 29, 2015; accepted February 15, 2016. Date of current version June 21, 2016. This work was supported in part by "Fundacão para a Ciência e Tecnologia" (FCT) through project UID/FIS/50010/2013.

M. Correia, J. Sousa, A. J. N. Batista, A. Combo, B. Santos, A. P. Rodrigues, P. F. Carvalho, B. B. Carvalho, and B. Gonçalves are with Instituto de Plasmas e Fusão Nuclear, Instituto Superior Técnico, Universidade de Lisboa, 1049-001 Lisboa, Portugal (e-mail: miguelfc@ipfn.tecnico.ulisboa.pt).

C. M. B. A. Correia is with Centro de Instrumentação, Dept. de Física, Universidade de Coimbra, 3004-516 Coimbra, Portugal.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TNS.2016.2531421

II. ATCA Features for High Availability

The ATCA form-factor offers generous layout area, power per slot and data throughput, including sub-specifications for PCIe, Serial Rapid IO or Ethernet transmission protocols. Most importantly, in the current subject, it provides resources to achieve high availability (HA), featuring redundant power modules, fans, shelf manager (ShM) units and backplane topologies. The ShM is responsible for handling these resources in order to implement fault-tolerant operation and achieve HA, monitoring and controlling the ATCA infrastructure health, cooling, power, and interfaces with the overall system manager controller.

The concept of Field Replaceable Unit (FRU) is essential to the ATCA architecture, meaning that each hardware component of the shelf can be replaced, adding modularity to the system. Furthermore, FRUs may be replaced (inserted or extracted) without a powering off the shelf, maintaining the desired level of service and increasing system availability. This protection mechanism is named Hot Swap. Fig. 1 represents a typical 14-slot ATCA shelf. Backplane connections may establish several network topologies (Star, Dual-star, and Full-mesh), depending on the shelf model, allowing to implement redundant paths for data and timing signals. In this case, the configuration is a dual-star, where logical slots 1 and 2 are hub-slots. Once installed, hub blades will provide point to point links (stars S1 and S2)

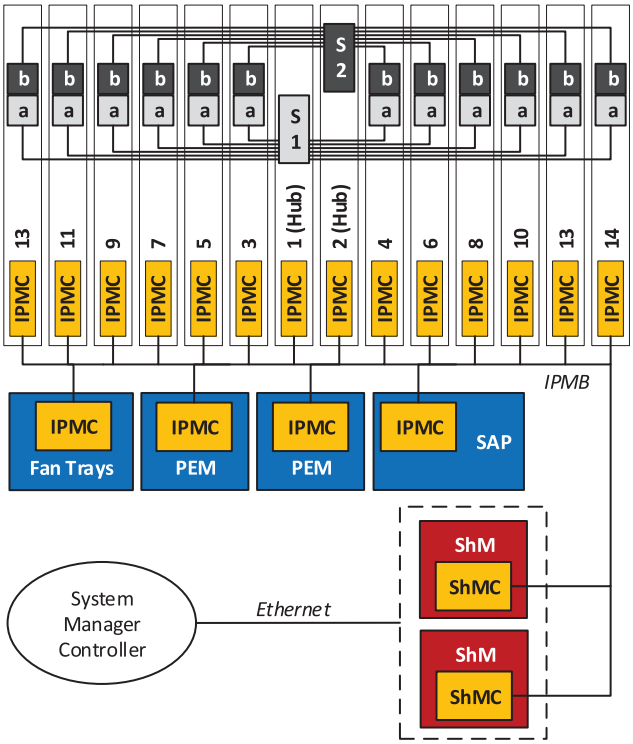

Fig. 1. Main functional blocks of a 14-slot ATCA shelf.

to node blade endpoints a and b of node slots 3 through 14. Each ATCA blade is a hot-swappable FRUs, managed by an on-board Intelligent Platform Management Controller (IPMC). Redundant Power Entry Modules (PEM), fan trays are equally hot-swappable FRUs. Also present is the Shelf Alarm Panel (SAP), which provides a serial console interface for each ShM and a set of alarm LEDs. All elements are controlled by the active ShM controller (ShMC) over the redundant I^2C-based Intelligent Platform Management Bus (IPMB) bus, using the Intelligent Platform Management Interface (IPMI) protocol.

III. System Architecture

The ATCA C&DA system is composed by two types of hardware modules. In the node slots of the shelf are digitizing units (ATCA-IOP) [6]. These modules may be configured either as ADC or DAC units, in a total of 48 input/output (IO) channels per board, with galvanic isolation up to 700 V. Digitized data are processed in a Virtex-6 Field Programmable Gate Array (FPGA) [7] and made available to the backplane PCIe network by a dual PCIe endpoint, also implemented in the FPGA. The hub slots are filled with ATCA-PTSW-AMC4 PCIe switch blades, which handle data from the node-endpoints (ATCA-IOP) and interface with an external host computer [8]. Both types of modules contain a CoreIPM OPMA2368 IPMC for local hardware management, including Hot Swap support [9]. This system was primarily developed to be used in dual-star topology, hence the dual PCIe endpoint per acquisition channel on the digitizing blades. A variety of redundancy schemes for the ATCA blades may be implemented within the dual-star setup by way of the redundancy of blades, assisted by the ATCA Hot-Swap mechanism.

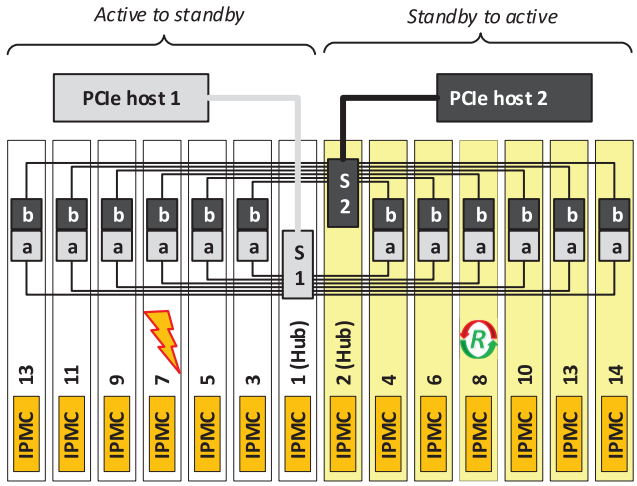

Fig. 2. 2N redundancy fail-over. Upon detection of failure of Node 7 the active half of the system enters in standby. The other half follows the reverse process. The redundant spare of Node 7 is Node 8.

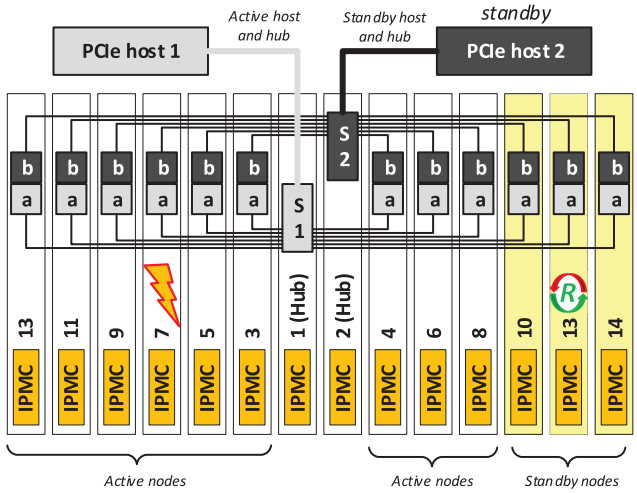

Fig. 3. N + M redundancy fail-over (M = 3). Upon detection of a failure, Node 7 enters in standby and its redundant spare (Node 13) takes over, going to active mode. All other nodes and hubs/hosts remain unchanged.

A. 2N Redundancy of all Blades

This configuration, depicted in Fig. 2, uses redundancy for every node and hub blade (and corresponding host computer). The shelf is split in two identical halves (two stars), typically operating in active-standby mode. In the case of failure of any active hardware module, the system manager controller toggles the active and standby roles of each half, allowing the replacement of the malfunctioning board without loss of service. 2N redundancy does not require any signal cabling changes for failover operation since each analogue signal cable is routed to each redundant pair of digitizer channels. This option has a higher cost per channel since every channel is duplicated. However, it ensures the highest level of availability.

B. N + M Redundancy of Node Blades

This configuration is shown in Fig. 3. It uses the same dual-star with dual hubs and host computers but there are now

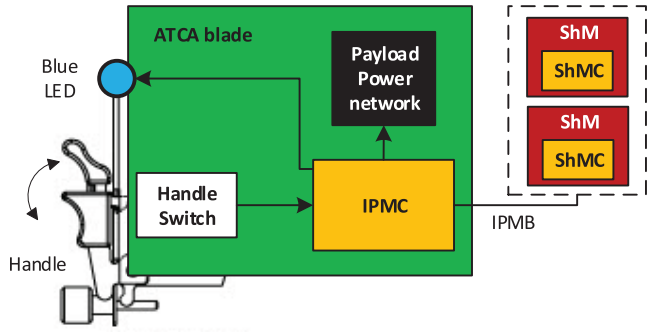

Fig. 4. ATCA Hot Swap mechanism.

TABLE I
ATCA IMPLEMENTATION OF PCIe HOT PLUG ELEMENTS

Hot Plug element	Implementation on ATCA form-factor
Indicator	Blue Led
MRL	Handle (lower)
MRL Sensor	Handle Switch
Electromechanical Interlock	(none)
Attention Button	(none)
Software User Interface	OS supported (optional)
Slot Numbering	Hardware supported (PEX 8696)
Power Controller	Hardware supported (PEX 8696)

only M spare (standby) node blades for N active node blades ($M < N$). There are more active than spare node blades so if one node has a failure, the fail-over process may require changing the cabling from one node blade to another. This may lower the availability since the operations of re-routing the analogue cabling have to be done manually. For the case of a hub blade failure there is no difference since there are still two (redundant) hubs. $N = M$ yields a 2N redundancy configuration, as in the preceding section, while $M = 0$ means no redundancy of nodes, which is the lowest cost per active channel. The $N + M$ scheme ($M < N$) is a compromise solution for cost, at the expense of availability of nodes.

IV. ATCA Hot Swap

The insertion and extraction of blades is supported by the Hot Swap features of the ATCA specification. It consists in a protection mechanism which turns on/off the payload power network of a given slot, allowing its FRU to be safely inserted or extracted, using the board's pair of Handles. The lower Handle activates a Handle Switch (HS), which is the Hot Swap sensor used to indicate to the IPMC if the Handle is open or closed. The IPMC negotiates with the ShMC permission for the requested FRU activation/deactivation processes, according to a state machine operation triggered by Hot Swap events, using the IPMI protocol. During this procedure, the IPMC informs the operator, through the FRU front panel Blue Led about the Hot Swap FRU state, with specified blinking pattern behaviours - especially important to signal if the board is ready be safely extracted. This is illustrated in Fig. 4.

V. PCI Express Hot Swap/Hot Plug

ATCA Hot Swap provides insertion/extraction of blades at the hardware level. However, at the host computer level, the operating system (OS) needs to be aware of these events, to correctly insert/remove the corresponding PCIe devices and successfully open or close related software applications. Nor the ATCA base specification or its extension "PCI Express Advanced Switching for ATCA" describes such a direct relationship between ATCA Hot Swap and PCIe [10]. The PCIe base specification indicates to support "Hot Plug and Hot Swap solutions" [11], yet further reading reveals more information regarding Hot Plug as "hot-add and hot-removal of adapters", while the term "Hot Swap" is no longer mentioned. The distinction between PCIe "Hot Plug" and "Hot Swap" is found on another author, stating that the PCIe Hot Plug is a standard "derived from revision 1.0 of the Standard Hot Plug Controller specification" for a "graceful" or "no-unexpected" methodology whereas PCIe Hot Swap does not have any standard and devices may be added or removed "without special consideration" [12]. Because ATCA Hot Swap is itself a "graceful" method (there is a warning notification to prepare the hardware for insertion/extraction of FRUs) it seems adequate to establish a relationship with the standard PCIe Hot Plug.

A. PCI Express Hot Plug

As mentioned previously, PCIe Hot Plug supports the hot-add/removal of adapters, meaning the standard PCIe form-factors, defining a set of Hot-Plug elements and respective behaviours. Other form-factors must define which and how these elements are to be implemented. The ATCA form-factor specification, as yet, does not define such implementation. The solution found was to establish a customized relationship between the ATCA native Hot Swap resources, the standard PCIe Hot Plug elements and the local hardware PCIe Hot Plug support. For the current architecture, the hardware for PCIe switching is based on the PLX PEX 8696 device [13], existing in the ATCA-PTSW-AMC4 hubs, which features PCIe Hot Plug support. Mechanical and sensor elements are provided by the ATCA Hot Swap mechanism (Handles, Handle Switch and IPMC). Table I shows which PCIe Hot Plug elements could be fulfilled by the current architecture. Attention Button and Electromechanical Interlock do not have a direct equivalent on the ATCA form-factor.

Hot Plug elements generate Hot Plug events, supported by the associated Downstream Port. For the current architecture, these are the Downstream Ports of the PEX 8696 of the ATCA-PTSW-AMC4 hubs, which link to each PCIe endpoint of the ATCA-IOP nodes. Hot Plug events, listed on Table II, are registered on the PEX 8696 Slot Control Register.

Once the Downstream Port receives a Hot Plug event, it must notify the software by a corresponding Hot Plug interrupt. Upon the reception of a specified sequence of Hot-Plug interrupts from a Downstream Port, the software layer proceeds to the insertion/removal of the corresponding device and/or taking the adequate actions to software applications using the device.

TABLE II
PCIe HOT PLUG EVENTS

Slot Events	Attention Button Pressed Power Fault Detected MRL Sensor Changed Presence Detect Changed
Command Completed Data Link Layer State Changed	

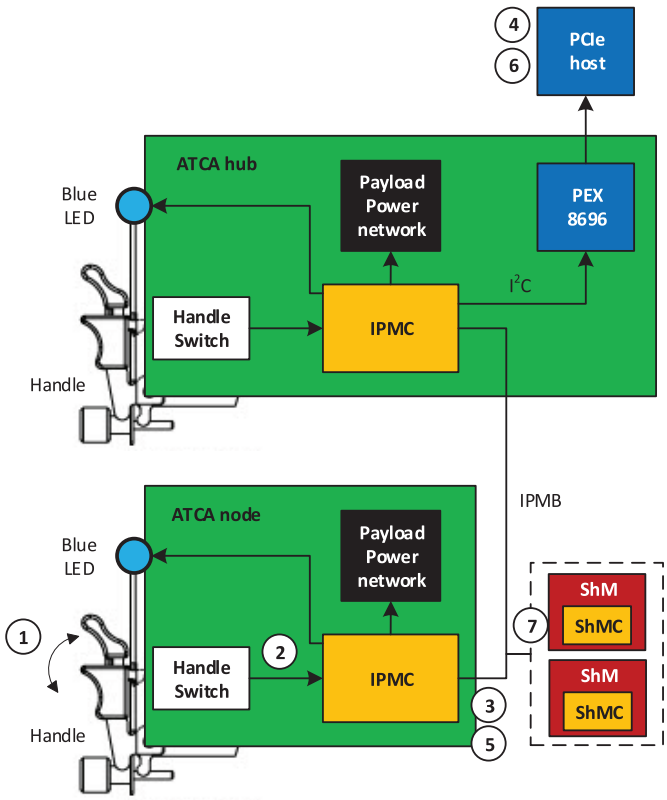

Fig. 5. PCIe Hot Plug example of node blade for the ATCA form-factor. This process occurs with the manual blade extraction, at the hardware level, supported by ATCA Hot Swap.

B. ATCA Form-factor Implementation of PCIe Hot Plug

This section describes an implementation for PCIe Hot Plug of the ATCA-IOP node blades linked to the Downstream Ports of the ATCA-PTSW-AMC4 hub, at the PEX 8696 switch. Consulting the Hot Plug process for this hardware component, two issues were found. First, the Hot-Plug event signaling interface is not physically implementable since the PCIe endpoints of the ATCA-IOP nodes do not implement these signals and there are not any sideband signals. The only PCIe signaling established between nodes and hubs on the ATCA Fabric Channel interface is in-band data. However, PEX 8696 allows Hot Plug events to be triggered by software, by I^2C access to its Slot Control register. The solution found was to use the ATCA-IOP IPMC not only to negotiate ATCA Hot Swap with the ShMC, but also to message the ATCA-PTSW-AMC4 hub IPMC the PCIe Hot Plug events, since the hub IPMC has I^2C access to the PEX 8696. With this methodology, the Hot Plug interrupts may then be generated by the PEX 8696 to the host. The second issue is that the specified Hot Plug process for the PEX 8696 uses the Attention Button event and the ATCA form-factor does not have a corresponding physical element. Therefore Attention Button event and interrupt were removed from the original process. The remaining required process events were maintained and implemented according to Tables I and II.

TABLE III
PCIe HOT PLUG EXTRACTION SLOT EVENTS

SEQUENCE OF EVENTS AND INTERRUPTS	
1.	Node blade Handle is opened.
2.	Handle Switch reports Handle open to node IPMC.
3.	Node IPMC messages hub IPMC to clear the PEX 8696 Slot Control register MRL State bit (OPEN); PEX 8696 generates MSI due to "MRL State Changed" event.
4.	PCIe Host receives the interrupt, terminating applications for the corresponding PCIe device.
5.	Node IPMC messages hub IPMC to clear the PEX 8696 Slot Control register Presence Detect State bit (OFF); PEX 8696 generates MSI due to "Presence Detect State Changed" event.
6.	PCIe Host receives the interrupt, removing the corresponding PCIe device.
7.	ShMC powers off the node slot and messages the node IPMC to signal the Indicator (Blue LED) that the node blade may be safely extracted.

C. Example of Node Blade Extraction

An example of the PCIe Hot Plug implemented for IPFN C&DA system, using the process presented in the previous section. Fig. 5 shows the setup for the extraction of an ATCA-IOP node blade, with a numbered sequence, described in Table III. The reverse process (node insertion) is analogous. In general, the node blade IPMC messages events to the hub IPMC, which generate interrupts to notify the necessary actions to be performed by the PCIe host - open/terminate applications and/or inserting/removing the corresponding PCIe devices. The node blade may then be physically inserted or removed, through ATCA Hot Swap, and the remaining nodes (not represented in the figure) keep working. In the case of a failure of the node board, a spare node takes over the failing one, using any of the redundancy schemes presented in III, preserving the service availability.

VI. RESULTS AND CONCLUSIONS

A C&DA system was conceived to exhibit HA properties, aiming to fulfil the demands of Nuclear Fusion diagnostics. The current hardware is part of the "ITER Catalog of I&C products–Fast Controllers" [14]. Preliminary tests for node insertion and extraction of node blades were successfully performed at IPFN, according to the developed procedures, described in the previous chapter. Blades were extracted in inserted from the shelf with correct activation/deactivation of respective software devices and applications, using Linux OS and in-house developed data acquisition software [15]. A demonstration of

the system was presented at the ANIMMA 2015 conference, using a 100 meter fiber optics link between the ATCA shelf and the host PC, allowing the computer to be placed farther away from the shelf and the effects of single event upsets [16]. A test plan is currently underway to test the availability, setting appropriate availability goals and create fault scenarios to assess system behaviour, starting from the redundancy scenarios herein described.

The ATCA form-factor not only contains intrinsic HA features but also provides redundancy resources that can be used to increase system availability, especially at the data transmission level, on the ATCA backplane. The current solution for PCIe device Hot Plug and ATCA Hot Swap relationship is customized for the particular application and hardware components but aims to be compatible with other types of hardware and OS. The ATCA specification could benefit from developing further definitions for the implementation of PCIe Hot Plug in future revisions, as well as further elaboration on the differences between PCIe Hot Plug and Hot Swap, as a means to standardize hardware and software procedures, increasing compatibility and facilitating the development of instrumentation systems and components.

REFERENCES

[1] B. Gonçalves *et al.*, "ITER fast plant system controller prototype based on ATCA platform," *Fus. Eng. Des.*, vol. 87, pp. 2024–2029, 2012.

[2] R. Buende *et al.*, "Reliability, availability, and quality assurance considerations for fusion components," *Fus. Eng. Des.*, vol. 29, pp. 262–285, 1995.

[3] D. van Houtte *et al.*, "ITER operational availability and fluence objectives," *Fus. Eng. Des.*, vol. 86, pp. 680–683, 2011.

[4] PICMG *PICMG 3.0 Revision 3.0 AdvancedTCA Base Specification*, Mar. 24, 2008.

[5] M. Correia *et al.*, "N + 1 redundancy on ATCA instrumentation for nuclear fusion, fusion engineering, and design," vol. 88, no. 6–8, pp. 1418–1422, 2013.

[6] A. J. N. Batista *et al.*, "ATCA/AXIe compatible board for fast control and data acquisition in nuclear fusion experiments," *Fus. Eng. Des.*, vol. 87, no. 12, pp. 2131–2135, 2012.

[7] Spartan-6 FPGA family. Xilinx Inc. [Online]. Available: http://www.xilinx.com/products/silicondevices/fpga/spartan-6/

[8] M. Correia *et al.*, "ATCA-based hardware for control and data acquisition on nuclear fusion fast control plant systems," *IEEE Trans. Nucl. Sci.*, vol. 58, no. 4, pp. 1701–1705, 2011.

[9] A. P. Rodrigues, "Intelligent platform management for fast control plant systems," *IEEE Trans. Nucl. Sci.*, vol. 58, no. 4, pp. 1733–1737, 2011.

[10] PCIMG *PICMG 3.4 R1.0 specification - PCI Express advanced switching for AdvancedTCA systems*, May 21, 2003.

[11] PCI-SIG *PCI Express Base Specification Revision 3.0*, Nov. 10, 2010.

[12] Hot-Swap in PCIe Based Systems, Application Note AN-701, IDT [Online]. Available: https://www.idt.com/document/apn/701-hot-swap-pcie-based-systems-1

[13] PEX 8696. PLX Technologies [Online]. Available: http://www.plxtech.com/products/expresslane/pex8696, 2011.

[14] P. Makijarvi, *ITER Catalog of I&C products - Fast Controllers, ITER*.

[15] P. F. Carvalho *et al.*, "PCI express hotplug implementation for ATCA based instrumentation," *Fus. Eng. Des*.

[16] A. Batista *et al.*, "SEU and Mitigation Tests in an ITER Relevant FPGA when Irradiated with Neutrons,".

[17] A. J. N. Batista *et al.*, "Control and data acquisition ATCA/AXIE board designed for high system availability and reliability of nuclear fusion experiments," *Fus. Eng. Des.*, vol. 88, pp. 1332–1337, 2013.