ISBN: 978-3-95450-214-1 ISSN: 2673-5490 doi:10.18429/JACoW-IPAC2021-THPAB287

PROVIDING COMPUTING POWER FOR HIGH LEVEL CONTROLLERS IN MicroTCA-BASED LLRF SYSTEMS VIA PCI EXPRESS EXTENSION*

P. Nonn†[1], A. Eichler, S. Pfeiffer, H. Schlarb, J. Timm
Deutsches Elektronen-Synchrotron DESY, Hamburg, Germany
[1]also at MicroTCA Technology Lab, Hamburg, Germany

Abstract

The MicroTCA.4 standard for crate architecture allows to use a PCI Express Generation 3 bus for data transmission between the modules in a crate. This enables a software, running on a CPU module in the crate, to directly access the data, processed, i.e., by a Field Programmable Gate Array (FPGA) on another module in the same crate. The CPU performance is limited, due to the limit of cooling capacity specified for each slot, by the MicroTCA.4 standard. This limitation can be circumvented, by extending the PCI Express bus from the crate to a high performance computer. This is already practised in Low Level Radio Frequency (LLRF) control systems. This article will discuss the advantages and disadvantages of this feature with a special focus on the use for high level control algorithms.

THE MicroTCA STANDARD

The Micro Telecommunications Computing Architecture (MicroTCA) standard [1], maintained by the *PCI Industrial Computer Manufacturers Group* (PICMG) provides guidelines and requirements for the design of reliable, remote maintainable computing infrastructures. It contains multiple sub-standards for different use-cases, with MicroTCA.4 being specialized towards scientific applications. Commonly, these are realized as 19" crates. The computing infrastructure of a MicroTCA.4 crate consists of a shelf manager, called *MicroTCA Carrier Hub* (MCH) and a backplane. The devices, connected to the backplane, have to adhere to the *Advanced Mezzanine Card* (AMC) standard [2], also maintained by PICMG, and are hence referred to as AMCs.

Processing Power Limitation in a MicroTCA.4 Crate

The maximal thermal power load per slot in a crate is limited to 80 W, due to cooling constraints (see Section 5.6 in [1]). Thus the Thermal Design Power (TDP) of a CPU, installed on an AMC, has to be well below 80 W, as other devices on the board might also produce some thermal load. As Fig. 1 shows, the processing power of a CPU, represented by its clock frequency and the number of cores, is limited by the TDP.

This limitation can be circumvented by installing multiple CPU modules into a MicroTCA.4 crate. The drawback of this work-around would be increased cost and the reduction of slots, available in a crate. Additionally, the modules, connected to different CPUs, can no longer communicate with each other via PCI Express.

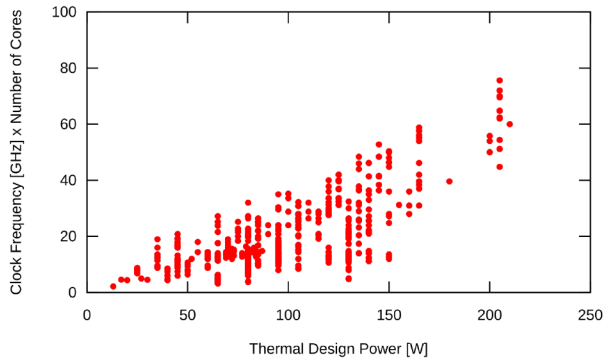

Figure 1: CPU Performance, represented by the product of clock frequency and number of cores, versus thermal design power (TDP) of various CPUs from the Intel Xeon family.

THE PCI EXPRESS BUS

For almost 20 years, the PCI Express standard is used to connect a computers CPU to its memory, peripheral devices (i.e. graphics adapter or Ethernet card) and increasingly also storage (i.e. Solid-State Drives (SSDs) via M.2). In this span of time the standard has developed multiple revisions, called generations. The overall throughput roughly doubled with each generation. The current generation (Gen5) has a maximal throughput of 3.928 GB/s per lane. A PCI Express lane is a full duplex serial connection between two devices, with usual numbers of lanes being 1, 4, 8 and 16.

CPU(s) and RAM are connected to each other and peripherals through a device called PCI Express Root Complex. It provides a logical separation between the interconnection of processor cores and RAM and the PCI Express bus with the connected peripheral devices. It is possible for a PCI Express device to access the RAM, without generating load on the CPU, which is called Direct Memory Access (DMA).

PCI Express in a MicroTCA.4 Crate

MicroTCA.4 in its current revision supports PCI Express links up to generation 3 (Gen3), with a throughput of up to 0.985 GB/s per lane. The number of lanes available to a module depends on the topology of the backplane. In the widely used dual-star backplane topology, each module has up to 4 lanes available. Thus the maximum data throughput for a module in a MicroTCA.4 crate with such a backplane is 3.94 GB/s.

* Supported by MicroTCA Techlab.
† patrick.nonn@desy.de

ISBN: 978-3-95450-214-1 ISSN: 2673-5490 doi:10.18429/JACoW-IPAC2021-THPAB287

The Root Complex in a MicroTCA.4 crate is usually provided by a CPU module. A configurable PCI Express switch, which is part of the crate manager, called MicroTCA Carrier Hub (MCH), allows to allocate slots in the crate to a slot, holding the CPU module. Figure 2 shows the PCI Express bus in a standard PC, a MicroTCA.4 crate with CPU module and PCIe switch and a crate with PCIe extension to a PC. The latter will be introduced below.

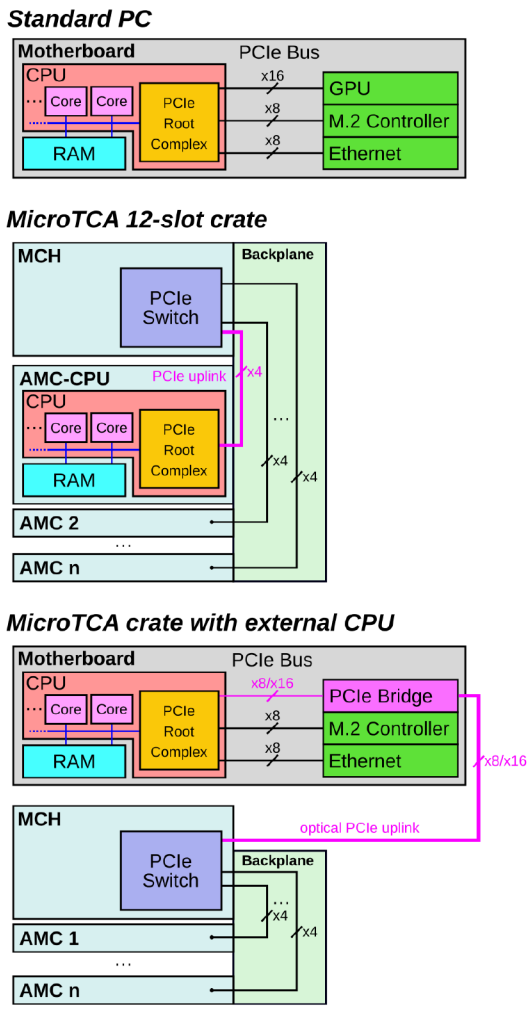

Figure 2: PCI Express bus architectures for standard PC motherboard, MicroTCA.4 crate with CPU module and MicroTCA.4 crate with external CPU.

PCI Express Bus Extension

To increase the available CPU power for a MicroTCA.4 crate, the CPU module can be replaced by a high performance computer, connected to the crates PCI Express bus. Figure 2 shows such an "extended CPU". The PCI Express bridge is located on an extension card, put into an appropriate slot on the Computers motherboard. The PCI Express uplink between Computer and MicroTCA.4 crate can be 8 or 16 lanes wide, and is usually established via an optical connection. In theory, multiple MicroTCA.4 crates could be connected to the same external CPU, but this has not been tested, yet.

Such an external CPU is tested for the new LLRF control system for the light ion injector LILAC of the NICA project at the Joint Institute for Nuclear Research [3].

HIGH LEVEL CONTROLLERS FOR LLRF

LLRF control systems usually control either one cavity [4] or, via vector-sum control, a group of same-type cavities [5]. In both cases LLRF controllers act isolated from each other and other devices. Thus LLRF controllers are not able to adjust themselves, to reflect changes in the state of the accelerator as a whole, i.e. when the beam parameters are changed. The interactions of devices along the beamline have to be learned by the operator. A High Level Controller (HLC), acting on multiple LLRF controllers, could be implemented, to, at least partially, automate this process.

A Model-based Predictive Controller (MPC) monitors a multitude of parameters and uses a model to predict the output, that would steer the system towards a desired state. Its ability to process large numbers of parameters makes it a candidate for HLC.

Computing Requirements of Model-Based Controllers

Model-based controllers are already used in LLRF. For example the Iterative Learning Controller (ILC) used at XFEL to optimize the feed forward. This "learning feed forward" solves an optimisation problem each iteration, resulting in a high CPU load. Thus it uses the Fast-Norm-Optimal ILC, to fit the hardware requirements, set by the AMC-CPU [6].

An HLC, using a model-based controller, like an MPC, would require a much more complex model, for which an optimisation problem needs to be solved. This raises the requirements for the hardware to execute the HLC to a level, beyond what an AMC-CPU could provide.

IMPLEMENTATION OF A HIGH LEVEL CONTROLLER

High Level Controllers, like MPCs, are usually implemented as software, running on a high performance computer, i.e. in a dedicated computing center, rather than firmware, running on an FPGA, as it is common for LLRF. In the case of a High Level RF controller for a MicroTCA.4 based LLRF, it would communicate with the software frontend of the LLRF controller over a dedicated, closed Local Area Network (LAN). Figure 3 depicts such a setup, where the LLRF controller software is running on an AMC-CPU interfacing with the controller firmware over PCI Express. The necessary infrastructure for such an implementation of an HLC are usually already present. Most facilities, operating a particle accelerator, also have the resources to provide high performance computing hardware. It is also common, that the LLRF hardware, together with other devices with an Ethernet interface (i.e magnet power supplies) share a dedicated, closed LAN (machine net). The implementation of an HPC can thus be cost effective.

ISBN: 978-3-95450-214-1 ISSN: 2673-5490 doi:10.18429/JACoW-IPAC2021-THPAB287

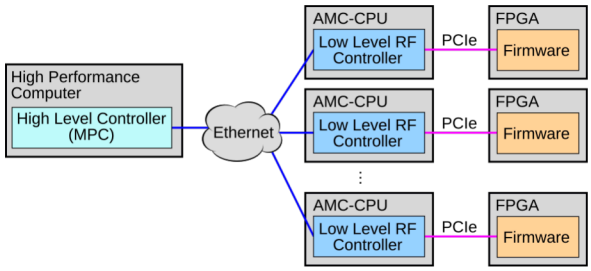

Figure 3: Communication infrastructure for a High Level Controller, running on a remote computer.

On the down side, the data throughput of even a dedicated Ethernet is limited to about 0.125 GB/s for the widely used Gigabit Ethernet, which is about $\frac{1}{8}$ of a x1 PCI Express link. Other Ethernet standards, like 10 Gigabit Ethernet increase the data throughput, but require a compatible infrastructure.

Alternative: External CPU

The HLC can also run on a high performance CPU, which acts as an external CPU for a MicroTCA.4 based LLRF control system, as described above. Figure 4 shows how such an HLC would interact with the LLRF. Both, LLRF controllers and HLC, would run on the same CPU. This increases the synchronicity of the data, provided by the LLRF controller software. It also allows the HLC to communicate directly with the firmware due to access to the PCI Express endpoints on the external CPU. This could allow the HLC to function independent from the LLRF controller software.

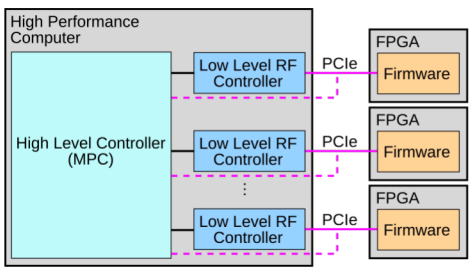

Figure 4: Integration of an HLC into an MicroTCA.4 external CPU, alongside LLRF control software.

DISCUSSION

The development of model-based predictive controllers to automate the RF control for for multiple cavities, up to and including whole accelerators, is still in its early stages. Hence there is little experience regarding the implementation of this kind of controller. Both implementations, presented in this article, have their advantages and disadvantages.

Scalability vs. Integration

The implementation of an MPC in a remote computing center, as shown in Fig. 3, is easier scalable. If the model gets to big to be computed with the desired speed, it can be split up (Distributed MPC), without the need to change the hardware. It is also possible to add, remove or exchange connected devices. On the other hand, because of that mutability, the model for an MPC might have to be re-trained more frequently, than in a more integrated system, as described in Fig. 4.

An MPC, implemented in an external CPU, is limited to the LLRF controllers connected to the external CPU. While it is possible to connect multiple crates to the same external CPU (see above), because each crate needs its own slot, the PCI Express infrastructure on the motherboard of the external CPU in combination with the number of slots per crate, limits the number of possible LLRF controllers, controlled by the same HLC. Additionally the distance between the crate(s) and the external CPU is limited to below 100 m. On the upside, the MPC could access parameters directly via PCI Express, without relying on the operation of the LLRF controller software. This could make the operation of the HLC more robust. For instance could an LLRF controller be restarted, without the need to halt the HLC, first.

SUMMARY

Extending the PCI Express bus of an MicroTCA.4 crate to a high performance computer allows to circumvent the power restrictions for AMCs, which come with the MicroTCA.4 standard. This is used to increase the number of cavities, which can be controlled individually (single cavity control) with the modules fitting into one 19-inch crate. In addition the CPU performance can be provided to a Model-based Predictive Controller, which acts as a High Level Controller for all cavities. But such an implementation of an High Level Controller also limits the scalability of the MPC.

REFERENCES

[1] PICMG, "Micro Telecommunications Computing Architecture Base Specification R1.0", Jul. 2006.

[2] PICMG, "Advanced Mezzanine Card Base Specification R1.0", Jul. 2006.

[3] P. Nonn, Ç. Gümü, C. K. Kampmeyer, H. Schlarb, Ch. Schmidt, and T. Walter, "MicroTCA Based LLRF Control Systems for TARLA and NICA", in *Proc. 10th Int. Particle Accelerator Conf. (IPAC'19)*, Melbourne, Australia, May 2019, pp. 4089-4091. doi:10.18429/JACoW-IPAC2019-THPRB115

[4] M. Steinhorst *et al.*, "Low Level RF ERL Experience at the S-DALINAC", in *Proc. 63rd Advanced ICFA Beam Dynamics Workshop on Energy Recovery Linacs (ERL'19)*, Berlin, Germany, Sep. 2019, pp. 52-55. doi:10.18429/JACoW-ERL2019-TUCOZBS05

[5] J. Branlard *et al.*, "The European XFEL LLRF System", in *Proc. 3rd Int. Particle Accelerator Conf. (IPAC'12)*, New Orleans, LA, USA, May 2012, paper MOOAC01, pp. 55-57.

[6] S. Kirchhoff, C. Schmidt, G. Lichtenberg, and H. Werner, "An Iterative Learning Algorithm for Control of an Accelerator based Free Electron Laser", in *Proc. 47th IEEE Conf. on Decision and Control*, Cancun, Mexico, Dec. 2008, pp. 3032-3037. doi:10.1109/CDC.2008.4739064