



European Coordination for Accelerator Research and Development

## PUBLICATION

# SURVEY OF COMMUNICATION LINKS FOR ATCA IN PHYSICS

Makowski, D (DMCS Lodz) *et al*

15 November 2010

The research leading to these results has received funding from the European Commission under the FP7 Research Infrastructures project EuCARD, grant agreement no. 227579.

This work is part of EuCARD Work Package **10: SC RF technology for higher intensity proton accelerators and higher energy electron linacs.**

The electronic version of this EuCARD Publication is available via the EuCARD web site <<http://cern.ch/eucard>> or on the CERN Document Server at the following URL :  
<<http://cdsweb.cern.ch/record/1306818>>

## SURVEY OF COMMUNICATION LINKS FOR ATCA IN PHYSICS

D. Makowski\*, G. Jablonski, A. Piotrowski, W. Cichalewski, W. Jalmuzna, DMCS, Lodz, Poland  
W. Koprek, S. Simrock, DESY, Hamburg, Germany

### Abstract

Modern machines used in high energy physics require sophisticated and complex control systems. The complex systems are usually built as distributed systems. Therefore, the connectivity and communication links between distributed subsystems play a crucial role in the control system. The Advanced Telecommunication Computing Architecture (ATCA) and Advanced Mezzanine Card (AMC) standards have attracted the attention of physics community because they offer various types of data communication channels with high bandwidth, redundancy, high reliability and availability. The standards allow using different types of communication interfaces like PCIe, Gigabit Ethernet, RapidIO. In real-time systems the data transmission latency is also important. The acquisition of real-time data from hundreds of analogue channels is required for the Low Level Radio Frequency (LLRF) controller of XFEL (X-ray Free Electron Laser) accelerator. The paper presents survey of the communication interfaces of the LLRF controller for XFEL. The discussion includes the properties of interfaces provided by ATCA and AMC standards and summarizes requirements for the data links and protocols required by LLRF controller.

### COMMUNICATION INTERFACES FOR HIGH-ENERGY PHYSICS

High Energy Physics Experiments require complex and reliable control systems. Data Acquisition (DAQ) systems usually demand transmission of large amounts of data whereas low and constant latency is more important in feedback, real-time control systems. The Advanced Telecommunication Computing Architecture (ATCA) and Advanced Mezzanine Card (AMC) standards offers various interfaces: Ethernet on Base Interface (10/100/1000 BASE-T) and Gb Ethernet, Infiniband, StarFabric, PCI Express on Fabric Interface [1]. The ATCA standards offers mainly switching protocols that enables gigabit per second throughput. The comparison of theoretical throughputs of ATCA links available on the backplane is summarized in Table 1.

The interfaces offered by ATCA standard have latency higher than a few  $\mu$ s. However, ATCA allows to use user-defines interfaces with lower latencies that can be implemented using LVDS standard or pure serial gigabit transceivers [2].

\* dmakow@dms.p.lodz.pl

### LLRF CONTROL SYSTEM OF XFEL ACCELERATOR

The LLRF system of XFEL (X-ray Free Electron Laser) needs to control almost 1000 accelerating cavities. The LLRF controller of each RF station, supervising 32 cavities, is connected to other accelerator components with a significant number of analogue and digital signals, e.g. 96 analogue cavity signals, 32 analogue and 32 digital signals for fast and slow piezo tuners, 10 digital interlock signals [2]. Therefore, the LLRF control system designed with application of ATCA standard can use various interfaces for data transmission with different throughputs and latencies. The complexity of LLRF system requires implementation of three different types of communication links: interfaces with latency lower than hundreds of ns used for the main LLRF controller loop, interfaces with gigabit throughput for DAQ system and transmission channel for controls data. In case of control data throughput is not so important because usually control data is sent in small packets, but latency is important especially for automation servers which must react between subsequent RF pulses [2].

### COMPARISON OF SELECTED INTERFACES

#### PCI Express

PCI Express (Peripheral Component Interconnect Express, PCIe) has been designed as a computer expansion card standard [3]. The PCIe bus can be considered as a high-speed serial replacement of the older (parallel) PCI/PCI-X bus. At the software-level, PCIe preserves compatibility with PCI. It requires a Root Complex (RC) for bus management and configuration and has a host-centric character, although direct point-to-point communication between the peripherals is also possible. The interface between two PCIe peers consists of one to up to 16 lanes, and is logically asymmetrical, with a distinction between the upstream (towards the root complex) and downstream (towards the peripheral) direction. To connect more than one peripheral to the root complex a switch is required. If the common reference clock is distributed among all peers in the system it is possible to use the spread-spectrum clocking to reduce electromagnetic interference. Currently the most widely used version of the standard is version 1.x, version 2.0 is being introduced to the market. In version 1.x of the standard each lane has the raw bit rate amounting to 2.5 Gbps using the 8/10 encoding, therefore the net

Table 1: Throughputs Comparison of Interfaces Offered by ATCA Standard

Interface	Base bit rate		Data bit rate	
	Single lane	Quad lane	Single lane	Quad lane
Base Gb Ethernet (1000BASE-T)	—	1.5 Gbps	—	1 Gbps
Fabric Gb Ethernet (1000BASE-BX)	1.25 Gbps	—	1 Gbps	—
Fabric 10 Gb Ethernet (1000BASE-BX4)	—	12.5 Gbps	—	10 Gbps
Infiniband	2.5 Gbps	10 Gbps	2 Gbps	8 Gbps
StarFabric	622 Mbps	2.5 Gbps	450 Mbps	1.8 Gbps
PCI Express	2.5 Gbps	10 Gbps	2 Gbps	8 Gbps

transmission rate is equal to 2 Gbps/lane, see Table 1. In the second version of the standard the signaling rate has been doubled. It is possible to use more than one switch in the system in the hierarchical manner. In general, only one root complex in the system is allowed. Some switches allow connecting more than one root complex to create a bus hierarchy visible from its upstream port as a single device or to enable the switchover on failure to increase the reliability of the system.

### Gb Ethernet

Ethernet is one of the interfaces offered by ATCA standard. A carrier blade with optional AMCs forms an ATCA subsystem which communicates with other blades over ATCA backplane using Ethernet. An Ethernet switch sitting in logical slot 1 or 2 of ATCA shelf enables connections between all blades in the shelf and external world. The topology of ATCA backplane gives possibility for implementation of two independent Ethernet networks: 1 Gb on Base Interface and/or 1 Gb/10 Gb on Fabric Interface. Each blade in the shelf may have interface to either Base or Fabric Interface or both. Connection of AMCs to Ethernet on the backplane requires implementation of local Ethernet switch on the blade. Ethernet switches on blades and in slot 1 or 2 separate internal traffic between AMCs and blades within one shelf. This feature significantly reduces protocol latency introduced by heavy-loaded switches.

The low latency links are essential in LLRF real-time systems where data transfer must be done before next will begin and with constant latency, e.g. for feedback in LLRF systems. The low latency links can be implemented in ATCA systems in different manner due to flexibility of PICMG 3.0 specification. Each port in 4-port channel in the ATCA backplane can be configured to different protocol. Besides protocols defined in PICMG 3.X, it is possible to define custom protocol. Implementation of custom protocols in point-to-point connections allows simplification of the protocol and thus reduction of the latency. Aggregation of four ports for low latency link gives 10 Gbs transfer between two carrier blades. This configuration is especially useful in ATCA systems with full-mesh backplane where every two carrier blades have dedicated communication channel. Implementation of low latency links

for communication between carrier blade and AMC can be done on Extended Options Region using ports from 12 to 20.

## EXPERIMENTAL RESULTS

### PCI Express Latency Measurements

The PCIe latency has been measured using the PC equipped with ASUS P5PL2 motherboard together with:

1. ML506 Xilinx Evaluation Board with the x1 interface put into the PCIe slot,
2. ML506 Xilinx Evaluation Board connected to the PEX8532RDK switch evaluation board put into the x16 PCIe slot,
3. Custom-made ATCA Carrier Board containing the PEX8532 switch and three TEWS TAMC900 AMCs connected using the x4 PCIe cable [4]. The topology of this connection is presented in Figure 1.

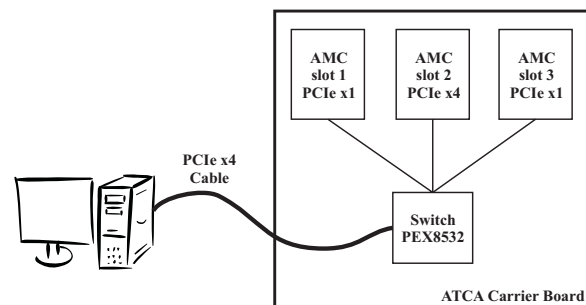


Figure 1: A block diagram of measurement system.

All the devices connected to the PCIe bus contain the Xilinx Virtex 5 FPGAs, programmed with the standard PCIe endpoint. The duration of a single read from these devices has been measured. For the first setup the round-trip latency of 1.5  $\mu$ s has been obtained. For the second setup the latency amounted to 2.1  $\mu$ s. In case 3 a series of measurements for x1 and x4 connections on an idle PCIe bus and with background DMA transfers from one, two and three AMC modules have been performed. The results of these measurements are presented in Tables 2 and 3. The results

show, that the latency in one direction through the single switch is of order of 300 ns. The background DMA transfer can greatly increase this latency up to a few microseconds, what is especially visible in case of x1 link.

Table 2: PCIe Read Latency [ $\mu s$ ]

		No DMA	DMA slot 1	DMA slot 2	DMA slot 1, 2	DMA slot 1, 2, 3
PCIe x1	min	1.7	1.7	2.3	2.1	2.0
	max	1.9	2.1	2.5	9.0	9.5
	avg	1.9	2.0	2.4	8.1	8.3
PCIe x4	min	1.6	1.7	2.1	2.1	2.3
	max	1.7	1.7	2.7	3.9	4.9
	avg	1.6	1.8	2.4	3.4	3.7

Table 3: PCIe DMA Throughput

	DMA slot 1	DMA slot 2	DMA slot 1, 2	DMA slot 1, 2, 3
DMA x1	155 MB/s	-	142 MB/s	126 MB/s
DMA x4	-	330 MB/s	284 MB/s	256 MB/s

### Gb Ethernet Latency Measurements

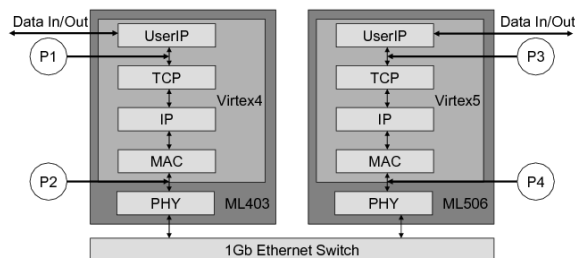


Figure 2: A block diagram of measurement system for Ethernet communication.

The latency test was performed using Ethernet component written in VHDL and implemented in Xilinx FPGA — so-called SiTCP (Silicon TCP) [5]. The component includes MAC, TCP/UDP and IP layers that are implemented in hardware. For the test only TCP protocol was used. The Figure 2 presents measurement setup. There were two evaluation boards used for the tests ML403 and ML506 — both equipped with Virtex FPGA chips. The data was sent through EIA RS 232 in ML403 and read out from ML506. There were six points from P1 to P6 where latency was measured. The results of measurement are collected in Table 4.

Table 4 presents the latency of VHDL component is comparable with latency of physical components. The minimum latency between input of TCP on one board and TCP output on the other is approximately 6  $\mu s$ .

Table 4: Ethernet Latencies with SiTCP

Points	Latency [ $\mu s$ ]	Comment
P1 – P6	6	From TPC input to TCP output
P2 – P5	2.8	PHY latency including switches
P3 – P4	2.5	Ethernet switch latency
P2 – P3	0.15	PHY latency
P1 – P2	1.2	From TCP input to MAC output

## SUMMARY

The interfaces offered by ATCA standard allows to obtain high throughput between carrier blades installed in ATCA shelf from 1 Gbps to 10 Gbps. The throughput depends on the interface type and links configuration (e.g. x1, x2, x4). However, the latency of the communication links can significantly vary. The PCIe allows to send data with latency lower than Gb Ethernet, especially when Ethernet stack is implemented in software. The latency for the hardware stack was in range of a few  $\mu s$ . The PCIe interface is usually used to design real-time systems and therefore it assures more predictable latency than Gb Ethernet. In real system the latency of Gigabit Ethernet can be only longer and it strongly depends on the network configuration such as number of switches, type of switches, throughput of up-links. The latency of the SiTCP should be also investigated under heavy load.

## ACKNOWLEDGMENT

The research leading to these results has received funding from the European Commission under the EuCARD FP7 Research Infrastructures grant agreement no. 227579. The authors are scholarships holders of project entitled "Innovative education ..." supported by European Social Fund.

## REFERENCES

- [1] PICMG. AdvancedTCA Base Specification. January 2003, PICMG 3.0.
- [2] Dariusz Makowski, Waldemar Koprek, Tomasz Jezynski, Adam Piotrowski, Grzegorz Jablonski, Wojciech Jalmuzna, and Stefan Simrock. Interfaces and communication protocols in ATCA-based LLRF control systems. In *Nuclear Science Symposium Conference Record, 2008. NSS '08. IEEE*, pages 32–37, October 2008.
- [3] PICMG. PCI Express Base Specification 1.1. 2005.
- [4] PICMG. PCI Express External Cabling 1.0 Specification. 2007.
- [5] T. Uchida. Hardware-based TCP processor for gigabit ethernet. *IEEE Transactions on Nuclear Science*, 55(3):1631–1637, June 2008.