Extending PCI-Express in MicroTCA Platforms

Whitepaper

Joey Maitra of Magma
&
Tony Romero of Performance Technologies

Introduction:

The introduction of MicroTCA® platforms has opened the door for AdvancedMC (AMC) modules to be deployed in numerous applications, and in many more industries such as telecommunications, wireless, aerospace/defense, medical, and industrial automation. AMC modules constitute the next generation of 'carrier grade' communication equipment. This series of specifications incorporate the latest trends in high speed interconnect technologies, next generation processors and improved Reliability, Availability and Serviceability (RAS).

As the AMC ecosystem has matured, now more than 34 vendors offer a wide base of AMC modules with functionality ranging from processors, storage, video, networking IO, FPGAs, DSPs, and more. However, with the new popularity of MicroTCA, some embedded designers are interested in configurations that include AMC cards operating alongside legacy PCI-Express cards or PCI cards. This opens the door even wider for updating processor power or IO performance with AMC cards, while leveraging older cards that have unique software development invested or offer functions not yet implemented in AMC modules. The solution is to extend the PCI-Express bus IO outside of the MicroTCA system to a platform that supports standard PCI-Express and PCI slots.

Alternatively, an external PCI-Express interface also allows a MicroTCA system to be an IO adjunct box to a Server Blade System, such that an application can leverage the large processing power of blade servers and interconnect to existing telecommunications IO such as T1/E1/J1 or OC-3.

This whitepaper discusses approaches to address these needs and significantly reduce time to market and development costs.

Scenario No. 1: IO Expansion

Embedded telecomm designers are constantly seeking to capitalize on COTS-based products and are gravitating to the performance, scalability, and flexibility of MicroTCA and AMCs, but they may have requirements for functionality that are not yet available in the AMC form factor. Until now, the only solution would be to design their own AMC solution to support their specific functionality, or to try to program an FPGA-based AMC solution. However, wouldn't it be nice if these designers could tap into the vast reservoir of COTS products offered in the PCI-X and PCI Express bus architectures? Examples of COTS card that are needed are Video Acquisition, Storage Controller, high end Single Board Computer (SBC)s, or specialized IO, to name a few.

The solution to the problem entails mechanisms that allow for the MicroTCA system to integrate with a platform that supports PCI-X and PCI-Express. This technique involves extending the PCI-Express fabric from the MicroTCA chassis, acting as a Root Complex, to an expansion box supporting either or both PCI-Express slots and PCI. PCI Express technology is rapidly being deployed in carrier fabrics. This interconnect technology is becoming practically ubiquitous in the embedded world.

MAGMA
MISSION TECHNOLOGY GROUP, INC.

PERFORMANCE TECHNOLOGIES

PCI Express technology is a low cost, highly scalable, switched, point-to-point, serial I/O interconnect that maintains complete software compatibility with PCI. The transfer rate is 2.5Gbs for 1st Gen and 5.0Gbs for 2nd Gen per lane per direction. Adding lanes to the link proportionately scales performance.

Magma's AMC Adaptor provides a new paradigm for connectivity to MicroTCA systems and by extension to ATCA shelf equipment. This method uses the AMC module as an enabler for connection to expansion chassis that provide slots for PCI, PCI-X and PCI Express. This then functions as the I/O sub-system consisting of the various I/O cards within the chassis. These expansion chassis are all rack-mountable with their own power supply and cooling resources. Many of them have space and power resources for multiple hard drives to be installed within the chassis as well. They come in various form factors ranging from 1U to 4U. Figure 1.0 depicts such a solution.

MAGMA
MISSION TECHNOLOGY GROUP, INC.

PERFORMANCE TECHNOLOGIES

Figure 1.0

Scenario No. 2: Extended Storage Options

Telecommunication companies must deal with a growing number of increasingly demanding application environments, including IMS, Security and Transport Control. These have put demands on equipment vendors for vast amounts of reliable storage facilities. While AMC modules exist that offer as much as 500GB of SATA storage per drive and offer sufficient storage capacity for numerous applications, there are storage-intensive applications that require Terabytes of capacity along with RAID capability.

With the proliferation of the Internet, the Telecom infrastructure is in need of storage that is used for data base applications that require extremely fast access time. In many instances, the latency involved in accessing data from external storage from the SAN (Storage Area Network) is not an acceptable solution. This is because in database applications, access times to storage resources translate into productivity which produces higher revenue streams contributing to the bottom line of the organization.

The Storage Area Network (SAN) consists of various categories of storage devices attached to a network appearing to the server as direct attached devices. There are two predominant technologies today comprising SAN; they are Fibre and iSCSI. Both of these technologies used layered protocols. With Fibre it is the Fibre Channel Protocol (FCP) which is a transport protocol similar to TCP used in IP networks. The maximum raw data rate can extend up to 20 Gigabits/sec. Similarly the iSCSI has its own protocol that runs embedded within the TCP/IP protocol. The maximum raw data rate is 480Mbits/sec for iSCSI. In both cases the data throughput is many-fold less than the data rate because of the software overhead.

The overhead associated with these protocols are costly and responsible for undesirable latencies on accessing attached storage devices; and the data rates are limited by the maximum bandwidth supported by the respective interconnect technology. On the other hand, PCI Express technology allows for a maximum of 256 nodes of attached devices on the network and supports a scalable data rate of up to 80 Gigabits/sec. The packet transport layer protocol specified in the PCI Express specification makes higher level protocols unnecessary thereby eliminating costly software overhead. Also this type of network topography capitalizes on all of the attributes of PCI bus architecture mainly the plug-n-play aspect.

Figure 2.0 is one example of how the AMC Adaptor from Magma can solve this problem.

Figure 2.0

Scenario No. 3: I/O Virtualization solutions

ATCA Blade Servers as well as Servers in general are increasingly faced with the situation as to what do with their idle CPU cycles. This situation is the result of ever increasing powerful processor architectures as well as with the advent of multi-core processor module. In many of these instances the CPU idle time is the direct result of the system being I/O bound.

The Telecom equipment providers, on the other hand, are in need of ever increasing compute power to address complex applications that need massive CPU resources. These applications are imposed on the equipment vendors by demanding Telecom service providers. Service providers are reluctant to invest massive amounts to upgrade their equipment to use an array of ATCA Blade servers. Their obvious preference would be to upgrade functionalities and feature set with some sort of a bridging solution.

Also, the aggressive deployment of Blade technology has put demands on the Telecom service providers to integrate Blade Servers to Telecom-based I/O. However, as mentioned, there is no graceful solution that exists on connecting Blade Servers to Telecom equipment allowing for the unleashing of the Blade power for telecommunication applications.

In today's economy, the tightening of capital spending for telecom infrastructure by organizations is causing Telecom equipment vendors to face the dual challenge of cutting costs and continuing to provide innovative newer solutions. One of most practical ways to accomplish this is to use I/O Virtualization technology. This allows multiple MicroTCA chassis to share the same I/O hardware resources and their use is transparent to each of the chassis and their Root Complex processors. This is supported with PCIe Multi Root I/O Virtualization techniques which are now an industry standard. Magma's flexible PCI Express chassis fully supports such applications. Figure 3.1 depicts such a scenario.

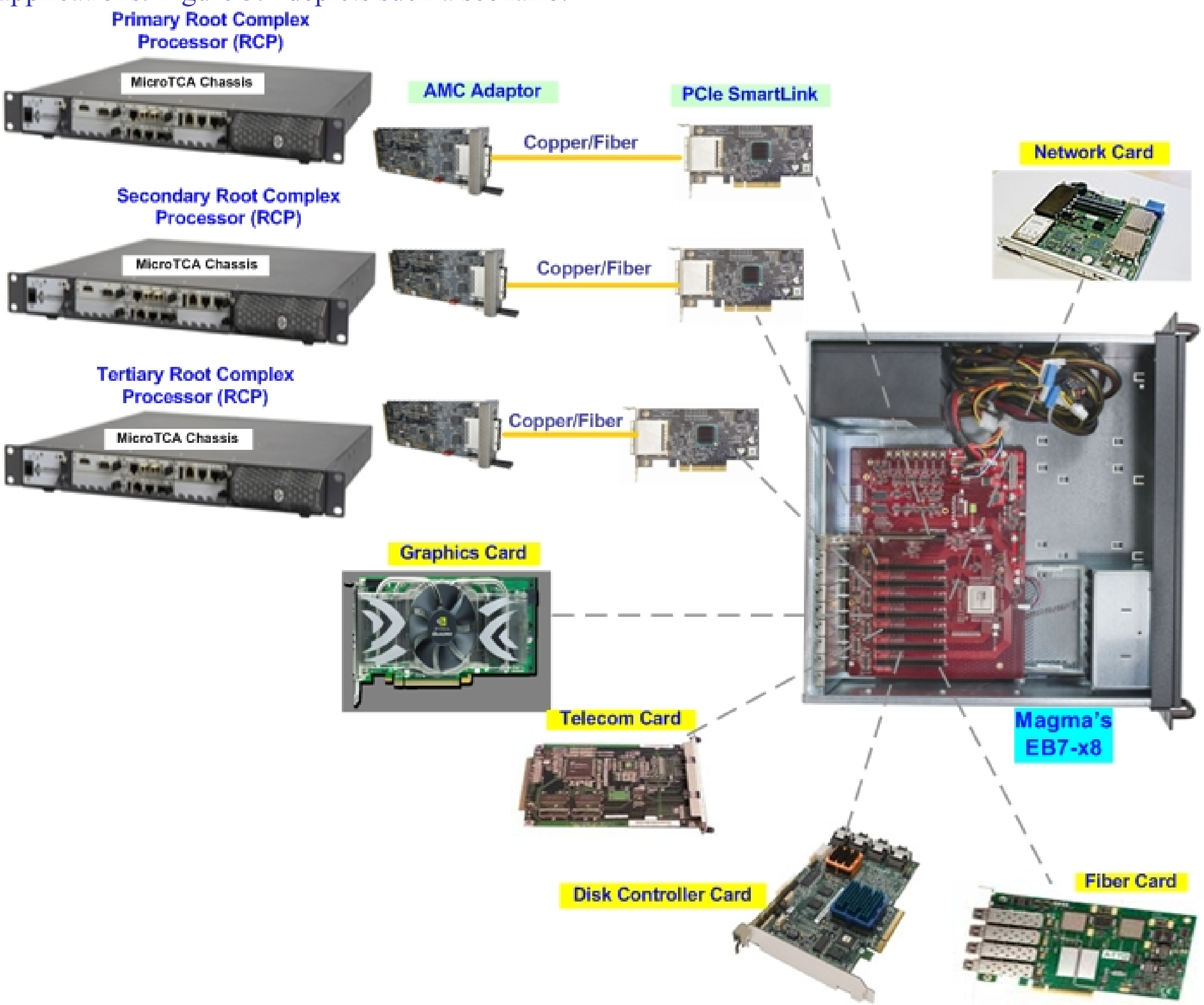

Figure 3.0

In addition, multiple MTCA systems with processors with PCI-Express adapters can communicate to multiple MTCA systems with AMC IO using this virtualization technique. Figure 3.1 depicts such a scenario.

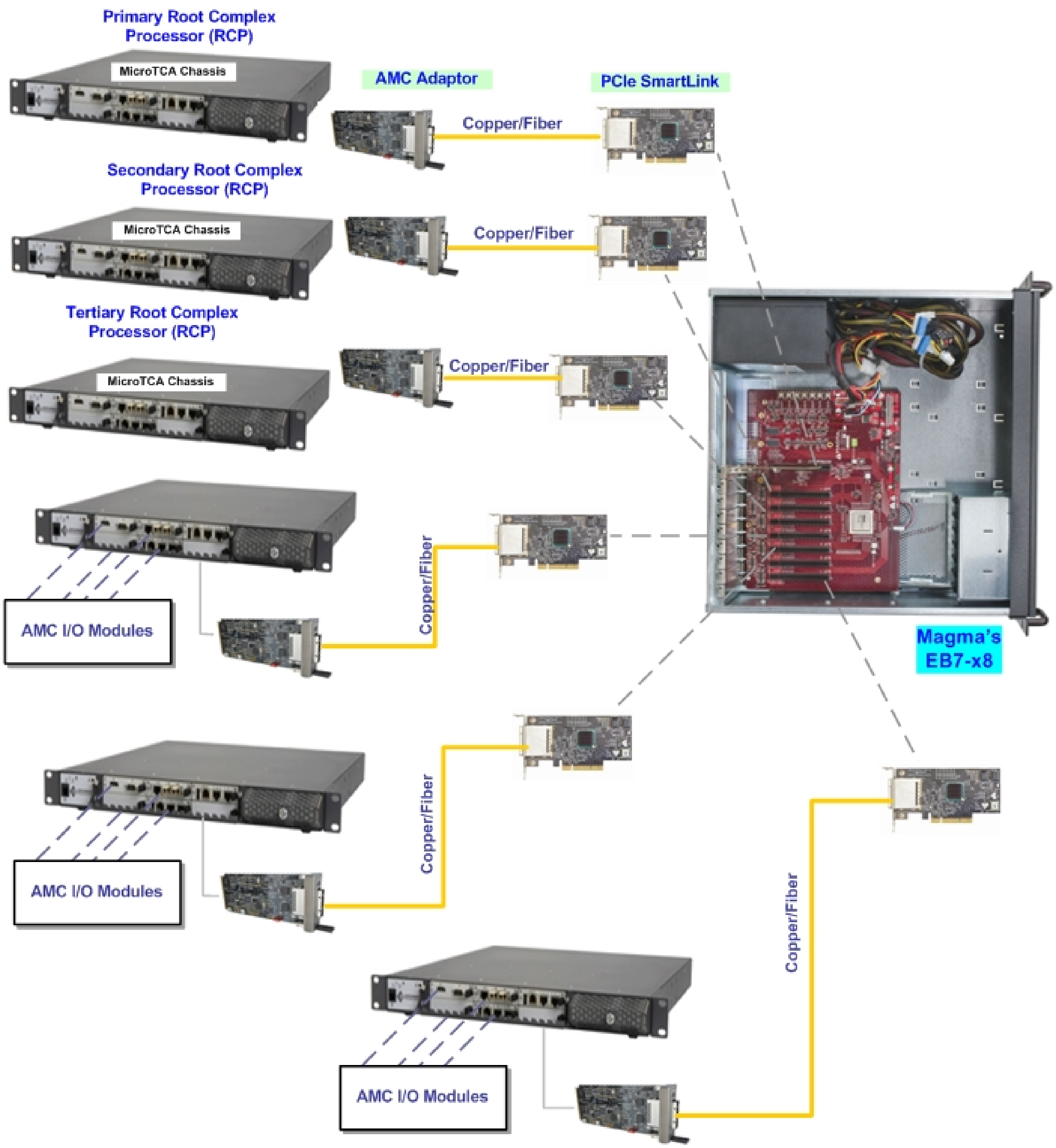

Figure 3.1

And finally, as more application services become prevalent in the Telecom network, and more computing power is required, Blade Server systems can utilize MicroTCA systems as an IO adjunct to provide Telecom interfaces such as T1/E1. Figure 3.2 depicts such a scenario.

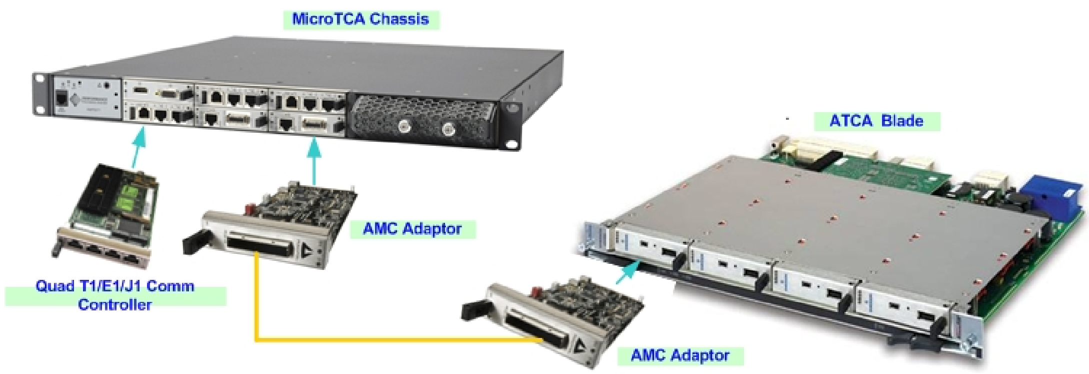

Figure 3.2

In the embedded world, there are always a vast range of product requirements. A PCI-Express adapter provides many bridging solutions, as highlighted in this whitepaper. It can bridge the world of AMCs to legacy PCI and PCI-Express hardware to support a broader level of functionality. Or bridge multiple MicroTCA systems together, to offer higher levels of functionality and redundancy. It can provide a high performance solution for storage solutions or provide an IO adjunct to server blade solutions. So if your application requires something more than current MicroTCA and AMC modules have to offer, consider "Bridging the Gap".

**The MicroTCA platform photographs are courtesy of Performance Technologies*
**The ATCA Blade photographs are courtesy of Sun Microsystems*