Implementation of an AdvancedTCA/ MicroTCA-compliant congestion management architecture

By Will Chu and John O'Day

In this article, Will Chu, President, CorEdge Networks, and John O'Day, Director of Technical Marketing, Celestica Communications Design Practice, discuss the need for a highly flexible 10 Gigabit Ethernet system in telecom systems with hardware-level congestion management to deliver a high-performance and manageable system.

Two separate but related megatrends currently exist in the telecom equipment industry. The first is the adoption of AdvancedTCA and MicroTCA (or on a combined basis, xTCA), which most Telecommunication Equipment Manufacturers (TEMs) are adopting for their next-generation architectures. xTCA promises to lower costs and speed time-to-market by freeing TEMs to focus on system architecture with high-value, high-layer software without the need to develop lower-layer, lower value-add infrastructure hardware and software from scratch. On the other hand, xTCA offerings to date have largely been designed with basic microprocessor blades that are good at doing data center-type computing tasks, but which are less equipped to handle heterogeneous networking wirespeed flows at 10 Gbps or above without interruption. These early implementations do not leverage the terabit capabilities that xTCA systems are able to deliver, and this has slowed the adoption of xTCA for next-generation telco networks.

The second megatrend is the need for wirespeed congestion management in order to deliver high-performance, scalable, and manageable systems across the network. This is becoming particularly important as wireless service providers begin to deploy 3G IP Multimedia Subsystems (IMS) that support voice, data, and video. At 1 Gbps speeds, supporting latency-sensitive quality of service is possible with software solutions. At 10 Gbps and above, the management of the network needs to be pushed down into lower-layer hardware, as software on microprocessor blades will be hard pressed to keep up with a network changing at 10 Gbps. Latency-sensitive applications such a voice and video will share the same physical infrastructure as less latency-sensitive data traffic, requiring the networks to manage traffic and congestion intelligently. When a single system is handling thousands of users, each with multiple megabits of data per users, the system quickly needs to handle multiple gigabits of data.

Unfortunately, until now, mainly expensive, proprietary techniques and hardware could support such wirespeed congestion management. To develop a lower cost, standardized mechanism for congestion management at Layer 2 over Ethernet, the 802.3ar committee has been established within IEEE, with members that include many of the world's tier one OEMs in telecommunications, semiconductors, and networking. It hopes to develop flow control standards for Ethernet by early 2007.

With a goal of assisting this standards effort, a group of companies from within the 802.3ar and xTCA efforts, CorEdge Networks, Celestica, and Fujitsu, a tier one TEM, along with other companies participating in future designs, partnered to create a prestandard implementation of wirespeed flow control, using xTCA form factors. The goal of this effort was to create a COTS implementation for this kind of equipment using AdvancedTCA-compliant equipment at the network core, and MicroTCA equipment at the edges, with AdvancedMC cards used ubiquitously.

The architecture was designed to meet several critical design requirements. These included:

Backplane fabric choice
The AdvancedTCA specification defines a number of backplane protocol and topology choices, and choosing the right set is a function of the features needed in a given application. From a protocol standpoint, the group decided to use Ethernet due to its low cost and universal deployment, scalable access speeds, and because it offers the best combination of technical characteristics and interface controller availability. Furthermore, moving to 10 Gigabit Ethernet (XAUI) for the fabric switch allows for a high-performance solution that ensures bandwidth is available to meet TEMs' future needs.

AdvancedTCA backplane topology
In terms of the backplane topology, PICMG specifies three options. The group chose the full mesh option as an effective choice for this application since it provides advantages in system scalability, system redundancy, and physical efficiency.

AdvancedTCA carrier card
For core AdvancedTCA switching, an AdvancedTCA carrier card was built that uses high-capacity XAUI fabric switch silicon supporting 60 SERDES lanes (4 lanes for 15 nodes) per AdvancedTCA blade. An example of this type of board is seen in Figure 1. It can mount two AdvancedMCs, with a 24- port XAUI switching fabric. The carrier card was designed by Celestica, using Fujitsu 10 Gigabit Ethernet switch chips and CorEdge Networks' IPMC. Customers developed the software layer required to deploy this solution for the customers' specific applications.

Figure 1

MicroTCA Carrier Hub (MCH) and fabric

For the MicroTCA system, the MicroTCA Carrier Hub (Figure 2) from CorEdge Networks was used along with an MCH fabric card that also used the Fujitsu switch chip.

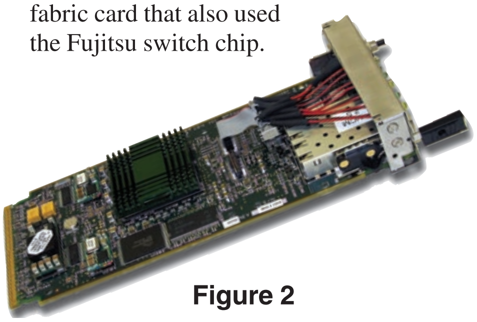

Figure 2

AdvancedMC/AdvancedTCA/ MicroTCA interoperability

A critical consideration when designing an architecture for congestion management across a wide range of packet types and network applications (and with different boards designed by different vendors) is interoperability. Creating such a solution requires *bridge* silicon that handles multiple protocols (SPI4.2, GMII, RGMII, PCI/PCI-X, PCI Express, 1 Gigabit Ethernet, 10 Gigabit Ethernet, RapidIO, and Serial RapidIO) so that it becomes a generic interface for application-specific silicon that does not have 10 Gigabit serial native interfaces (for example, NPU, Microprocessor, Security, Semantic Processing, and DSP). This bridge silicon also must support popular telco backplane protocols (1/10 Gbps Ethernet and Serial RapidIO) so that AdvancedMCs can become reusable in Ethernet or Serial RapidIO applications and in both AdvancedTCA carriers and MicroTCA.

The bridge silicon used in this project was developed by CorEdge Networks, and is designed to provide Ethernet Layer 2 hardware based flow control using IEEE 802.3ar techniques to provide guaranteed packet delivery and congestion management at low latency at 10 Gbps using Ethernet. For this demonstration, several CorEdge Networks 10 Gbps AdvancedMC Line Cards (Figure 3) were used.

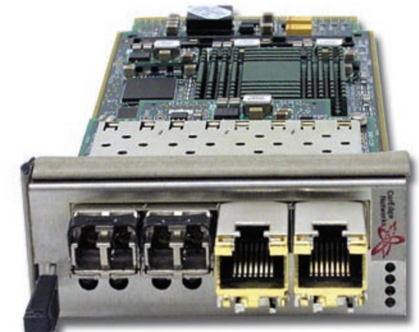

Figure 3

The CorEdge Networks bridge chip and IPMI MMC solution are also being used by third parties to develop next-generation AdvancedMCs that will interoperate within the overall architecture. These next-generation AdvancedMCs use different application silicon (NPU, Microprocessor, Security, Semantic Processing, and DSP) in conjunction with the bridge silicon so that they are enabled with 10 Gigabit Ethernet (XAUI) with IEEE 802.3ar hardware flow control.

At GLOBALCOMM 2006, a working version of this architecture will be demonstrated at the PICMG pavilion. Streaming video from a source laptop (along with a second-source 10 Gbps data stream) will go into a MicroTCA chassis, then switch through an AdvancedTCA switch, into a second MicroTCA chassis, and then play on a second laptop. When flow control is engaged, the video stream comes through without degradation. However, when flow control is disengaged with the second 10 Gbps data stream contending with the video for priority, the video becomes unwatchable.

We believe that this first working demonstration of a prestandard IEEE 802.3ar network in an xTCA-compliant form represents a milestone for the telecom industry, and we hope that it spurs accelerated adoption of the IEEE 802.3ar standard.

Will Chu *is the president of CorEdge Networks, a leading developer of AdvancedTCA, MicroTCA, AdvancedMC, and IPMI products, and he helped to start the company in 2004 while he was a principal at Signal Lake Ventures. At Signal Lake Ventures he focused on early stage venture capital investments since 2001. Prior to that, he held various technical and managerial positions at Fidelity Investments, Texas Instruments, ITU Ventures, and a number of startups. He holds an MBA from the MIT Sloan School of Management and graduate EE degree from Tufts University.*

John O'Day *is the director of technical marketing for Celestica's Communications Design Practice, of which the company's building blocks for AdvancedTCA systems and boards for Celestica's design services are a part. Prior to joining Celestica, John worked with Arrow Electronics. Throughout his career, John has held roles in technical marketing, business development, sales, and product management in the embedded computing market – all relating to embedded and high-performance computing in VME, CompactPCI, and most recently, AdvancedTCA.*

For more information, contact the authors at:

Will Chu
CorEdge Networks, Inc.
50 Commonwealth Avenue, Suite 504
Boston, MA 02116
Tel: 617-267-5205
E-mail: info@coredgenetworks.com
Website: www.coredgenetworks.com

John O'Day
Celestica Communications Design Practice
1150 Eglinton Avenue East
Toronto, ON, Canada M3C 1H7
Tel: 888-899-9998 or 416-448-5800
E-mail: joday@celestica.com
Website: www.celestica.com