» Whitepaper «

Embedded Systems at 10 Gigabit Speed

Beyond the speed limit – Double-Width MicroTCA Systems

Irene Hahner, Claudia Bestler

Embedded Systems at 10 Gigabit Speed

Beyond the speed limit – Double-Width MicroTCA Systems

High-Speed Ethernet allows simplifying the structure of computer systems in a significant way. 10G is available in advanced CPUs and Ethernet MACs for processing and storage subsystems. One pipe may handle both the operation of clusters as well as high-speed communications inside and outside of the system. Resources may be allocated by virtualization, including network interfaces. 10 Gigabit technology is now available in embedded systems. This paper presents a summary of 10G embedded systems based on the world-wide industrial standards AdvancedMC™ and MicroTCA.

CONTENTS

Why 10 Gigabits?

Multi-core CPUs come with 10 Gigabit/sec speed for inter process communication. Among the reasons to use 10 Gigabit Ethernet are:

» Interfaces: Prices have come down. There are SFP+ transceivers available in the same small form factor as SFP with passive copper links for short range, price sensitive applications (such as chassis interconnect, up to several meters), as well as optical connections for longer distances. At the same time, embedded standards such as MicroTCA™ and AdvancedMC™ support 10 Gigabit speed over the backplane, including multiple interfaces, as well as other high-speed fabrics such as PCIe or Serial RapidIO.

» Single Pipe: Different traffic classes can be allocated to one single pipe (with QoS providing many service levels). With protocols such as iSCSI, storage can use the same interface technology as inter process communication or exterior services (such as Web-traffic or Voice over IP). For aggregation of traffic, 10G avoids multiple 1G links with complex link aggregation mechanisms.

» Virtualization: The number of virtual machines per processor is increasing. Virtual machines are becoming a design principle on multi-core CPUs. 10G Ethernet MACs support virtualization, i.e. network interfaces can be shared between VMs in a hardware independent way.

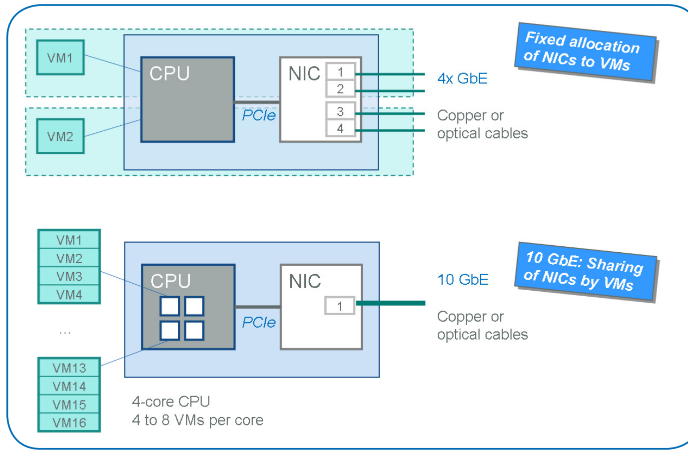


Figure 1: Allocation of Network Interfaces to Virtual Machines

VMs in embedded computing are increasingly being used to encapsulate parallel jobs. Today's four core CPUs easily support four to eight VMs per core. As shown in Figure 1 (lower part), this results in a number of network interfaces exceeding 16 ports. In addition, the traffic capacity of the CPU exceeds 1Gigabit speed, requiring multiple 1GbE ports per VM. 10G NICs with support of virtualization solve many problems by sharing of 10G Ethernet ports between VMs without need of link aggregation. Most importantly, inter-VM traffic is handled within the NIC without impact on the Hypervisor.

Virtualization drives 10G

On many desktop devices, virtualization has been used to boot one operating system or the other. In this case, one Virtual Machine (VM) may own the complete hardware, with no need to share. In IT server infrastructure and in embedded systems, virtualization is increasingly used to run multiple VMs in parallel on the same hardware. With current Network Interface Cards, Ethernet ports cannot be shared easily between parallel VMs. The usual solution is to allocate Ethernet ports to individual VMs, as shown in Figure 1 (upper part). While this type of solution works with a limited number of VMs and a sufficient number of NICs, it has major drawbacks with increasing number of VMs:

» The number of network interfaces needs to match the number of VMs, generating extra cost, space, cabling, as well as the hassle of configuration and maintenance.

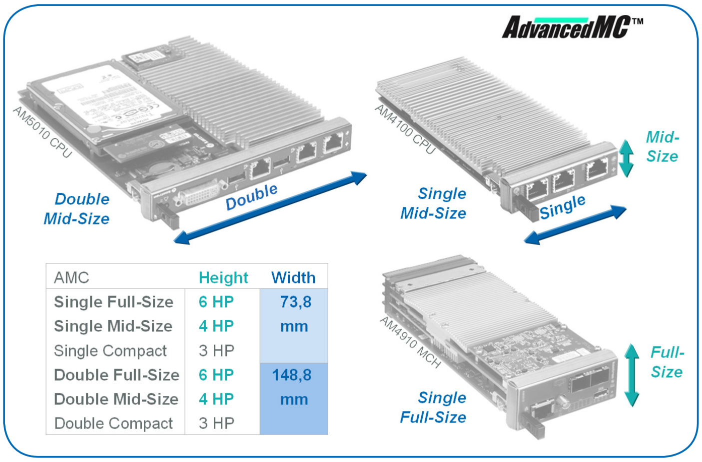
» The Hypervisor – the software layer running on the hardware and providing the virtualization - needs to manage traffic between VMs (and thus becomes a multi-port Ethernet switch), or alternatively traffic between VMs is channeled over exterior network interfaces.

10G Embedded Technology Ready to Use

While 10G has been available for embedded server technologies like AdvancedTCA® for some time (and is moving to 40G now), 10G is becoming available in smaller form factors with lower thermal budget and hence lower power consumption such as MicroTCA™ and AdvancedMC.

AdvancedMC

AdvancedMCs (AMCs) are specified in different form factors. The depth of the boards is 183.5 mm. From a horizontal perspective, the boards vary between 3HP, 4HP and 6HP in height. The 4HP Mid-Size form factor and the 6HP Full-Size form factor represent the most popular implementations. Like CompactPCI® boards, which are available in 3U and 6U form factor, AdvancedMCs are also available in two different form factors: Single (2U) or Double (4U) sizes. Figure 2 shows a summary of form factors.



AMC	Height	Width
Single Full-Size	6 HP	73,8
Single Mid-Size	4 HP	mm
Single Compact	3 HP	
Double Full-Size	6 HP	148,8
Double Mid-Size	4 HP	mm
Double Compact	3 HP	

Figure 2: AdvancedMC™ Form Factors

The fundamental difference between the basic form factors is the thermal budget. Single AMCs may dissipate up to 40 Watts, double AMCs up to 80 Watts. Single Mid-Size AMCs (2U, 4HP) such as ATCA® processor boards or ATCA® hubs are frequently used in ATCA® systems, where they are placed as modules on ATCA® carrier boards. The name "Advanced Mezzanine Card" reflects this type of application. The mid-size form of 4HP factor keeps 2U space on the ATCA® blade, which provides a better mechanical stability. AMCs on ATCA® are used as system controllers, interface cards or hard drives, not as main processors. Hence the limited thermal budget is not critical.

For implementations outside of ATCA, such as MicroTCA, both single and double form factors are used. The implementation of main processors benefits from the much higher thermal budget of the double form factor: 80 Watts instead of 40 Watts. Also, the extra PCB space can be used to mount hard disk drives or solid state drives directly on the processor board. This way, the double AMC represents a complete single board computer, i.e. one serviceable entity, rather than placing the drive on a separate AMC and wasting another AMC slot. The double front space allows extra interfaces, such as additional USB ports, or graphics via DVI, VGA or Display Port.

Equipment practice with AMCs shows, that 120 Watts are fairly easy to cool in a single air stream. 120 Watts corresponds to 3 processor AMCs in single format in a serial configuration within the same air stream. The double format of maximum 80 Watts per AMC does not represent any challenge for cooling and can simplify the systems design considerably. Figure 3 shows some samples of double AMC boards.

The Kontron AdvancedMC server blade AM5030 shown on the top left reflects the benefits of the double form factor. The thermal budget allows a server class Intel® quad core Xeon® processor to be placed on the blade at 1.73 GHz speed. The double full-size board (4U, 6HP) provides sufficient space for an efficient heat sink, a solid state drive module, plus 3 banks of DDR3 memory (including ECC) over three channels on VLP DIMM modules for a total of up to 24 GB. The Kontron AM5030 also features a server class chipset, as well as the latest generation of 10G NIC. The 10G NIC directly connects to the CPU via eight lanes of PCI-Express and provides dual 10GbE interfaces to the fabrics on ports 8-11, as well as ports 17-20 (according to the SCOPE profile for MicroTCA™ fabrics). Being a server class processor, graphics is not the focus of the Kontron AM5030. Still, it provides a graphics controller and VGA port on the front, so it may be operated with operating systems using a graphic user interface.

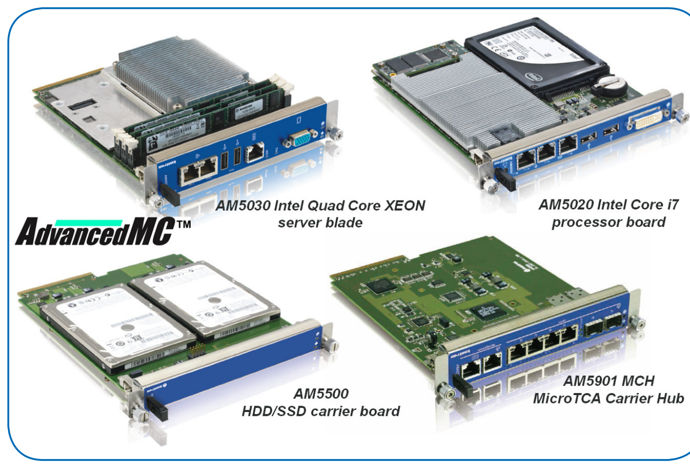


Figure 3: Double Wide AMCs

The Kontron AdvancedMC processor board AM5020 on the top left represents a single board computer based on an embedded Intel® Core™ i7 dual core processor at 2.0 GHz and 2.53 GHz speed. It uses dual channels of soldered DDR3 memory (incl. ECC) up to 8 GB. The Kontron AM5020 can also be equipped with a hard disk drive or solid state drive. The Intel® Core™ i7 processor provides integrated graphics; a DVI interface is accessible on the front. The AM5020 provides high-speed PCI-Express fabrics with a total of 8 PCIe lanes on AMC ports 4-7 and 8-11. In a MicroTCA™ system, those fabrics can be used in combination with a high-speed interface board, like frame grabbers or customized high-speed interfaces.

For mass data, the Kontron AdvancedMC HDD/SSD carrier board AM5500 provides space for two SATA hard disk drives or sold state drives on a double mid-size form factor (4U, 4HP). The port mapping corresponds to the AMC and MicroTCA™ standards, which define SATA connections on AMC ports 2 and 3. This way, the Kontron AM5500 can be combined with any standard compliant processor AMC, such as the Kontron AM5030 or AM5020. Mounting multiple hard disk drives or solid state drives on a double AMC also represents a way of cost improvement: AMC slots represent a valuable resource, which should not be wasted for a single HDD. Modifications on a double carrier board in order to accommodate more than two drives are fairly easy, also the combination with a RAID controller or a small processor to provide network attached storage (NAS) via 10 GbE.

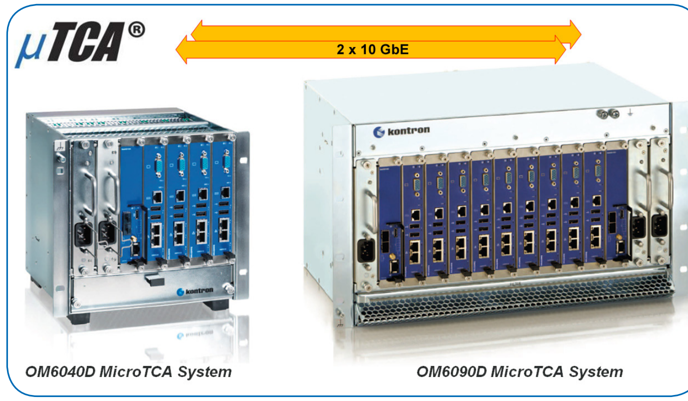


Figure 4: MicroTCA™ at full speed

The bottom right of Figure 4 shows the Kontron MicroTCA™ Carrier Hub (MCH) AM5901. The Kontron AM5901 contains all the management functions according to the MicroTCA™ standard, as well as a non-managed Ethernet switch, which connects to the GbE base interfaces of any standard compliant AMC (on AMC ports 0 and 1). The double form factor eliminates one bottle neck of MicroTCA™ MCHs: It provides sufficient front space for uplinks to the system. The Kontron AM5901 provides a total of 4 GbE uplink ports, which can be used over 4 RJ45 connectors, or alternatively, 2 SFP connectors plus 2 RJ45 connectors. The Kontron AM5901 targets entry level MicroTCA™ systems with point-to-point fabrics of PCI-Express and Serial Rapid IO.

MicroTCA

Like CompactPCI, MicroTCA™ represents a PICMG® standard for embedded systems. MicroTCA™ is the first standard using serial interfaces on the backplane, rather than the traditional parallel bus architectures. The AMC connector provides 21 ports, being used for high-speed links at 2.5 Gigabit/sec each. Thus, there is ample space for high-speed fabrics like 10 GbE, PCI Express, Serial RapidIO, SATA and others.

MicroTCA™ defines 1GbE as basic communication infrastructure between boards over the backplane. In addition, extra fabrics can be used. 10 GbE is currently implemented on the backplane using 4x 2.5G lanes using the XAUI interface (according to 802.3ap -KX4). Future implementations will use 10 Gigabit/sec speeds per lane (according to 802.3ap -KR). Systems with 10G backplanes are on the market for some time already. What will drive 10G is the next generation of processor blades on AMC form factor.

For high-end multi-processor configurations, matching systems are the Kontron MicroTCA™ platform OM6090D for redundant configurations and the Kontron OM6040D for smaller, non-redundant configurations. The Kontron OM6090D is a 6U high 19" system with front to back cooling for up to 9 AMCs and 2 MCHs. All components are front pluggable, incl. the AC power supplies. The Kontron OM6090D supports dual 10GbE links on the backplane, which can be used for dual-star fabrics in combination with two 10G MCHs. The Kontron OM6040D represents a half-size system for 4 AMCs plus one MCH with bottom up cooling and single-star 10 GbE fabric.

The Kontron AM4910 is a fully MicroTCA™ compliant MCH with 1 GbE base switch and 10 GbE fabric switch for up to 12 AMCs plus MCH interlinks. It provides two 10 GbE uplinks on the front plate via SFP+. The Kontron MicroTCA™ Carrier Hub AM4910 is a member of the Kontron family of embedded switches with ATCA, CompactPCI® and MicroTCA™ form factors. Thus, the feature sets for configuration and management including exterior management interfaces is compatible. The Kontron AM4910 comes with carrier grade support of clocks and redundancy. For other high-speed fabrics like PCI-Express or Serial Rapid IO, the Kontron AM4904 provides PCI-Express fabric switches, respectively Serial RapidIO fabric switches. The front plate is available in Single and Double form factors.

Areas of Application

MicroTCA™ represents the most advanced embedded standard today, which provides high-speed serial fabrics. PCI-Express can be used between processor AMCs and high-speed IO boards. PCI-Express is frequently implemented in a cost efficient way by point to point connections between boards on the backplane, or, alternatively, as switched fabric via an MCH. Serial Rapid IO provides high-speed connectivity for applications which demand extremely low latency between processors. The vast majority of inter-processor communication can be handled by 10 G Ethernet. Due to the high-speed fabrics, MicroTCA™ is particularly well suited for application which demand high-performance and high throughput. The range of applications covers telecommunications, industrial automation, military communication systems, as well as medical applications. In telecommunications, systems may be implemented in a completely redundant way, with no single point of failure, and all components exchangeable on the running system (hot-swappable).

Such configurations apply for communication servers, such as IPTV, professional mobile radio, or military communication systems. Beyond military communications, other demanding areas of application are radar or sonar data processing. In telecommunication systems, one important area is test equipment, such as load testers and conformance testers for UMTS or the next generation of mobile radio, LTE.

In industrial automation, MicroTCA™ allows the set-up of multi-processor systems with high performance and fast inter-processor communications. Among the applications are image processing for machine vision, surface inspection and sorting of goods, as well as real-time motion control with low latency. Image processing also applies for medical applications, such as computer tomography, x-ray scans in angiography, ultrasonic and spectroscopy.

As shown in this article, 10 Gigabit technology is getting more and more important in the Embedded world. With double-width AdvancedMCs and the respective fitting MicroTCA systems there is already a ready-to-use, commercial-off-the-shelf solution available today to meet this requirement.

About Kontron

Kontron is a global leader in embedded computing technology. With more than 30% of its employees in R&D, Kontron creates many of the standards that drive the world's embedded computing platforms. Kontron's product longevity, local engineering, support, and value-added services helps to create a sustainable and viable embedded solution for OEMs and system integrators.

Kontron works closely with its customers on their embedded application ready platforms and customer solutions, enabling them to focus on their core competencies. The result is an accelerated time-to-market, reduced total-cost-of-ownership and improved overall application with leading-edge, highly-reliable embedded technology.

Kontron is listed on the German TecDAX stock exchanges under the symbol "KBC".
For more information, please visit: **www.kontron.com**

CORPORATE OFFICES

Europe, Middle East & Africa
Oskar-von-Miller-Str. 1
85386 Eching/Munich
Germany
Tel.: +49 (0)8165/ 77 777
Fax: +49 (0)8165/ 77 219
info@kontron.com

North America
14118 Stowe Drive
Poway, CA 92064-7147
USA
Tel.: +1 888 294 4558
Fax: +1 858 677 0898
sales@us.kontron.com

Asia Pacific
17 Building,Block #1,ABP.
188 Southern West 4th Ring Road
Beijing 100070, P.R.China
Tel.: + 86 10 63751188
Fax: + 86 10 83682438
kcn@kontron.cn