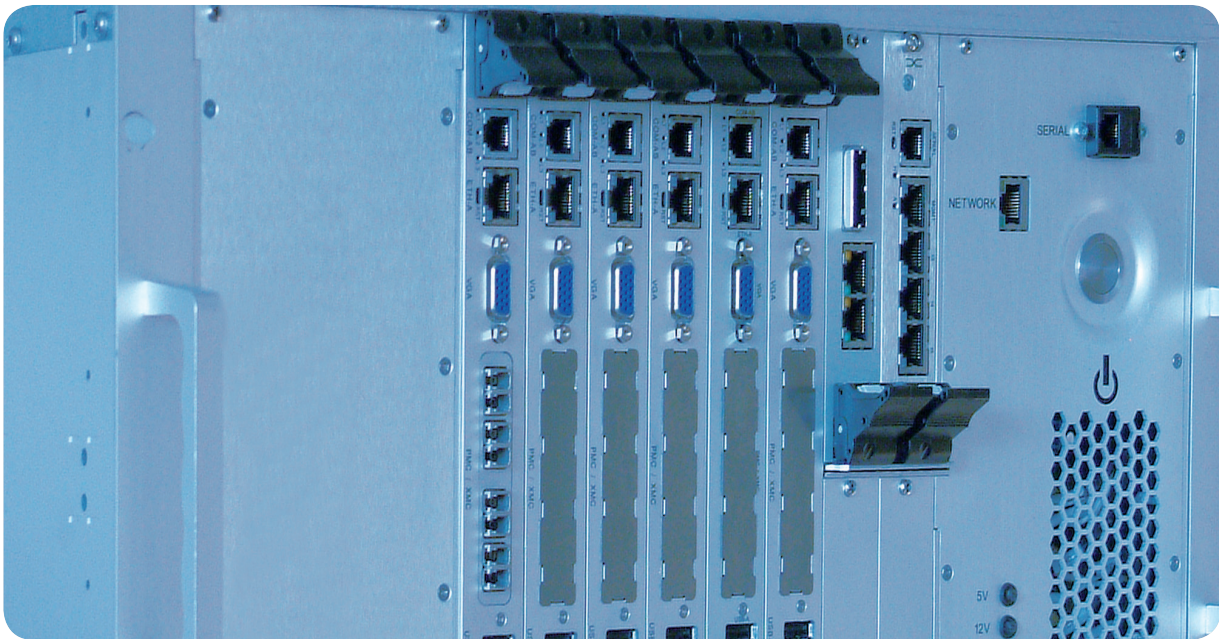# » Application Story «

PCI Express® Fabric breakthrough

Up to 4.4 GB per second TCP/IP communication

VPX based multi-processor applications with PCI Express® and VxFabric

Recently Kontron and PLX announced an industry breakthrough in the deployment of PCI Express® (PCIe®) and VXFabric technology as a VPX backplane interconnect for TCP/IP communication. Up to 4.4 gigabytes per second performance is now available for TCP/IP and up to 5.6 for raw data. This represents a tremendous jump in processing power and I/O data bandwidth for high performance embedded computing applications such as radar and sonar systems as well as video based situational awareness systems.

PCI Express® (PCIe) technology is being used across multiple applications – storage, server, graphics, communications, instrumentation, medical, embedded and consumer. However, more recently PCIe has been picked as the choice of interconnect technology in multi-processor applications. This is made possible by the incarnation of Gen 3 speeds of PCIe now supporting up to 8 gigatransfers per second (GTps). With a single PCIe switch it is now possible to support 1.5 Tbps of data transfer, which clearly enables newer usage models. Additionally, the latest PCIe switches support features such as two non-transparent (NT) ports for redundancy, direct-memory access (DMA) for faster data transfer, and spread spectrum clock (SSC) isolation for connections across different clock domains. All these features support the design of PCIe interconnect technology

in newer applications, such as VPX based multi-processor communication.

VPX is a VITA standard specification for high-bandwidth military and aeronautical systems, which allows board computers to move away from parallel bus architectures on the backplane and implement high-speed serial link point-to-point connections such as PCI Express®, GbE or Serial RapidIO between boards. VPX offers a well defined ecosystem for multi-vendor, multi-module, integrated system environments, providing developers and system integrators a high level of performance and interoperability for their specific COTS-based rugged systems.

Need for advanced connectivity and bandwidth

VPX based systems are deployed at the forefront of many demanding image-intensive aerospace and defense applications. In these application areas the need for advanced connectivity and bandwidth is constantly evolving because of requirements for higher fidelity of data (higher resolutions), different types of data (video versus audio), and the amount of data sources (higher numbers of sensors). In particular, next generation radar, targeting and surveillance systems for UAVs, and broadband electronic warfare monitoring and jamming systems are requiring enhanced resolution imagery, higher I/O rates, and higher performance switch-based serial communication fabrics. For all these size-, space- and weight-constrained applications it is essential to increase performance without exceeding the thermal budgets.

Most of these already field-deployed applications use TCP/IP via Ethernet for inter-processor communication. But native Ethernet is typically limited to what the processor and/or chipset provide and currently 10-gitabit Ethernet and even 40-gigabit Ethernet aren't natively supported by highly energy-efficient embedded processors and chipsets. If higher bandwidth is a must, radar and sonar engineers, as well as those developing video based situational awareness applications, currently have to use an alternative data plane to increase performance, which also usually requires modifying the application code. One way that helps to bridge this inconvenient gap in technology deployment of higher bandwidths without modifying application code is to facilitate the ongoing evolution of PCIe. Switches, such as PLX Express Lane Gen 3 devices with up to 96 lanes and 24 ports, are designed for such deployment, as each lane can transfer 8GTps. By using latest PCIe interconnect technology to implement Ethernet over PCIe (TCP/IP over PCIe) as well, VPX system designers can now apply considerable innovation to enable a higher performance while application software stays protected from obsolescence. Applications can facilitate native PCIe ports on all the components of the VPX single-board computer (SBC) to keep the costs down and reduce the latency of the system. The reduced latency comes as a result of using PCIe for inter-processor communication without the need to convert from multiple technologies.

From the cooperation of Kontron and PLX engineers, comes VXFabric, which helps those engineers reduce development costs by simplifying and accelerating application development of inter-CPU communication in VPX system architectures. VXFabric provides the required software between a PLX Technology PCIe Gen 3 switch and the bottom of the standard TCP/IP stack, which allows the boards to run their existing TCP/IP-based application without having to be modified, while also enabling migration to emerging standards, such as 10 gigabit and 40 gigabit Ethernet.

Inter-board communication at hardware speed

VXFabric delivers engineers an open infrastructure which implements efficient inter-board communication at hardware speed. The architecture is compliant with the OpenVPX standard (VITA65) which defines two main hardware topologies of the backplane: distributed and centralized topologies. In its current form, VXFabric can simultaneously interconnect up to 12 nodes via PCI Express®. The physical interconnect through the backplane can be made in various implementations according to the choice of backplane style:

» A distributed PCIe backplane can be used without the need of any PCIe switch. In this case, very compact architectures can be proposed, as long as the relevant data path between nodes is implemented in the backplane. This data path is mostly application dependent, making distributed backplane the right solution for cost effective deployments.
» When a higher bandwidth and more flexibility in the data plane are required, a centralized topology is recommended.

With a PCIe switch board, such as the Kontron 3U VPX PCIe and Ethernet Hybrid switch VX3905, it is possible to interconnect up to 12 nodes with the VXFabric via PCI Express® and establish communication links between any nodes using the same backplane. This approach fits well the HPEC application domain, as well as lab use where multiple application data paths can be evaluated with the same equipment.

Multiprocessor architectures

From the hardware point of view, the architecture is based on several CPU boards, each featuring several processing cores, interconnected through PCIe via the VPX backplane, using a PCIe switch. Software wise, VXFabric is equivalent to an Ethernet network infrastructure mapped over a switched PCIe express fabric. It implements the layers allowing the user to handle the communication with an IP socket programmatic interface. This API allows direct access to all classic protocols like TCP or UDP. Furthermore, VXFabric requires no modification of existing applications, which helps reduce development efforts and simplifies migration to the new VPX architecture.

The standard user programming model of Kontron VXFabric is based on the IP protocol and implements a socket layer API through an emulation of an Ethernet interface over PCIe –

similar to implementations of pseudo-Ethernet found in virtual machines. This API is the main reason why the compatibility of existing applications with VXFabric is guaranteed. The task to migrate from a Gbit Ethernet TCP/IP infrastructure towards Kontron VXFabric avoids the low level complex and proprietary APIs – like most of serial RapidIO™ or Infiniband™ implementations for example – and is straightforward. VXFabric is available under Linux for all Kontron VPX CPU boards and it is designed to be portable on other operating systems offering a modern TCP/IP stack as well as other architectures such as FPGAs. This implementation is scalable. Furthermore, VXFabric does not require any other infrastructure than the VPX backplane and a VPX PCIe switch to interconnect VPX boards. 100% of the necessary hardware and silicon involved comes from the main IT market and is not at the mercy of a little group of suppliers.
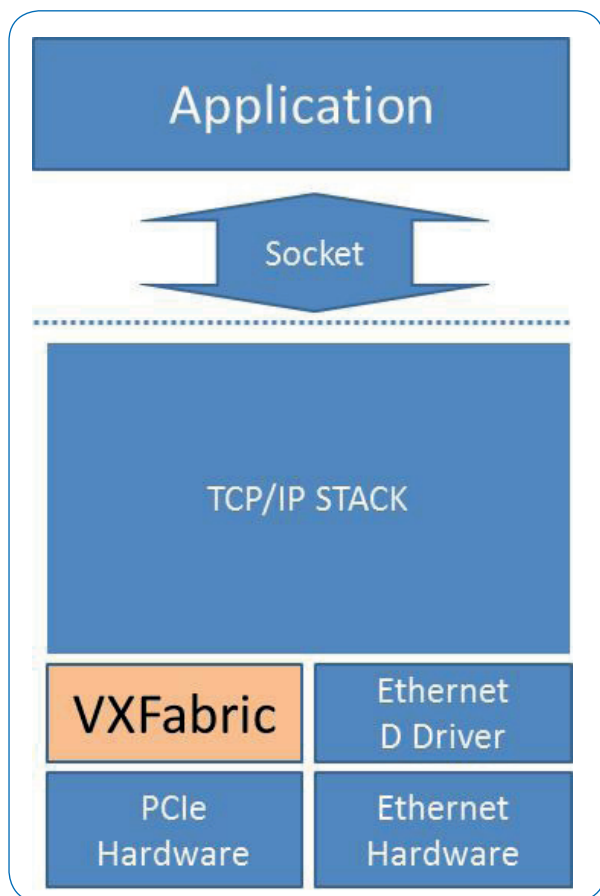

Figure 1: VX Fabric™ implementation

Data communication through Socket API

There is no difference when using the socket API on top of VXFabric or on top of a standard Ethernet interface. The datagram socket (UDP) or streaming socket (TCP) API is the preferred way to address the VXFabric which is usually dedicated to data communication. A good coding example can be found in the source code of the "iperf" (open source GPL) tool which has been used unmodified to get benchmark and throughput measurement of the VXFabric implementation.

For specific communication links where real-time data transfers have to be used, an application can choose to use the VXFabric raw mode to move large data sets directly from user memory to user memory. Driving the PCIe silicon DMA engines, the utility layers of VXFabric take care of handling the low level programming and the management of scatter/gather DMA job lists while the VXFabric infrastructure maintains the necessary system-wide address mapping coherency. With this raw mode, the user cannot rely on all the intrinsic features of TCP/IP communication (guaranteed delivery, data check, flow control) and must find other means to synchronize between the multiple parts of a distributed application.
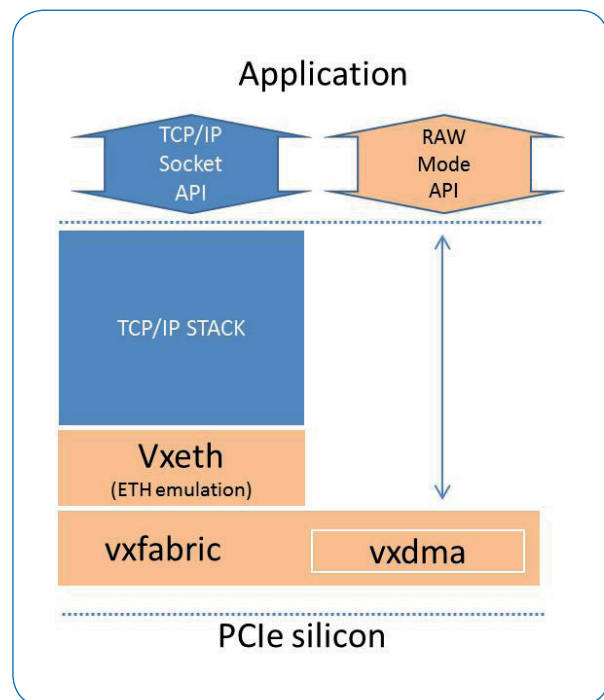

Figure 2: VXFabric™ TCP and RAM API

Performance Figures

Up to 5.6 gigabytes per second raw bandwidth can now be made available for VPX form factors. The configuration is capable to support up to 4.4 gigabytes per second for TCP/IP communication. This represents a tremendous jump in processing power and data I/O bandwidth for high-performance embedded computing applications, such as radar and sonar systems as well as video based situational awareness systems. Sustained throughputs have been measured with the standard iperf tool. PCIe Gen 3 performances have been demonstrated on VX3042/VX3044 cards hosting the Intel® Core™ i7 3rd-Gen processor (Ivy Bridge). For PCIe Gen 2, results were obtained for communications between VX3035 cards based on the Intel® Core™ i7 2nd-Gen processor (Sandy Bridge). And the VX3030/VX6060 cards, featuring the Intel® Core™ i7 1st-Gen (Arrandale CPU), were used to exercise the bandwidth in PCIe Gen 1 mode. During the measurements, Turbo boost and hyperthreading were disabled.

	Sustained TCP bandwidth	CPU load TCP (load as a percentage of one multicore CPU)	Sustained raw bandwidth (1)
PCIe x8 gen3	4.4 GBytes/s	10% (quad core)	5.6 GBytes/s
PCIe x4 gen3	2.5 GBytes/s	6% (quad core)	2.8 GBytes/s
PCIe x4 gen2	1.3 GBytes/s	8% (dual core)	1.5 GBytes/s
PCIe x4 gen1 or x2 gen2	0.6 GBytes/s	5%	0.8 Gbytes/s
(1): Very light CPU load, transfer handled directly by switch DMA engine.			

Figure 3: VXFabric™ sustained bandwidth and CPU usage

Conclusion

Based on PCIe Gen 3, VXFabric is a technology ideal to be deployed in many military and aeronautic applications both for high performance embedded computing (HPEC) with up to 12 quad-processor nodes, or in SWAP optimized designs with a few computer nodes and many I/Os. The pervasion of PCIe and TCP/IP in all computer technologies are positioning PCIe and VXFabric as the most efficient, inexpensive and perennial technologies for switch fabric in embedded systems. Looking forward, PCIe Gen 4, with speeds of up to 16Gbps per link, is expected to dramatically accelerate and expand the adoption of PCIe technology even in new market segments, while making it easier and economical to design with and use. As VXFabric is bridging PCIe and the 'internet of things' protocol TCP/IP together, application software can be protected from obsolescence for minimum the next 20 years.

Vincent Chuffart

Mil & Aero Product Manager, Kontron

Krishna Mallampati

Senior Director of Product Marketing, PLX Technology

About Kontron

Kontron is a global leader in embedded computing technology. With more than 40% of its employees in research and development, Kontron creates many of the standards that drive the world's embedded computing platforms. Kontron's product longevity, local engineering and support, and value-added services, helps 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 custom solutions, enabling them to focus on their core competencies. The result is an accelerated time-to-market, reduced total-cost-of-ownership and an 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	**North America**	**Asia Pacific**
Lise-Meitner-Str. 3-5	14118 Stowe Drive	17 Building,Block #1, ABP.
86156 Augsburg	Poway, CA 92064-7147	188 Southern West 4th Ring Road
Germany	USA	Beijing 100070, P.R.China
Tel.: +49 (0) 821 4086-0	Tel.: +1 888 294 4558	Tel.: +86 10 63751188
Fax: +49 (0) 821 4086 111	Fax: +1 858 677 0898	Fax: +86 10 83682438
sales@kontron.com	info@us.kontron.com	info@kontron.cn