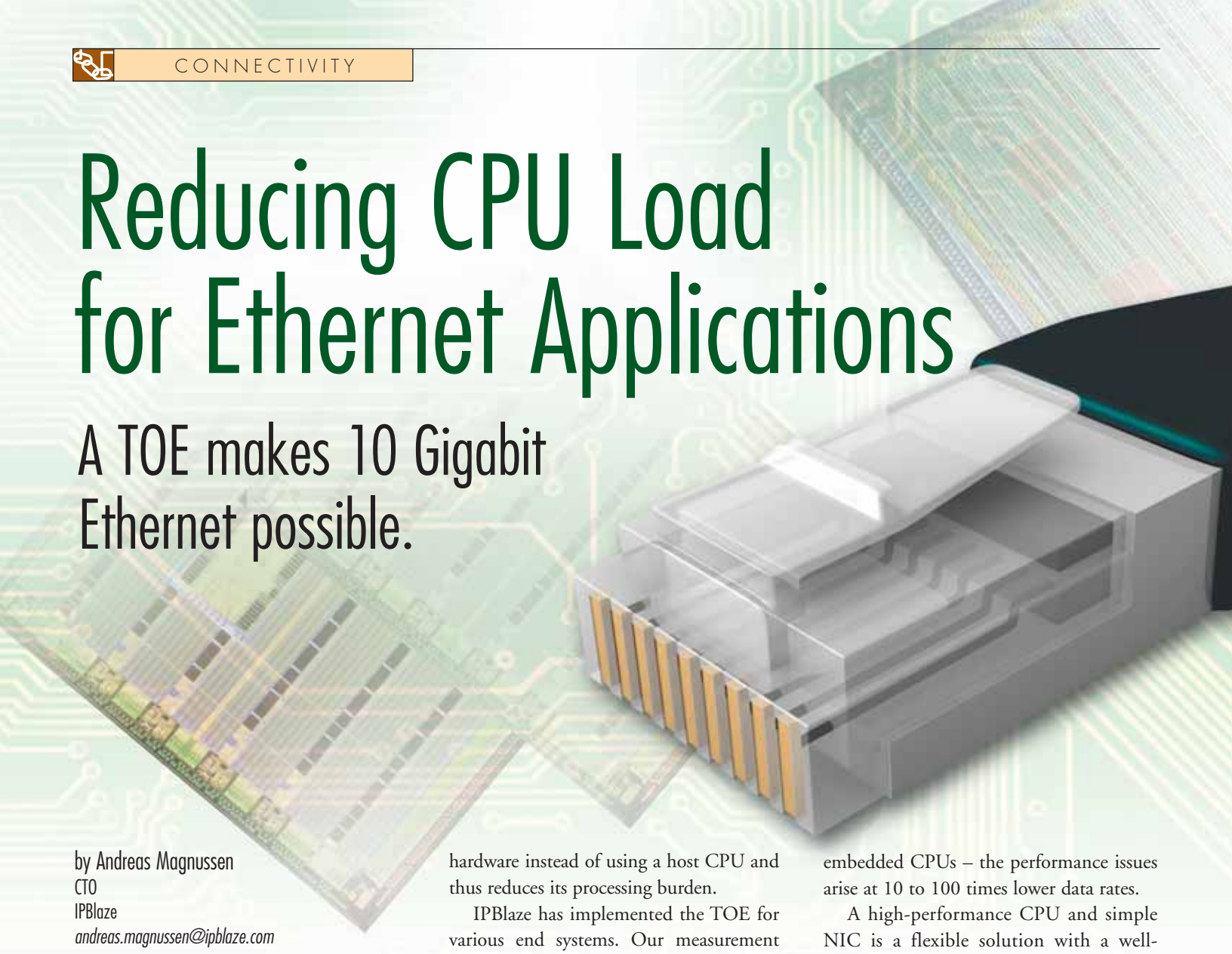
Reducing CPU Load for Ethernet Applications

A TOE makes 10 Gigabit Ethernet possible.



by Andreas Magnussen
CTO
IPBlaze
andreas.magnussen@ipblaze.com

Ethernet is playing an increasingly important role today, as it is used anywhere to connect everything. Not surprisingly, bandwidth requirements are increasing in the backbone as well as in end systems.

The implication of this increasing bandwidth is an increased traffic processing load in end systems. Today, most end systems use one or more CPUs with an OS and a network stack to implement network interface functions. For many applications, the increasing traffic load leads to performance issues in the network stack implementation. As these performance issues are seen already at 1 Gigabit Ethernet (GbE), implementing 10 GbE using a software stack is not a viable solution.

To solve these problems, IPBlaze has developed a unique and highly configurable TCP/IP offload engine (TOE). The TOE processes the TCP/IP stack at wire speed in hardware instead of using a host CPU and thus reduces its processing burden.

IPBlaze has implemented the TOE for various end systems. Our measurement results show that using an IPBlaze TOE reduces latencies and CPU utilization considerably. In this article, I'll give an overview of the implementation options available today for Ethernet applications and show where the IPBlaze TOE can give you the upper edge in terms of performance. Figure 1 shows performance examples with and without the IPBlaze TOE.

Ethernet Configurations

Table 1 shows the five different implementation options available today.

A CPU plus a simple network interface card (NIC) is a very general solution used in most PCs and servers today. A new CPU generation always provides higher performance and thus also increases network performance. However, the increase in network processing performance is significantly lower than the increase in CPU power. The same problem exists with embedded CPUs – the performance issues arise at 10 to 100 times lower data rates.

A high-performance CPU and simple NIC is a flexible solution with a well-known API (socket), and it is easy to implement applications in common PC/server environments. One of the downsides is the difficulty in scaling efficiently to 10 GbE because the load distribution between CPU cores and process intercommunication creates a bottleneck. Power dissipation (heat) is also a limiting factor in many systems.

Some NIC acceleration ASICs on the market perform protocol offload for specific applications such as storage. A high-performance CPU and NIC acceleration ASIC is a good solution if the ASIC supports all of the features needed (iSCSI, TCP offload). The functionality and bandwidth is fixed, however, which makes it very hard to add functionality or adapt to protocol changes without a huge performance impact.

FPGAs are so powerful today that it is possible to implement an accelerated 10

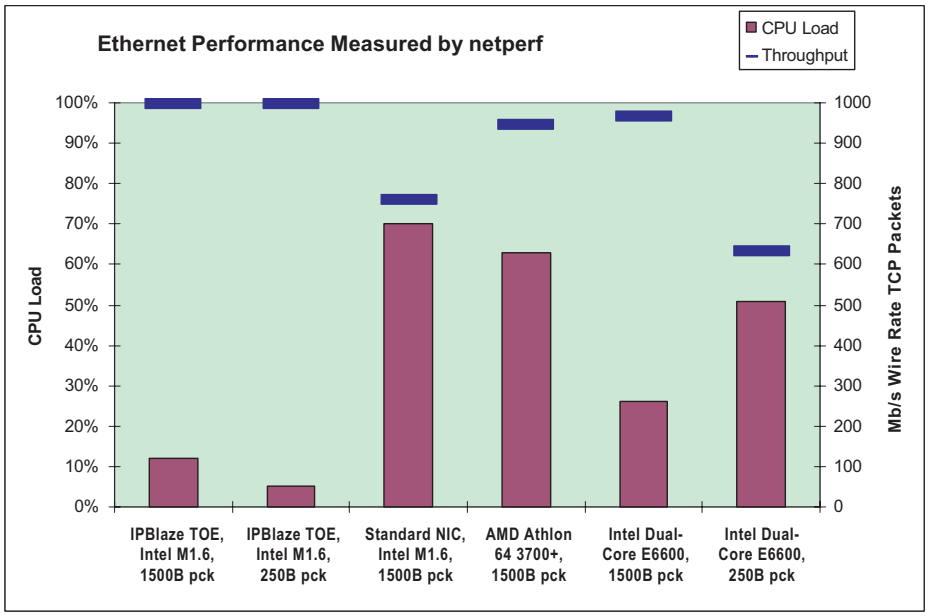


Figure 1 – The CPU power needed to process network data and throughput, tested at 1 GbE links, using the netperf performance measurements program.

High Performance	Properties
High-performance CPU + simple NIC	High CPU load, CPU limits performance
High-performance CPU + NIC acceleration ASIC	Good with in supported feature set, but inflexible solution
High-performance CPU + NIC acceleration FPGA	Flexible, scalable solution
Medium Performance	**Properties**
ASIC SoC (communication controller CPU, network if)	General solution, performance limited by software
FPGA SoC (integrated CPU, network acceleration)	Good for high data rates, flexible and scalable

Table 1 – Overview of implementation options

GbE NIC in an FPGA at a competitive cost point. One example is the IPBlaze TOE implemented in Xilinx® Virtex™-4 and Virtex-5 devices. The Virtex-5 LXT device has hardware support for PCIe (PCI Express), which makes it an attractive choice to use in PCs and servers.

Examples of ASIC SoCs are controller-type ASICs and communication controller ASICs. This is a good solution for multi-function systems with limited bandwidth (up to 10-20 Mbps). However, it is limited by low bandwidth.

A powerful example of an FPGA SoC solution is an FPGA with an IPBlaze TOE and an embedded CPU (a PowerPC hard-core or MicroBlaze™ soft-core processor). Applications processing can be done in software by the PowerPC while the FPGA hardware, together with the TOE, handles bulk transfers like video or images at wire speed (1 or 10 GbE). This solution makes it possible to use FPGA logic for fast data processing and fast data communication and software for complex low-bandwidth operations. Software in the embedded CPU can be Xilinx microkernel, Embedded Linux, or VxWorks.

The IPBlaze 10 GbE TOE NIC FPGA solution has the same performance as a TOE NIC ASIC while providing much higher flexibility. High-performance TOE NICs are now available using cost-effective FPGAs. FPGAs help you with fast time to market and ensure compatibility and interoperability.

Application Examples

The IPBlaze TOE is a general protocol offload engine including TCP offload. The IPBlaze TOE product lines include a number of TOEs with different functionality targeting different applications (Figure 2).

One example is an FPGA working as an intelligent high-performance TOE NIC at n-times-1 GbE or n-times-10 GbE connected to a PC through the PCIe interface. The NIC can be customized with add-on functions such as advanced switch flow control. You can also easily add protocol processing functions such as RDMA and iSCSI for storage applications.

Another example would be embedded video applications where Ethernet is used as a backplane. The TOE can transfer the images or video streams without the need for a CPU.

High Flexibility with FPGAs

FPGAs hold a number of advantages compared to ASICs and offer scalable solutions with high data rates: 1 GbE, n-times-1 GbE, 10 GbE, and n-times-10 GbE. The flexibility of the FPGA makes it possible to add new functionality and adapt to protocol changes. Advanced network functions like switch-specific flow control, or performance-enhancing features can be added. In load distribution, the IPBlaze TOE can distribute traffic to multiple CPU cores for higher system performance. Maintenance and upgrade of the TOE hardware functions located in the FPGA is done by firmware upgrade – much like driver software updates.

IPBlaze 10 GbE TOE core

The IPBlaze 10 GbE TOE core is an implementation of a full-featured TCP/IP networking subsystem in the form of a "drop-in" silicon IP core. It includes a standards-compliant TCP/IP stack implemented as a compact high-performance synchronous state machine.

The TOE core is built from a collection of synthesizable Verilog modules, which are customized before delivery to provide the best possible performance and feature set in a given system.

The TOE core can be instantiated along with your own IP and configured to operate without the need for host CPU support.

The IPBlaze TOE core can support as many as 1,000 concurrent TCP connections in on-chip memory. When a higher number of connections are required, a connection cache will allow you to access the most recently used connection states on-chip. The state information for non-cached TCP connections is stored in host memory or in dedicated external memory.

For applications where the TOE does not terminate connections but monitors all connections on a backbone, an external dedicated high-speed memory is used to support a very high number (1 million and above) of concurrent TCP connections.

The IPBlaze TOE core uses three well-defined interfaces:

- Xilinx 1 GbE and 10 GbE MAC network interfaces for either 1 GbE or 10 GbE support
- A high-speed RAM option for dedicated external memory for TCP connection state information
- Xilinx PCIe type or memory host processor interface

The host interface is essential for system behavior and performance and represents a significant part of the system complexity.

The host CPU has register access to TOE registers while the TOE has access to main memory. The main memory contains the send and receive data buffers, queue structures, and TCP connection state information if connection caching is enabled. An event system is used for fast message signaling between the host and the TOE. The event system includes a number of optimizations for high throughput and low latency.

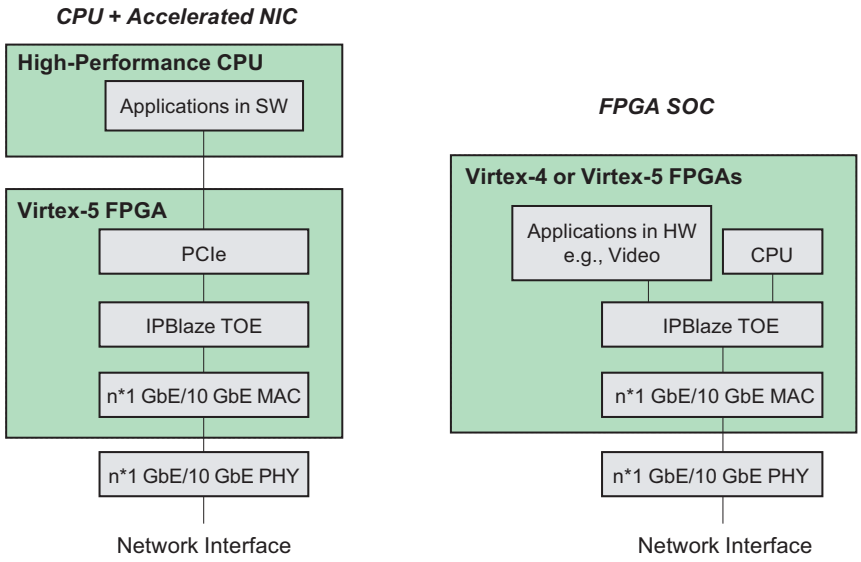


Figure 2 – Block diagram for a high-speed TOE solution for a PC/server and embedded system

OS-Compliant Socket API

Applications software interfaces to the TOE through a fully standards-compliant socket API, with significantly higher performance than a standard software socket implementation. The socket API is implemented at the CPU kernel level. The CPU load is reduced by 5-10 times and latency is improved by 4-5 times.

Hardware-based acceleration should be able to scale well. The evolution of performance in future Xilinx FPGA generations provides a good upgrade path. Furthermore, continuous improvements in the IPBlaze TOE technology also provide increased performance.

The flexibility in an FPGA solution allows easy upgrade of the FPGA code to support new protocol features.

The IPBlaze TOE core is designed to target a variety of FPGA families depending on the performance and functionality requirements. The IPBlaze TOE core has been implemented in Virtex-II, Virtex-4, and Virtex-5 devices.

Here are some example numbers from a 10 GbE TOE implemented in Virtex-5 devices:

- Core clock: 125 MHz
- TOE latency: 560 ns
- Throughput: 10G line rate (10G bidirectional)
- Packet processing rate: 12 million packets per second

Xilinx and IPBlaze provide TCP offload solutions that can be implemented as is or customized for functionality, size, speed, or target application. IPBlaze also offers a comprehensive set of high-performance TOE solutions for different market segments. Current TOE solutions include:

- IPBlaze General TOE for NIC applications.
- IPBlaze Security TOE for high-performance (2-times-10 GbE) bump-in-the-wire network monitoring. The security TOE supports more than 1 million concurrent TCP connections.
- IPBlaze Embedded TOE for SoC applications that require high-speed network communication (1 GbE and 10 GbE).

Conclusion

With standards-compliant TOE cores, IPBlaze has created a technology platform for the 10 GbE TOE market, working with industry leaders in high-performance computing solutions. Programmable solutions enable system architects to add functionality as needed. Integrating multiple IP cores into a single FPGA can reduce board costs and time to market.

To learn more about protocol offload solutions, visit *www.ipblaze.com.*