Different TOEs for different folks

By Al Basseri

As Moore's Law has proven, processor performance doubles every 18 months. This has inevitably caused an application performance challenge. It has taken a compute bound problem and changed it to an I/O bound problem. I/O storage and server manufacturers have recognized this and have been addressing this issue as the I/O bottleneck grows worse each day. Compare this problem to that of city and state transportation planners who have realized the rapid growth and subsequent daily traffic bottlenecks in major metropolitan areas such as Silicon Valley. These bottlenecks have hurt the long-term growth and stability of the area. As a result, businesses have partially or completely moved away from the Silicon Valley because of the high cost of operations and overall diminished ROI.

Just as transportation planners struggle to come up with alternative solutions to address the bottleneck, many hardware manufacturers have faced similar challenges. Unfortunately, most have not learned from the mistakes that city planners have made in the past. Simply building more housing or roads has not addressed the problem. Building new roads and then restricting the types of traffic that can move through these roads only creates new standards and adds another level of complexity. Yet, the fundamental problem still exists. Similarly in computer architecture, the problem is not simply storage, but it also involves server-to-server traffic.

The majority of I/O bound problems today are Ethernet based TCP/IP traffic. It is the processing for the transport layer and the link layer that are increasing the load on the host CPU and indirectly affecting the application performance. Mission-critical network applications require support for a high number of concurrent sessions while maintaining acceptable throughput.

Common TOE designs

TCP/IP Offload Engine (TOE) solutions offload TCP/IP processing from the host CPU, increase network throughput, and in essence, create more CPU cycles for application processing. Refer to Figure 1 for an overview of the key benefits derived from utilizing TOE. Many TOE developers have addressed the bottleneck issue strictly from a storage point of view by focusing on applications with few concurrent sessions and large data transfers. Just as most roads are packed with lots of cars and not larger vehicles, the typical mission-critical network application is being occupied by an increasing number of concurrent TCP sessions carrying small TCP packets over IP. Most available TOE designs cannot address the issue in the high-session count environment. Yet a majority of the I/O bottleneck occurs in such an environment.

Another drawback of many TOE designs in today's market involves the need for large buffer memory for TCP segment reassembly to address dropped, or out of order segments. Since the required buffer size is dependent on the TCP connection bandwidth and the end-to-end delay, the buffer grows with the network's speed resulting in higher costs. In addition, as the number of connections increases the performance drops. As it turns out, the bandwidth required to support that level of memory is at least twice the wire speed, requiring a complicated high-speed memory design. In this case, the TOE needs hardware in order to inter-

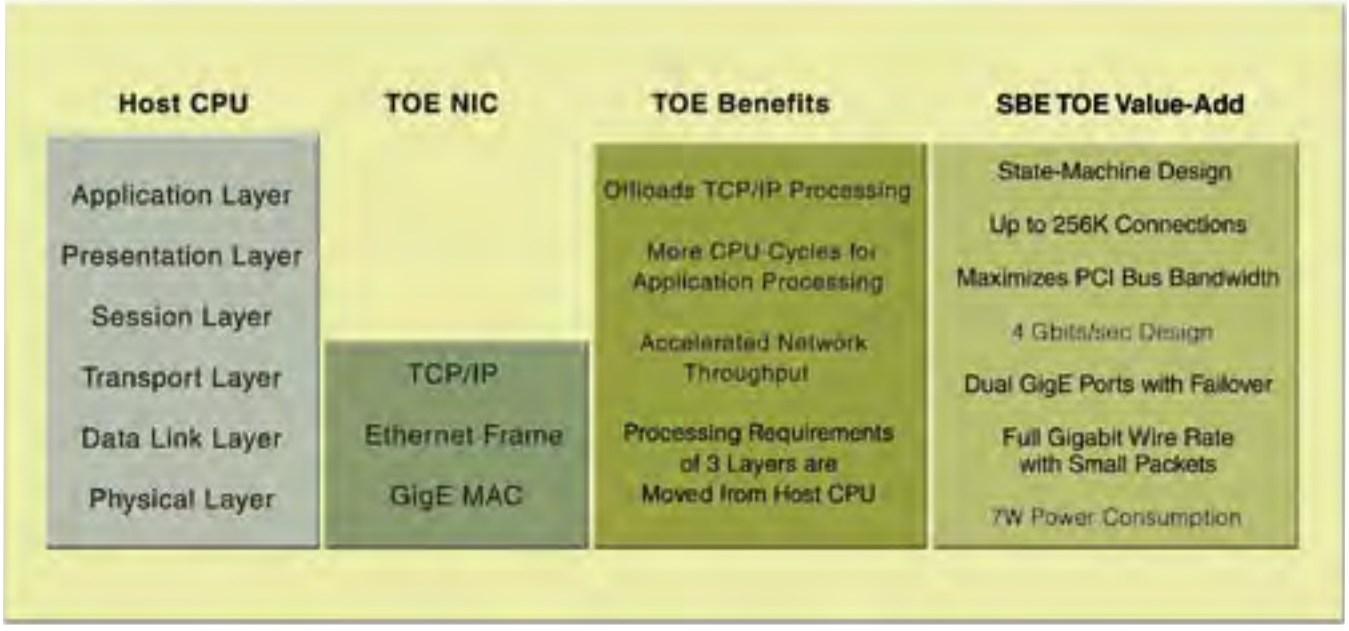

Figure 1

face to larger and wider memory, further elevating development time and costs.

Optimizing TOE for high session count applications

The optimal solution to the high session count problem is to design a fast-through TOE. A fast-through TOE allows all of the data, whether in-order or out-of-order, to be processed immediately from the TOE to host memory. This eliminates the additional cost and complexity associated with introducing a TCP reassembly buffer. To achieve this immediate result, SBE is developing a TOE solution that processes information from each TCP segment it receives immediately, without requiring a larger buffer. Figure 2 illustrates the general architecture of the SBE TOE solution. SBE's TOE board is based on a state machine design that supports up to 250,000 concurrent sessions while maximizing the bandwidth of the existing bus. More importantly, all of this occurs while consuming as little as 7W of power. It is a simple and effective solution for addressing high traffic loads on individual servers. Now as far as Silicon Valley traffic goes that problem will be for the new California Governor.

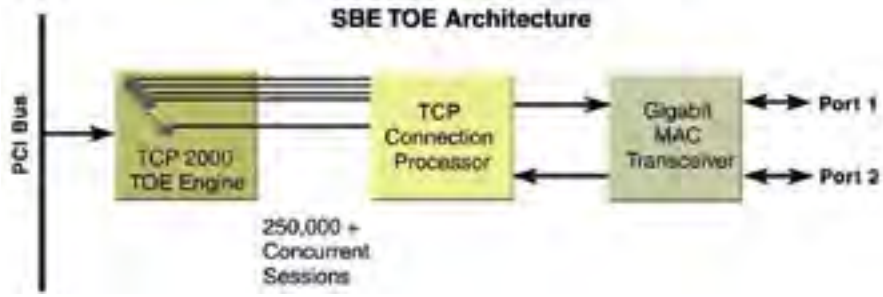

Figure 2

Al Basseri *has more than 10 years of management experience in marketing and support involving enterprise software, security, and storage area networks. At SBE, Al manages the strategic business development and product strategy for enterprise solutions, including TCP/IP Offload Engine. Prior to joining SBE, Al held key positions at industry leading companies, including BEA Systems. He holds a Computer Science degree from San Jose State University and has written numerous papers on enterprise support methodologies.*

For further information, contact Al at:

SBE, Inc.
2305 Camino Ramon, Suite 200
San Ramon, CA 94583
Tel: 925-355-7611
E-mail: abasseri@sbei.com
Web site: www.sbei.com