TCP Off-Load Engines
Dipping a TOE in the water

A White Paper

James J Hipperson September 2004

Assiette Ltd, Rasa Sayang, Ricketts Hill Road, Tatsfield, TN16 2NA
Email: info@assiette.co.uk, Tel: 01959-571-367

There is no denying that almost all of the constituent parts of a network have changed dramatically over the last 10+ years. From the co-axial cable of Thin & Thick Ethernet through to Switched infrastructures and UTP cabling and high speed interconnects for both LANs and WANs we have seen change, so it's nice to have a constant in TCP/IP. But even the way this works is now set to change.

Assiette

TCP OFF-LOAD ENGINES

Dipping a TOE in the Water

By James J Hipperson

Introduction

We have seen the design concepts of networking change dramatically over the last 10+ years. We have moved from the use of co-axial cables over which Ethernet ran at 10 Mbps, to high speed switched networks running at Gigabit speeds, even 10 Gigabit for the more demanding user. However, underlying all of this technology is a protocol that has made it all possible – TCP/IP. Without it we would not have the ability to run our LAN, our WAN or the Internet as we know them today. So before we discuss TOE technology lets look, at a high level, at TCP/IP.

What is TCP/IP?

The Internet protocols are the world's most popular open-system protocol suite because they can be used to communicate within and across any set of interconnected networks and are equally well suited for LAN and WAN communications. The Internet protocols consist of a suite of communication protocols, of which the two best known are the Transmission Control Protocol (TCP) and the Internet Protocol (IP).

Internet protocols were first developed in the mid-1970s, when the Defence Advanced Research Projects Agency (DARPA) became interested in establishing a packet-switched network that would facilitate communication between dissimilar computer systems at research institutions (the start of the Internet). DARPA funded the required research. The result of this development effort was the Internet protocol suite, completed in the late 1970s. TCP/IP was later included with Berkeley Software Distribution (BSD) UNIX and has since become the foundation on which the Internet and the World Wide Web (WWW) are based. Table 1 is a simplified description of the Internet Protocol Suite and illustrates how TCP/IP relates to the other OSI Layers.

Layer	Name	Internet Protocol Suite
7	Application	FTP, Telnet, SMTP, SNMP
6	Presentation	
5	Session	
4	Transport	TCP
3	Network	IP
2	Link	Ethernet, Token Ring etc.
1	Physical	Coax, UTP etc.

Table 1 – OSI Layers

IP – Internet Protocol

The Internet Protocol (IP) is a network-layer (Layer 3) protocol that contains addressing information and some control information that enables packets to be routed. Along with the Transmission Control Protocol (TCP), IP represents the heart of the Internet protocols. IP has two primary responsibilities: providing connectionless, best-effort delivery of datagrams through an

internetwork; and providing fragmentation and reassembly of datagrams to support data links with different maximum-transmission unit (MTU) sizes.

TCP - Transmission Control Protocol

TCP provides reliable transmission of data in an IP environment. TCP corresponds to the transport layer (Layer 4) of the OSI reference model. Among the services TCP provides are stream data transfer, reliability, efficient flow control, full-duplex operation, and multiplexing.

With stream data transfer, TCP delivers an unstructured stream of bytes identified by sequence numbers. This service benefits applications because they do not have to chop data into blocks before handing it off to TCP. Instead, TCP groups bytes into segments and passes them to IP for delivery.

TCP offers reliability by providing connection-oriented, end-to-end reliable packet delivery through an internetwork. It does this by sequencing bytes with a forwarding acknowledgment number that indicates to the destination the next byte the source expects to receive. Bytes not acknowledged within a specified time period are retransmitted. The reliability mechanism of TCP allows devices to deal with lost, delayed, duplicate, or misread packets. A time-out mechanism allows devices to detect lost packets and request retransmission.

TCP offers efficient flow control, which means that, when sending acknowledgments back to the source, the receiving TCP process indicates the highest sequence number it can receive without overflowing its internal buffers.

Full-duplex operation means that TCP processes can both send and receive at the same time.

Finally, TCP's multiplexing means that numerous simultaneous upper-layer conversations can be multiplexed over a single connection.

Add them together and we have TCP/IP, two protocols made for each other!

Networking overheads

As you can see from these descriptions IP, relative to TCP, has little to do, it is fast and efficient. However TCP has much to do and can present a major overhead/bottleneck to busy I/O intensive servers such as File and Database servers. This overhead takes the form of CPU cycles on the Server and can be anywhere from 30% through 90%. The solution is simple, take this overhead off of the Server CPU and let the Ethernet NIC do the work instead.

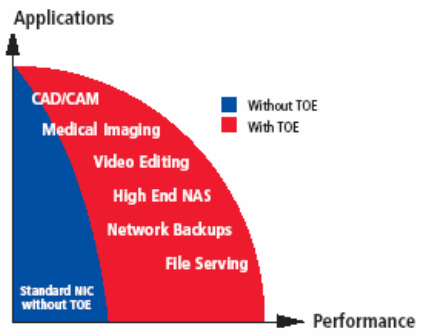

Fig. 1 – TOE Performance benefits

Hence the development of TCP Off-Load Engines or TOEs.

How and where can they be used?

In the following sections we give examples of different situations in which TOE cards can be used and illustrate their effect.

Doing More With Less

Many IT departments are currently faced with reduced capital budgets as well as decreased staffing levels. Considerable financial, staffing, and time investments have already been made in a typical data centre including original capital equipment costs, installation, configuration, and ongoing server maintenance and deployment costs.
Preserving these investments by extending the useful life of these systems can be critical to the success of many projects.

With Windows-based servers handling increasing network traffic, one focus is to find a solution that moves network data as quickly and effectively as possible.

Providing mission critical network availability with minimal server resource usage is the ideal.

The TCP/IP Bottleneck

File servers utilise TCP/IP as the underlying transport protocol for client communications with file protocols such as NFS and CIFS. These systems spend a significant portion of their resources preparing data for transmission over TCP/IP via Ethernet networks. Well balanced file servers frequently spend half their CPU resources performing network protocol processing. This resource bottleneck makes them prime candidates for upgrade of their network adapter with a TCP Offload Engine (TOE).

The TOE Solution

TOE cards provide significant performance boosts and server efficiencies independent of other improvements to the system. Installation of a TOE card allows the greatest number of clients to be served using the existing server. The upgrade is as simple as any other PCI card installation – simply install the hardware and driver – and one can experience immediate performance improvements including:

- Reduced latency
- Increased throughput
- Reduced CPU utilization

Quick, Easy, and Low Risk

Optimised file server platforms with TOE cards now support more users with the same or even fewer management resources. The quick, easy and low risk installation of a TOE solution avoids tedious, time consuming, and expensive server upgrades and data migrations.

Triple File Server Performance

Upgrading with a TOE also prepares the file server to take maximum advantage of Microsoft Windows Server 2003. Recent testing by Microsoft as shown that upgrading a file server to Windows Server 2003 with a TOE can triple file server performance over the same platform running Windows NT4 and using standard NIC cards.

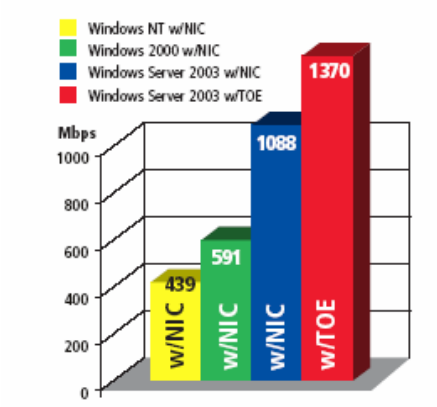

Fig. 2 – Peak File Server Performance

An illustration that a TOE can extend the scalability of one of the lowest cost storage solutions available – Windows based file servers. A TOE based NIC complies with all of the Ethernet standards so can support industry network standards for failover and link aggregation, additional performance benefits can be achieved such as:

- Added reliability
- Increased availability
- Improved scalability

Reduced Data Centre Costs

By implementing TOE into existing servers and achieving performance gains equal to more expensive servers, it is also possible to reduce the total cost of ownership by:

- Consolidating servers and maintenance fees
- Lowering power and facilities requirements
- Handling more users/clients with existing IT staff

It is always more cost-effective to implement a board level upgrade on an existing server than to replace it.

Backup Servers - Improving Their Performance and Efficiency

Most networks environments are increasingly constrained by large transfers of data backup, long backup times to storage devices, and less efficient use of server resources as a result. System and database availability is heavily impacted during backup periods, and with the rapid growth

of storage and the need to archive customer data, IT managers are actively seeking solutions for their backup problems.

TOE Enhanced Solution

The use of a TOE in backup servers can provide an increase in throughput to storage devices processing large amounts of data. In addition, server resources are returned for additional backup application processing. Medium to large IT departments have used TOE in their backup servers for both multi-streaming tape devices as well as virtual and physical tape libraries. In configurations where multiple systems are archiving their data simultaneously on a backup server with multiple tape drives, TOE offers immense benefits over conventional NICs in backup and restore performance.

Backup Server Performance

In recent backup server tests with TOE versus conventional NICs, TOE increased backup performance substantially in megabytes per minute (MB/min). The test also showed a reduction in system CPU resources needed to run the backup application, leaving resources to add even more tape drives to the backup server to further reduce backup time or manage storage growth.

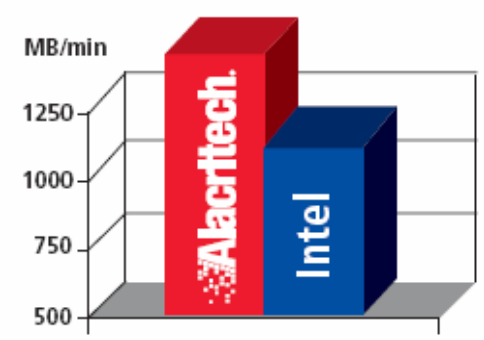

Fig. 3 – Backup Throughput Gains

Test Results

The chart in Figure 3 shows throughput results when running Veritas Backup Exec 9.0 using Intel PRO/1000T and an Alacritech 1000x1 Server and Storage Accelerator against IBM LTO tape drives. The TOE figures show an increase of 30%.

As was mentioned earlier a TOE also reduces CPU cycles by taking this overhead off of the Server and onto the TOE.

Figure 4 shows the CPU savings for each backup run. The TOE figure shows an appreciable drop of 34% in CPU utilisation over the 'standard' Intel card.

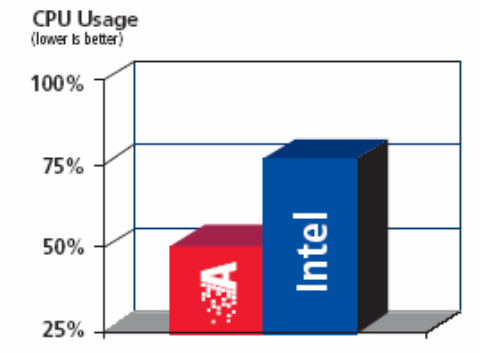

Fig. 4 – CPU Cycles Reduced

The TOE card used is this test, as with all other results quoted in this paper, is an Alacritech Accelerator and a comparison of the data shows Alacritech with a 30% increase in overall backup efficiency with industry standard software and high performance tape devices. Server environments with more efficient data sets could increase their overall backup performance even more.

Results from Computer Associates BrightStor Enterprise Backup, VERITAS NetBackup, CommVault Galaxy and Legato NetWorker also show similar increases in backup performance. With additional system performance tuning and configuration changes, even greater throughput and efficiencies could be achieved.

iSCSI Lowers Storage Consolidation Costs

Since distributed, direct attach storage has a heavy support burden and frequent unbalanced capacity allocation, storage consolidation has become a primary focus. Storage Area Networks (SANs) using expensive Fiber Channel (FC) interconnectivity promise cost reductions for large data centres, but leave a large class of un-served systems due to the excessive

capital and human resource costs SANs place on IT budgets. Managers with large, distributed SCSI installations are searching for economical and easily managed SAN solutions to effectively consolidate storage. One of the latest emerging technologies for storage consolidation is Internet SCSI, or iSCSI.

iSCSI Delivers Excellent TCO

iSCSI enables network storage connectivity at Ethernet prices. While traditional SANs have been cost prohibitive due to expensive FC Host Bus Adapters (HBAs) and switches, iSCSI utilises existing Ethernet infrastructures to access block storage devices on the local network, see figure 5, lowering capital investment requirements. Management of a FC SAN is even more expensive due to training, software and hardware costs needed to run a proprietary storage network.

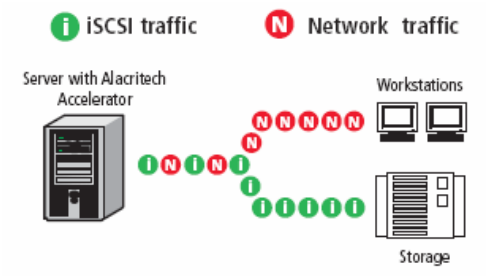

Fig. 5 – iSCSI data movement

On the other hand, Ethernet configuration software and management experience is standardised and used throughout the world. This allows IT managers to consolidate storage and extend the storage capacities of their Windows servers by simply using the Microsoft supplied iSCSI driver to access iSCSI storage. iSCSI is a prime candidate for the consolidation of storage in today's cost-critical environments.

Benefits of iSCSI Deployment on Windows Systems

Due to the relatively high costs of HBA-class products, some organisations may choose to implement iSCSI by simply deploying the Microsoft iSCSI software driver on their Windows-based servers. With this driver, iSCSI becomes just another Windows based service using TCP/IP. Windows servers must now handle existing TCP/IP network traffic needs as well as the additional bandwidth used by iSCSI storage traffic. When looking at the broader issue of efficient data delivery on Windows based servers, the need for a solution to move both network and storage data as quickly and effectively as possible becomes clear. iSCSI HBA products do not address broader connectivity needs, since they only process iSCSI storage traffic, not general networking traffic.

TCP/IP Offload Engine (TOE) Acceleration

The ideal network and storage data delivery solution provides mission critical network and storage availability while optimising server resource usage. TCP/IP Offload Engines (TOEs) accelerate network and storage data delivery across Ethernet environments with unparalleled levels of performance and efficiency. Figure 6 compares the throughput of a TOE enabled iSCSI Ethernet NIC against a standard iSCSI only Intel Pro 1000T NIC.

Integrating a TOE into a Windows server provides the fastest and most effective solution for optimising data delivery and enabling server consolidation in file servers, backup servers, Database servers, video servers and other network applications while also absorbing any iSCSI overhead.

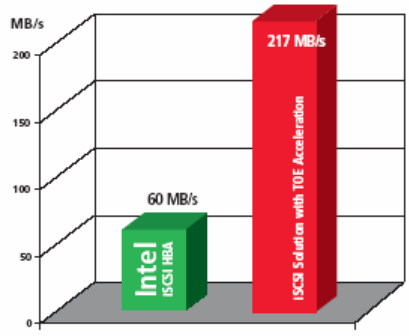

Fig. 6 – iSCSI throughput with TOE

Systems administrators will see dramatic increases in network and storage performance while simultaneously saving server processing cycles, interrupts, and memory bandwidth.