TOE : TCP Offload Engine

Speaker: Chang chin-her
Outline

- Introduction
- TOE over TCP/IP
- The implementation of TOEs
- Performance with TCP Offload
- Conclusion
Today, the growth of Ethernet from 10 Mbit/s to 10 Gbit/s.

I/O is becoming a major bottleneck in delivering high-speed computing.

The performance degradation problem can be particularly severe in Internet SCSI (iSCSI).
Thumb law: For every one bit per second of network data processed, one hertz of CPU processing is required.

TCP/IP offload Engine (TOE) that can reduce the amount of TCP/IP processing handled by microprocessor and server I/O subsystem.
TOE over TCP/IP (1/5)

Figure 1. Comparing standard TCP/IP and TOE enabled TCP/IP stacks
TCP/IP helps ensure reliable, in-order data delivery.

- Reliability
- In-order data delivery.
- Flow control.
- Multiplexing.
Traditional methods to reduce TCP/IP overhead offer limited gains:

- TCP/IP checksum offload
- Large send offload (LSO) = TCP segmentation offload (TSO)
TOE over TCP/ IP(4/ 5)

- TOEs reduce TCP overhead on the host processor
  - CPU interrupt processing
  - Memory copies
    - RDMA
    - zero-copy algorithms
  - Protocol processing
TOE over TCP/IP (5/5)

Figure 2. Transmitting data across the memory bus using a standard NIC
The implementation of TOEs (1/4)

- Network adapters that can handle TCP/IP processing operations.
  - partial versus full offloading

- Extensions to the TCP/IP software stack that offload specified operations to the network adapter.
  - completely transparent to the higher-layer protocols.
Processor-based vs Chip-based

Processor-based:
- expensive
- still can create bottlenecks at 10 Gbps and beyond.
- partial or full offloading
- extensions

Chip-based:
- better performance
- cheap
- partial offloading
Partial versus full offloading

A partial TOE implementation does not handle the following:
- TCP connection setup
- Fragmented TCP segments
- Retransmission time-out
- Out-of-order segments
The host software uses *dynamic* and *flexible* algorithms to determine which connections to offload.

In addition, the host software is responsible for preventing denial of service (DoS) attacks.
Performance with TCP Offload

- Throughput
- CPU utilization
- Latency
There is no standard driver interface for major operating systems and TOE adapters.

The market expects network adapters to be inexpensive.
Memory bandwidth and bus bandwidth are just two of the most critical system dimensions that need to be monitored as systems adopt 10 Gigabit Ethernet.

TCP Segmentation Offload and Zero-Copy function has been included in Linux 2.6.x.