



Network Storage Group Host Products Technology Brief March 20, 2003

ISCSI TERMINOLOGY: TCP/IP

With the iSCSI market looming, we will find ourselves needing to speak a new language. The ISP4010 is an iSCSI and TCP/IP offload engine, but what do TCP and IP do? What does TCP/IP refer to? This brief is intended to put some meaning behind some of the terms used in IP SANs and TCP/IP networks.

TCP/IP Protocol Suite

The TCP/IP suite forms the basis for the Internet and refers to a suite of protocols that actually includes more than just the Transport Control Protocol (TCP) and Internet Protocol (IP). Networking protocols are built up in layers, each being responsible for some distinct aspect of communication. TCP/IP is a four-layer system spanning the upper six layers of the seven layer Open System Interconnect OSI networking model. This table compares the various TCP/IP protocols against the OSI layers and makes an approximate comparison to Fibre Channel layers.

OSI Layer#	Name	Example Protocols in TCP/IP Suite	Analogous Fibre Channel Layers
5-7	Application	ISCSI, Telnet, FTP, e-mail, SCSI-3	SCSI-3, VI, IP
4	Transport	TCP, UDP	FC-4
3	Network	IP, ICMP, IGMP	FC-3
2	Data Link	Ethernet, Token Ring,	FC-2 and most of FC-1

Data Link Layer

The data link layer handles the logical interface to the cable. This is where arbitration for network access occurs. Most local area networks use Ethernet, although some Token Ring still exists. Historically, 10Mbps Ethernet used a shared bandwidth, collision based, access arbitration method and only achieved about 30% bandwidth utilization. Most Gigabit networks today are switch based and achieve bandwidth utilization levels over 90%. This evolution is similar to the progression from Fibre Channel Loop (shared bandwidth) to Fabric (switched).

Addressing at the data link layer is called a "MAC" address or hardware address. This address is assigned at the factory and is unique to each network node. The MAC address is similar to the Fibre Channel World Wide Name. Data can be switched at a local level based on the hardware address.

In the ISP4010, the Ethernet data link layer is implemented as hardware logic in the Media Access Control (MAC) block.

Network Layer

The network layer deals with moving packets around larger networks. Network layer addressing (IP Address) is an abstraction of the MAC address and simplifies the routing of traffic through the network. This routing simplification enables more efficient routing, worldwide addressability and independence from data link layers (e.g. IP over Fibre Channel, Token Ring or Ethernet). The

network layer is also responsible for cutting up frames (fragmentation) and putting them back together (reassembly) if there are links in the path that only support small frames (historically this was true on noisy WAN links).

In the ISP4010, IP is implemented as hardware logic in the IP blocks.

Transport Layer

The transport layer deals with the flow of data between systems. There are two major protocols used at this layer that provide different services: transport control protocol (TCP) and user datagram protocol (UDP). iSCSI is dependent on TCP for its transport.

TCP is a complex protocol that guarantees the delivery of data, in order, to the application layer. Guaranteed delivery is referred to as “reliable” and is accomplished by setting up timers and a system of acknowledgements for each packet that is sent, so the TCP layer can be sure that each packet makes it to the destination. In-order delivery to the application layer is administered by tagging frames with an ordering number, which the receiving side reads and uses to give the application data in the order it was sent. Since TCP has the flexibility to receive frames out of order, allowance have to be made to buffer received data until any missing data is in place. This requirement drives the need for large amounts of memory in iSCSI that does not exist in Fibre Channel. TCP also has methods to detect network congestion and controlling transmission rates. *These bulk data movement features of TCP are implemented in the ISP4010 as hardware logic in the TCP blocks.*

In addition, TCP also includes aspects that are administrative in nature, setting up connections between systems, for instance. *These administrative aspects of TCP, that don't affect the data flow performance, are implemented in firmware that runs in the ISP4010 embedded processors.*

UDP is a simpler transport level protocol than TCP and is not used for iSCSI. UDP is “unreliable,” (sometimes referred to “spray and pray”) meaning UDP doesn't have any mechanism to know if the frames make it to their destination and doesn't care about the order of delivery to the application. The application layer, if it is necessary, must assure delivery of packets (some applications such as video streaming can tolerate a certain amount of lost data). *In the ISP4010, the UDP checksums are implemented in hardware logic as part of the IP block, but remaining UDP functions are left up to the host device.* Some customers for non-storage networking applications may use UDP.

Application Layer

The term iSCSI is often used to describe the entire stack from Ethernet through TCP/IP and iSCSI, but the iSCSI specification only defines an application layer that sits on top of TCP. Within the “TCP/IP suite” there are commonly a number of applications such as: telnet (used for remote login), FTP (file transfer protocol), SMTP (Simple Mail Transfer Protocol) and SNMP (Simple Network Management Protocol). *In the ISP4010, applications run in firmware on the embedded processors. We implement the iSCSI application as well as SNMP, ARP (Address Resolution Protocol, and Ping. Other applications will be implemented in the future.*