iSCSI Storage Networks on Cisco Nexus Switches

What You Will Learn

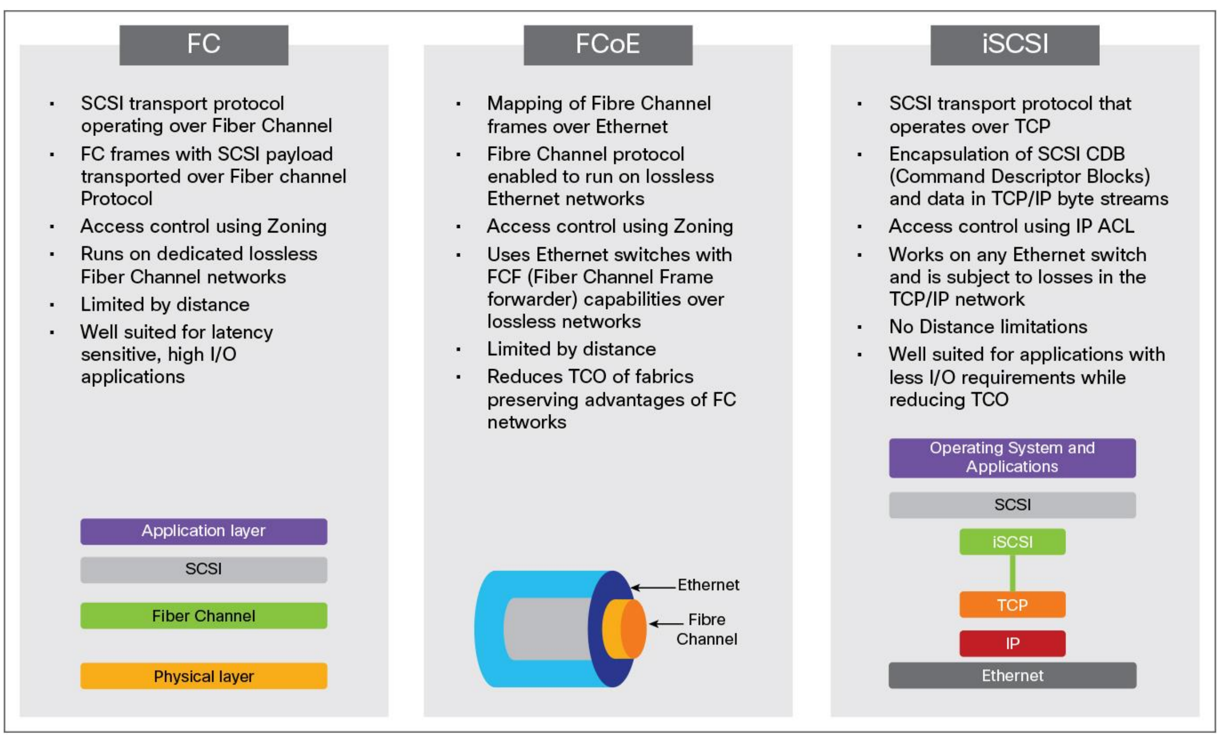
IP and Small Computer System Interface over IP (iSCSI) storage refer to the block access of storage disks across devices connected using traditional Ethernet and TCP/IP networks. iSCSI is an IP-based storage networking standard for linking data storage facilities.

Ethernet’s simplicity, resilience, and cost effectiveness allow it to complement traditional Fibre Channel networks, thereby making IP storage solutions increasingly popular in enterprise deployments. IP-based storage solutions provide low-cost, flexible options for deploying SANs.

This document explains the main factors in iSCSI storage network design and the benefits of Cisco Nexus® switches in building highly efficient data center storage networks.

SCSI is a block-transfer protocol that enables data transfer between various independent peripheral devices and computers. SCSI connects disks in a storage array and the tape drives of a tape library to the servers. Figure 1 shows a comparison of Fibre Channel, Fibre Channel over Ethernet (FCoE), and iSCSI, the most widely deployed block-storage protocols.

Figure 1. Comparison of Block Storage Protocols



iSCSI Overview

iSCSI protocol enables transport of SCSI commands over TCP/IP networks. By transmitting SCSI commands over IP networks, iSCSI can facilitate block-level transfers over the intranet and Internet. The iSCSI architecture is similar to a client-server architecture in which the client initiates an I/O request to the storage target.

The TCP payload of an iSCSI packet contains iSCSI protocol data units (PDUs), all of which begin with one or more header segments followed by zero or more data segments.

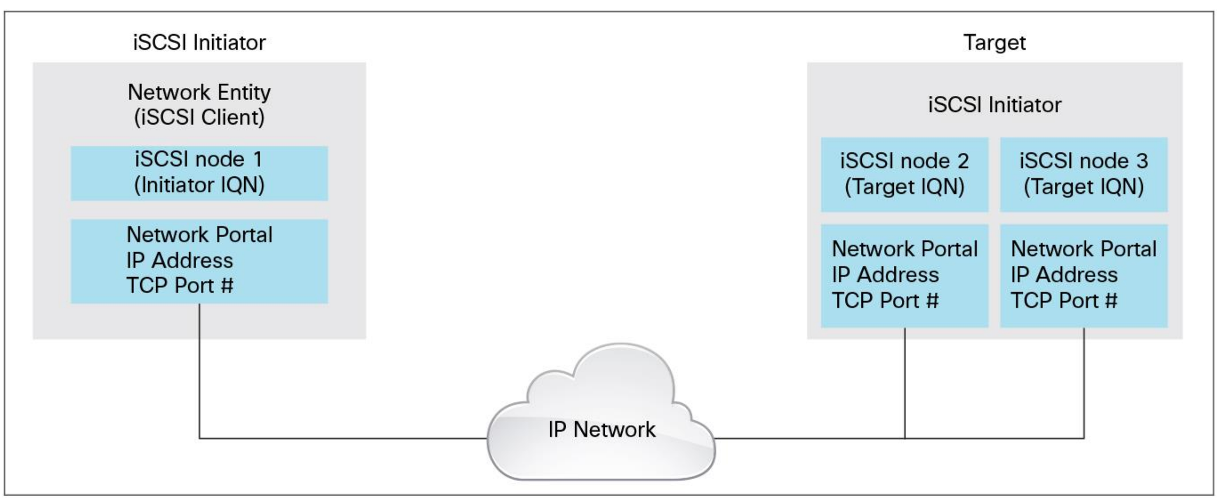
iSCSI Terminology

An iSCSI network consists of multiple devices such as iSCSI initiators and iSCSI targets. Each device has various components associated with it. Some of the iSCSI components that make up an iSCSI storage network are listed here:

- iSCSI name: The iSCSI name is a unique worldwide name (WWN) by which the iSCSI node is known. The iSCSI name uses one of the following formats:
    - iSCSI qualified name (IQN): for example, iqn.1987-05.com.Cisco.00.9f9ccf185aa2508c.taget1
    - Extended unique identifier (EUI): for example, eui.02004567A425678D
- iSCSI node: The iSCSI node represents a single iSCSI initiator or iSCSI target.
- Network entity: The network entity represents a device or gateway that is accessible from the IP network (for example, a host or storage array).
- Network portal: The network portal is a component of a network entity that has a TCP/IP network address that is used by an iSCSI node. A network portal in an initiator is identified by its IP address, and a target is identified by its IP address and its listening TCP port.

A network entity contains one or more iSCSI nodes. The iSCSI node is accessible through one or more network portals. An iSCSI node is identified by its iSCSI name. Figure 2 shows the components of an iSCSI network.

Figure 2. iSCSI Network Components



iSCSI Session

The highest level of an iSCSI communications path is a session that is formed between an iSCSI initiator and an iSCSI target. Two types of sessions are defined in iSCSI:

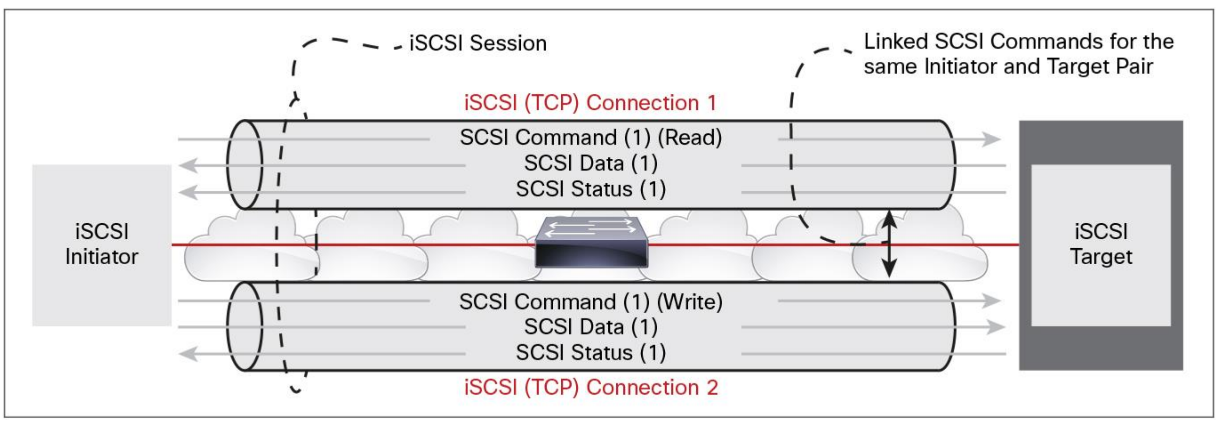
- iSCSI discovery and login session used by the initiator to discover available targets
- General iSCSI session to transfer SCSI data and commands after the login

A session is identified by a session ID (SSID), which consists of initiator (ISID) and target (TSID) components. TCP connections can be added and removed within a session; however, all connections are between the same unique initiator and target iSCSI nodes. Each connection within a session has a unique connection ID (CID).

An iSCSI session is established through the iSCSI login process. This session is used to identify all TCP connections associated with a particular SCSI initiator and target pair. One session can contain one or more TCP connections. The login process is started when the initiator establishes a TCP connection to the desired target either through the well-known port (3260) or a specified target port. The initiator and target can authenticate each other and negotiate a security protocol. During the login phase, numerous attributes are negotiated between the iSCSI initiator and target. Upon successful completion of the login phase, the session enters the full-featured phase.

Figure 3 shows the relationship between an iSCSI session and TCP connections

Figure 3. iSCSI Session



Lossless iSCSI Networks

Ethernet networks are highly susceptible to broadcast and multicast storms, leading to congested networks. The Data Center Bridging (DCB) feature extends lossless capabilities to Ethernet networks, hence providing an option to define networks well suited for storage traffic. When used with traditional Ethernet networks provide lossless iSCSI networks, DCB provides the following features:

- Priority Flow Control (PFC): PFC enables eight virtual queues on a single wire and helps send Pause frames to a single type of traffic instead of all traffic on the wire. This feature helps prevent head-of-the-line blocking while maintaining lossless capabilities. In consolidated networks, in which a mix of traffic is sent on the same wire, PFC helps prioritize the traffic and assign it to either the drop or no-drop class. PFC is useful for iSCSI networks when they are designed for lossless, oversubscribed networks.

- Enhanced Transmission Selection (ETS): ETS provides the capability to allocate bandwidth to each traffic class on the same wire. ETS helps prioritize and optimize the throughput for IP storage traffic, which shares a medium with other traffic on the same link. The guaranteed bandwidth also aids in performance calculations to tune applications during peak traffic.
- Data Center Bridging Exchange (DCBX) Protocol: DCBX is used to exchange all the DCB features across the devices and maintain consistency. DCBX helps ensure consistent quality-of-service (QoS) parameters across the network and servers. The features are advertised to the servers in type-length-value (TLV) format using the Link Layer Discovery Protocol (LLDP).

DCB functions, when enabled for iSCSI networks, help ensure bandwidth, throughput, and performance for the storage traffic.

Traditional iSCSI Networks Compared to Lossless iSCSI Networks

The choice of a traditional or a lossless iSCSI network deployment depends on the application requirements and the oversubscription ratio of the storage network.

Figure 4 lists some of the differences between traditional and lossless iSCSI networks.

Figure 4. Traditional iSCSI Compared to Lossless iSCSI Networks

Traditional iSCSI	Lossless iSCSI
Works on all Ethernet switches	Works only on DCB enabled switches
No Prioritization of iSCSI traffic	Can prioritize iSCSI traffic as "no_drop" using PFC
Bandwidth Shared with all the traffic on wire	Bandwidth allocation to iSCSI traffic using ETS
TCP layer takes care of flow control	Additional Flow control using pause frames at the Ethernet layer along with TCP flow control
Useful for large networks with multiple hops between Initiator and Target	Works well with lossless networks with predictable traffic patterns

IP Storage on Cisco Nexus Switches

As business needs change and additional demands are placed on data centers, today's IT operators and administrators are seeking ways to build efficient storage networks and get extra value from their investments. IP storage networks require careful design because traffic shares the network with other applications. Architects must take into account not only how to handle the existing data that traverses the network, but also how to handle data growth over time.

Common factors to consider in designing a highly effective data center storage network include:

- Ease of running iSCSI storage traffic along with other traffic on the same wire
- Bandwidth allocation and distribution of resources to the iSCSI traffic
- Consistent and predictable performance across the network
- High availability with fault tolerance
- Effective use of switch ports while performing load balancing
- High level of security
- Low-latency and high-throughput fabric

Cisco Nexus switches provide these capabilities and the features needed for an efficient IP storage network.

The Cisco Nexus fabric provides low-latency, resilient, high-throughput switches and a comprehensive feature set well suited for networks that need consistent performance. Cisco Nexus switches along with Cisco® MDS 9000 Family products provide options for combining traditional Fibre Channel SANs with iSCSI storage without the need to change the network or server topologies. The unified ports on the Cisco Nexus 6000 and 5000 Series Switches enable transparent migration between Fibre Channel, FCoE, and iSCSI protocols and hence reduce total cost of ownership (TCO).

Cisco Data Center Network Manager (DCNM) provides a unified tool for managing, maintaining, and troubleshooting Cisco LAN and SAN fabrics. DCNM provides easy-to-use tools for configuring and managing the features of iSCSI storage networks.

Table 1 shows the features provided by Cisco Nexus and Cisco MDS 9000 Family switches and their benefits in iSCSI storage networks.

Table 1. Cisco Nexus and Cisco MDS 9000 Family Platform Features

Feature/Function	Platform	Benefits
10/40/100Gb port speeds	Nexus 7000, 6000/5000, MDS	Line rate and High Speed Networks and Interconnects. Investment Protection with 40/100Gb port speeds
Fabric Extender (FEX) Nexus 2000 Series platform	Nexus 7000, 6000/5000	Higher port densities, Ease of Management, Highly Scalable Networks
Virtual Port Channel (vPC)	Nexus 7000, 6000/5000	Multi-chassis link aggregation and load balancing. Reduced link blocks and higher port resiliency and link level redundancy
Quality of Service (Qos)	Nexus 7000, 6000/5000, MDS	Prioritize the iSCSI Storage Traffic and allocate bandwidth requirements for various types of traffic based on Priority
Fabric Path	Nexus 7000, 6000/5000	Ethernet routing with load balancing of traffic across the fabric. Eliminates the need for Spanning tree and hence the blocked links
VLAN/VSAN Separation	Nexus 7000, 6000/5000, MDS	Logical iSCSI networks with the ability to assign storage traffic to the iSCSI VLAN consistent across the network. Mapping of VLAN and VSAN for Multi-protocol features
Multi-Protocol Support Fiber Channel over Ethernet (FCoE)	Nexus 7000, 6000/5000, MDS	Allows seamless integration and Interoperability with the other storage protocols
Multi-Protocol Support Fiber Channel (FC)	Nexus 6000/5000, MDS	Fiber Channel ports to run the traditional FC protocol for accessing FC Storage devices
Unified Ports (FC, Ethernet/FCoE/iSCSI)	Nexus 6000/5000, Nexus 2000	A single port that can be configured as FC or Ethernet/FCoE/iSCSI port. Multi-protocol flexibility and investment protection
iSCSI Gateway	MDS 9250i	Ability to interop the traditional FC SAN and the iSCSI SAN. Provides FC storage access to the iSCSI hosts

Feature/Function	Platform	Benefits
Virtual Device Context (VDC)	Nexus 7000	Provides ability to logically divide the chassis into multiple switches. Process level high availability and segregation of traffic at the switch level
Dual Supervisor	Nexus 7000, MDS	Processor level High availability and fault tolerance along with line card level redundancy
Security (TrustSec, RBAC, ACL, DHCP)	Nexus 7000, 6000/5000	Ability to build highly secure networks with multiple options for authentication, authorization and avoid snooping attacks
Data Center Network Manager (DCNM)	Nexus 7000, 6000/5000, MDS	Unified Management tool to manage all the Cisco LAN and SAN fabrics

Summary

Cisco Nexus switches are manageable, secure, highly available, and resilient Ethernet switches with a comprehensive feature set. These switches give data center designers and administrators the flexibility to deploy any storage device anywhere in the network. In particular, IP storage devices benefit from the increased cross-sectional bandwidth, segregation of storage traffic, and high availability provided by these switches. These features enable data center designers to deploy IP storage arrays for mission-critical applications. In addition, by enabling the consolidation of I/O traffic, these platforms allow administrators to reduce their overall capital expenditures (CapEx) and operating expenses (OpEx) while taking advantage of both mature and future Ethernet features. This flexibility allows Cisco Nexus switches to dynamically meet the changing business and performance requirements in the data center.

For More Information

- Cisco Unified Fabric products and services
- Cisco Unified Fabric technologies
- Cisco Nexus 7000 Series Ethernet storage directors white paper
- iSCSI protocol details
- FCoE and iSCSI comparison

Americas Headquarters
Cisco Systems, Inc.
San Jose, CA

Asia Pacific Headquarters
Cisco Systems (USA) Pte. Ltd.
Singapore

Europe Headquarters
Cisco Systems International BV Amsterdam,
The Netherlands

Cisco has more than 200 offices worldwide. Addresses, phone numbers, and fax numbers are listed on the Cisco Website at **www.cisco.com/go/offices.**

Printed in USA

C11-734014-00 03/15