



Q&A

Cisco SFS 7000D Series InfiniBand Server Switches

Q. What are the Cisco® SFS 7000D Series InfiniBand switches?

A. A. The Cisco SFS 7000D Series InfiniBand switches are a new series of high-performance InfiniBand switches that deliver next-generation InfiniBand performance and flexibility with autosensing dual-speed DDR/SDR InfiniBand 4X ports that support 20-or 10-Gbps per port. These switches, in common with the SFS-7000P series switches, are designed with best-in-class reliability, high availability, and serviceability features. The Cisco SFS 7000D Series provides the foundation for interconnecting servers with industry-standard DDR (20-Gbps) or SDR (10-Gbps) full duplex, low-latency InfiniBand. The Cisco SFS 7000D Series switches are ideal foundation for high-performance server clusters or for enterprise grids when used inconjunction with Cisco VFrame Server Fabric Virtualization Software.

Table 1 shows the complete list of Cisco SFS 7000D Series switches and components.

Table 1. Cisco SFS 7000D Series Products

Part Numbers	Product Name and Description
SFS7000D-SK9	Cisco SFS 7000D InfiniBand Server Switch, 24 4X InfiniBand ports, (2) power supply/cooling units
PWR-SFS7000P=	Cisco SFS 7000D Power Supply/Cooling FRU
SFS-7012D	Cisco SFS 7012D IB Server Switch - 144 Port DDR – Standard Configuration
SFS-7024D	Cisco SFS 7024D IB Server Switch - 288 Port DDR – Standard Configuration
SFS-7012D-X	Cisco SFS 7012D IB Server Switch - 108 Port DDR – Upgradeable Configuration
SFS-7024D-X	Cisco SFS 7024D IB Server Switch - 156 Port DDR – Upgradeable Configuration
SFSX7012/24D-4X12	Cisco SFS 7012P/7024D InfiniBand 4X 12-Port DDR Line Card
SFS-7012/24D-FM	Cisco SFS 7012D/7024D DDR Switch Fabric Module – No Management
SFS-7012/24D-MM-K9	Cisco SFS 7012D/7024D DDR Switch Fabric Module – With Management
FAN-SFS7012/24=	Cisco SFS 7012/7024 Fan Tray
PWR-SFS7012/24=	Cisco SFS 7012/7024 Power Supply
SFS-7012-RACK=	Cisco SFS 7012 Rack Mount System
SFS-7024-RACK=	Cisco SFS 7024 Rack Mount System

Q. What business benefits does high-performance, low-latency InfiniBand switching provide?

A. The Cisco SFS 7000D Series switches provides a new class of data center infrastructure that interconnects discreet server resources together into a high-performance, 20- or 10-Gbps, low-latency fabric. The increased bandwidth and support for Remote Direct Memory Access (RDMA) and low latency (<4 usec user-space latency) switch is ideal for applications that demand high-bandwidth and low-latency. These technologies enable industry-standard servers to be combined into effective high-performance clusters and grids. In addition to high-performance computing, InfiniBand also offers the performance, security, and control required for utility computing. The Cisco SFS 7000D Series interconnects servers through a high-performance, unified compute fabric, and when combined with Cisco SFS 3000 Series multifabric server switches, allows shared pools of I/O and storage resources that can be mapped together by Cisco VFrame Server Fabric Virtualization Software to deliver virtual “compute services.” Another tangible benefit delivered by the SFS-7000D series InfiniBand switches is providing a future proof compute switching fabric investment. By investing in DDR capable infrastructure today, as server technologies evolve and provide greater PCI-Express performance,

the potential of InfiniBand DDR can be realized without upgrading the entire InfiniBand fabric, thereby delivering significant savings over the life-time of the infrastructure.

Q. Do the Cisco SFS 7000D Series switches support and work with Cisco VFrame Server Fabric Virtualization Software?

A. Yes. The Cisco SFS 7000D Series switches fully support enterprise and research server grid deployments based on Cisco VFrame Server Fabric Virtualization Software. A node does not have to be connected directly to a Cisco SFS 3000 Series Multifabric Server Switch to participate in a server fabric managed by VFrame software. However, to use VFrame, the network design must include a Cisco SFS 3000 Series Multifabric Server Switch in the overall fabric.

Q. What type of cables does the Cisco SFS 7000D Series switches support?

A. The Cisco SFS 7000D Series uses copper CX-4 cables. Additionally, it is possible to construct large InfiniBand switch fabrics by installing pluggable optical modules (POMs) on to the CX-4 ports. See the question about optical connections in the section about building SFS fabrics.

HIGH AVAILABILITY

Q. Do the Cisco SFS 7000D Series switches have redundant power and cooling?

A. Yes. The Cisco SFS 7000D Series comes standard with redundant power supply and cooling modules.

Q. How does redundant fabric management work and what does this mean?

A. An InfiniBand network is a pre-routed switched network, much like the telephone system. This gives InfiniBand many desirable properties including resiliency, predictable performance, and monitoring capabilities. As with any managed, switched network, there is a fabric management entity, called the subnet manager, which oversees the fabric. The subnet manager routes, manages, monitors, and responds to individual node requests in the fabric. Because the subnet manager is critical to operation of the fabric, every Cisco SFS 7000D switch comes with an embedded subnet manager capable of running fabrics with thousands of nodes. To make sure there are no single points of failure in the fabric, the Cisco SFS 7000D Series switches synchronize the subnet manager state amongst each other, such that at any given moment, any of the switches is capable of managing the fabric. Thus, if one of the switches fails or has to reboot, the overall fabric is not disturbed because any of the other switches can take over fabric management. The Cisco SFS-7012D and 7012D do not have an embedded subnet manager and require an external subnet manager, such as the Cisco High-performance Subnet Manager for correct operation.

Q. How does the Cisco SFS 7000D Series switches safely manage system upgrades?

A. The Cisco SFS 7000D Series switches support separate system image loading and activation. For upgrading multiple switches, it is possible to pre-stage new system images, and activate them simultaneously in your fabric during a maintenance window. Additionally, the Cisco SFS 7000D Series supports a default "recovery image" that will be booted if a switch fails to successfully load the target image. If an upgrade fails, the switch will boot in recovery mode, allowing the administrator to revert back to a known good state.

BUILDING SFS FABRICS WITH CISCO SFS 7000D SERIES SWITCHES

Q. How do Cisco SFS 7000D Series switches integrate with Cisco SFS 3000 Series switches to provide Ethernet and Fibre Channel connectivity over InfiniBand?

A. The base fabric for interconnecting the servers in the cluster is typically built out of one or multiple Cisco SFS 7000D Series switches. After that fabric is built, Cisco SFS 3000 Series switches are added to the cluster by attaching directly to Cisco SFS 7000D Series switches through connecting available InfiniBand ports on each switch. The Cisco SFS 3000 Series provides wire-speed I/O connectivity to Ethernet and Fibre Channel networks using Ethernet-to-InfiniBand and Fibre Channel-to-InfiniBand gateways. Each port-to-port connection between a Cisco SFS 7000D Series switch and a Cisco SFS 3000 Series switch provides 20- or 10-Gbps full duplex connectivity for line-rate communications with the entire base InfiniBand server cluster. A server can be connected anywhere

in the InfiniBand fabric and still have full access to Ethernet or Fibre Channel I/O connectivity provided by the Cisco SFS 3000 Series.

Q. How do I use the Cisco SFS 7000D Series to build InfiniBand fabrics larger than the size of a single switch?

A. InfiniBand fabrics, for both high-performance computing and unified I/O grids, which are larger than a single switch, are built with 2-tier designs by connecting multiple switches together. Switches are connected together using the standard ports on the switch; there are no special ports for connecting Cisco SFS 7000D Series switches together. When building these large fabrics, Cisco SFS 7000D Series switches fulfill two different roles in the big fabric: edge switch and core switch. Any one of the Cisco SFS 7000D Series switches can serve as either an edge switch or a core switch, but not both at the same time. Edge switches use some of their switch ports to connect to servers, and the remaining ports to connect to core switches. Each core switch brings the edge switches together in a web-like fashion by connecting to every edge switch. The number of core switches in the fabric and number of connections between an individual core switch and an edge switch will vary based on fabric design. Core switches do not connect to servers directly. Using this 2-tier model, it is possible to build very large InfiniBand fabrics without sacrificing server-to-server performance.

Q. What is the maximum distance between two ports, either switch ports or HCA ports, in a fabric?

A. Currently, the Cisco SFS 7000D Series supports copper cable distances of up to 15 meters (m) port-to-port in SDR mode and up to 10 meters in DDR mode. It is possible for two servers to be much farther apart than that by stringing switches together, but the distance between two ports cannot exceed 15 m over copper links.

Q. Does the Cisco SFS 7000D Series support optical connections?

A. Yes, the Cisco SFS 7000D Series switches can use optical connections through external pluggable optical modules (POMs). A POM enables an individual port on a Cisco SFS 7000D Series switch to use 50-micron or 62.5-micron fiber cables. With 50-micron cables, port-to-port fiber distances can exceed 150 m, and with 62.5-micron cables, port-to-port fiber distances go up to 75 m. One POM is needed for each port involved in the connection; so, one switch-to-switch connection would require two POMs. The POMs can be ordered directly from vendors such as Emcore or Fujitsu.

Q. What is a blocking factor in an SFS InfiniBand fabric?

A. InfiniBand fabrics can be designed to be fully nonblocking, which provides full 20-Gbps bandwidth from every server to every other server in the fabric, or partially blocking. Partially blocking designs require fewer switches to build, but provide less average bandwidth to each server. Blocking factors are typically introduced on the edge switch by using more ports to connect servers than to connect the edge switch into the core switches. On the edge switches, if an equal number of ports is used to connect servers and to connect to the core, the fabric is said to be nonblocking. If twice as many ports are used to connect servers as used to connect the edge switch to the core, the blocking factor is said to be two to one or a 50 percent blocking factor. If three times as many ports are used to connect servers as used to connect the edge switch to the core, the blocking factor is said to be three to one or a 67 percent blocking factor. All single switch designs are nonblocking because Cisco SFS 7000D Series switches are internally fully nonblocking.

Q. Can you combine SDR and DDR in a fabric?

A. Yes. The Cisco SFS 7000D Series DDR InfiniBand switches have autosensing ports that will negotiate port speed to be double data rate (DDR–20-Gbps) or single data rate (SDR–10-Gbps).

Q. What is the impact of mixing SDR and DDR on a fabric's blocking factor?

A. The Cisco SFS 7000D Series DDR InfiniBand switches support autosensing of SDR or DDR attached devices and also supports the capability to forward packets between SDR and DDR connections that provides better utilization of resources and investment protection for existing InfiniBand 4X devices. To exemplify how this impacts blocking factor, the 24-port SFS 7000D can support 16 4X SDR server connections with 8 4X DDR uplinks to the InfiniBand fabric core and offer non-blocking capability while offering a higher server-to-fabric density than a pure SDR switch.

Q. What effect does the blocking factor have on SFS InfiniBand fabric performance?

A. The blocking factor has absolutely no effect on fabric latency or ability to access I/O resources. However, blocking factor does affect worst-case throughput from server to server. Servers on a 50 percent blocking factor fabric will have half of the maximum bandwidth, 20 Gbps, or 10 Gbps available in the worst case. Servers on a 67 percent blocking factor fabric will on average have one third of the maximum bandwidth or 6.67 Gbps available in the worst case. The reason to consider the worst case is there are many fabric-usage factors that can affect the actual blocking experienced by a server. For example, all communications between servers attached to the same edge switch will be fully nonblocking because, as mentioned in the previous answer, the blocking factor is introduced between the edge switches and the core switches. In addition, blocking will only be experienced if there is contention on the links from the edge switch to the core switch. If other servers on the same edge switch are using the network at less than maximum capacity, it is possible for a server to still have full bandwidth in the fabric because there will not be contention for the uplinks from the edge to the core. Thus, the blocking factor will be the worst-case bandwidth experience by a server node.

Q. How does one calculate the blocking factor?

A. The blocking factor is calculated by dividing the aggregate throughput of ports on an edge switch used for uplinks to core switches by the aggregate throughput of ports used to connect servers. Thus, if on a 24-port Cisco SFS 7000 used as an edge switch, 16 ports are used to connect servers and 8 ports are used to uplink to the core all at DDR speeds, the blocking factor is $8/16 = 50\%$. All single-switch installations are fully nonblocking because all Cisco SFS 7000D Series switches are fully nonblocking internally.

FABRIC MANAGEMENT

Q. How do I manage the Cisco SFS 7000D Series?

A. Management of the Cisco SFS 7000D Series switches is integrated into the management suite for all Cisco SFS products. Management options include the Java-based Element Manager GUI, the Web-based chassis manager GUI, and a command-line interface (CLI). For management security, the Web-based management supports HTTPS and the CLI supports Secure Shell (SSH) Protocol v2. Additionally, the switches fully support SNMPv2 and v3 MIBs for integration with other network management utilities. The Cisco SFS 7000D Series management features proactive notification of system events through syslog and SNMP traps. The administrator can configure these active messages to many levels, ranging from frequent alerts to alerts for critical failures only.

Q. What fine-grained user management is supported on the Cisco SFS 7000D Series?

A. The Cisco SFS 7000D Series management supports fine-grained user roles with access control lists (ACLs). Each user can be assigned a different role that grants ACL privileges to view and modify different settings on the switch. These ACLs are strictly enforced, making it possible for diverse groups within an organization to safely manage and monitor a Cisco SFS 7000D Series switch.

Q. What management security features does the Cisco SFS 7000D Series support?

A. The Cisco SFS 7000D Series switches are hardened secure devices. The switches support SSHv2 and HTTPS for secure remote management. Additionally, the switches support SNMPv3 encryption. For user control, the switches support RADIUS for external authentication. Each user is assigned a role with a fine-grained ACL. These ACLs are strictly enforced for each user, making it safe for diverse groups within an organization to manage these switches.

Q. Is it possible to observe the entire fabric and its topology through the Cisco SFS 7000D Series management suite?

A. Yes. It is possible to visually view and monitor the entire InfiniBand fabric from a Cisco SFS 7000D Series switches management connection. Fabric information such as number of switches, servers, error rates across the fabric, and general performance information are all available in the topology viewer in the Cisco SFS 7000D Series switches management tools. Additionally, information about overall fabric health, and individual switch and server health is available. This information can be viewed in the Cisco SFS management tools, or accessed through SNMP or CLI-based scripts.

Q. What type of monitoring is supported on the Cisco SFS 7000D Series switches?

A. The Cisco SFS 7000D Series has many different monitoring options. It is possible to get gather performance information on a port-by-port basis in real time. This information includes raw performance data as well as error rates and packet statistics. This data can be expressed as updated text in tables or as live moving graphs over time. It is also possible to graph and compare a number of different data sets at the same time. Additionally, this data can be monitored at a high level to answer questions like, "I want to monitor all ports between server 1 and server 2 for performance and error rates." The administrator can either observe these monitored connections manually, or set triggers that will proactively notify the administrator of certain conditions such as transmission rates dropping below a threshold or error rates exceeding a threshold.

Q. Do I need any additional fabric management software to operate my Cisco SFS InfiniBand fabric?

A. To take advantage of advanced server virtualization and network provisioning, unified I/O fabrics, and enterprise grid capability, Cisco VFrame Server Virtualization Software must be added to the SFS fabric. For extremely large SFS fabrics, in excess of 1500 servers, the Cisco SFS High Performance Subnet Manager must be added to the SFS fabric for performance and advanced management capabilities.

SERVICE AND WARRANTY

Q. What is the warranty for the Cisco SFS 7000D Series switches?

A. Cisco SFS 7000D Series switches come with the Cisco Limited Lifetime Hardware Warranty.

This warranty automatically comes with the purchase of eligible Cisco SFS products, free of charge. Ongoing software updates are also available to customers on Cisco.com free of charge. For specific details on the Limited Lifetime Hardware Warranty, visit:

[http://www.cisco.com/univercd/cc/td/doc/es_inpk/lh2cen .htm](http://www.cisco.com/univercd/cc/td/doc/es_inpk/lh2cen.htm)

Q. What types of service and support are available for the Cisco SFS 7000D Series switches?

A. In addition to worldwide, 24-hour Cisco Technical Assistance Center (TAC) support, a full complement of implementation-based services are available at this time for the Cisco SFS 7000D Series switches.

TECHNICAL SUPPORT SERVICE

Technical Support Service is available through Cisco SMARTnet[®] support and Cisco SMARTnet Onsite. Cisco SMARTnet support augments the resources of the operations staff by providing access to expertise, both online and on the telephone, and a range of hardware Advance Replacement options. Cisco SMARTnet Onsite complements the hardware Advance Replacement feature by adding the services of a field engineer-services that can be critical for those locations where staffing is insufficient or unavailable to perform parts replacement activities. For more information about Cisco SMARTnet support, visit:

http://www.cisco.com/en/US/products/svcs/ps3034/ps2827/ps2978/serv_group_home.html

ADVANCED SERVICES

Total Implementation Solutions (TIS) offers a full range of implementation solutions including project management, project engineering, configuration, staging, and rollout coordination, to help ensure correct installation and deployment. Configuration services now include development and verification of configuration for intelligent services such as Cisco VFrame Server Virtualization Software. For more information about Total Implementation Solutions, visit: <http://www.cisco.com/go/services>

PRODUCT AND CONTACT INFORMATION

- Q. Where can I find technical and product specifications and other additional information about the Cisco SFS 7000D Series switches?**
- A.** For product literature including data sheets and product specifications, see the Server Networking and Virtualization Website at <http://www.cisco.com/en/US/products/ps6418/index.html>.
- Q. What are the part numbers for the Cisco SFS Series products?**
- A.** The part numbers and ordering information for the new Cisco SFS Series products is available at <http://www.cisco.com/en/US/products/ps6418/index.html#products>



Corporate Headquarters
Cisco Systems, Inc.
170 West Tasman Drive
San Jose, CA 95134-1706
USA
www.cisco.com
Tel: 408 526-4000
800 553-NETS (6387)
Fax: 408 526-4100

European Headquarters
Cisco Systems International BV
Haarlerbergpark
Haarlerbergweg 13-19
1101 CH Amsterdam
The Netherlands
www-europe.cisco.com
Tel: 31 0 20 357 1000
Fax: 31 0 20 357 1100

Americas Headquarters
Cisco Systems, Inc.
170 West Tasman Drive
San Jose, CA 95134-1706
USA
www.cisco.com
Tel: 408 526-7660
Fax: 408 527-0883

Asia Pacific Headquarters
Cisco Systems, Inc.
168 Robinson Road
#28-01 Capital Tower
Singapore 068912
www.cisco.com
Tel: +65 6317 7777
Fax: +65 6317 7799

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

Argentina • Australia • Austria • Belgium • Brazil • Bulgaria • Canada • Chile • China PRC • Colombia • Costa Rica • Croatia • Cyprus • Czech Republic
Denmark • Dubai, UAE • Finland • France • Germany • Greece • Hong Kong SAR • Hungary • India • Indonesia • Ireland • Israel • Italy
Japan • Korea • Luxembourg • Malaysia • Mexico • The Netherlands • New Zealand • Norway • Peru • Philippines • Poland • Portugal
Puerto Rico • Romania • Russia • Saudi Arabia • Scotland • Singapore • Slovakia • Slovenia • South Africa • Spain • Sweden
Switzerland • Taiwan • Thailand • Turkey • Ukraine • United Kingdom • United States • Venezuela • Vietnam • Zimbabwe

Copyright © 2006 Cisco Systems, Inc. All rights reserved. CCSP, CCVP, the Cisco Square Bridge logo, Follow Me Browsing, and StackWise are trademarks of Cisco Systems, Inc.; Changing the Way We Work, Live, Play, and Learn, and iQuick Study are service marks of Cisco Systems, Inc.; and Access Registrar, Aironet, BPX, Catalyst, CCDA, CCDP, CCIE, CCIP, CCNA, CCNP, Cisco, the Cisco Certified Internetwork Expert logo, Cisco IOS, Cisco Press, Cisco Systems, Cisco Systems Capital, the Cisco Systems logo, Cisco Unity, Enterprise/Solver, EtherChannel, EtherFast, EtherSwitch, Fast Step, FormShare, GigaDrive, GigaStack, HomeLink, Internet Quotient, IOS, IP/TV, iQ Expertise, the iQ logo, iQ Net Readiness Scorecard, LightStream, Linksys, MeetingPlace, MGX, the Networkers logo, Networking Academy, Network Registrar, Packet, PIX, Post-Routing, Pre-Routing, ProConnect, RateMUX, ScriptShare, SlideCast, SMARTnet, The Fastest Way to Increase Your Internet Quotient, and TransPath are registered trademarks of Cisco Systems, Inc. and/or its affiliates in the United States and certain other countries.

All other trademarks mentioned in this document or Website are the property of their respective owners. The use of the word partner does not imply a partnership relationship between Cisco and any other company. (0601R)