

Intelligent Traffic Services with the Cisco MDS 9000 Family Modules

The second-generation Cisco® MDS 9000 Family modules offer extensive tools for traffic engineering.

INTRODUCTION

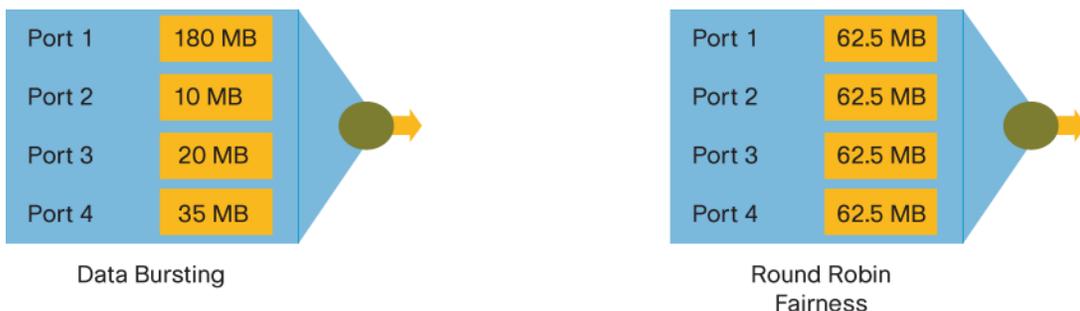
As storage networks continue to mature and grow in scale, the level of service provided by those networks must scale as well. The demands of today’s high-performance storage subsystems and the growing number of servers in the data center require a storage area network (SAN) that can provide robust performance and on-demand flexibility while still maintaining an attractive cost structure. Efficient utilization of network bandwidth is crucial in building a large-scale enterprise data center SAN. Even as 4-Gbps and 10-Gbps Fibre Channel interfaces become available, most servers, storage subsystems, and applications continue to run at 1 Gbps or 2 Gbps. Because of this fact, it is very important to be able to effectively allocate the network bandwidth to accommodate actual I/O throughput at each interface dynamically. Cisco MDS 9000 Family intelligent oversubscription technology, using a flexible crossbar-based modular design, allows for a mix of price and performance to meet any SAN requirements.

OVERSUBSCRIPTION

With the introduction of the Cisco MDS 9000 Family platform, Cisco Systems® changed the way that SANs were designed and deployed. One of the most compelling design options that it offers is the choice between a full-rate module and an oversubscribed module. This choice allows the best mix of performance (storage, tape, and Inter-Switch Links [ISLs]) and price per port (large number of servers). Because 95 percent of Cisco MDS customers purchase oversubscribed modules, these modules have pushed the cost of deploying a SAN down while maintaining critical application performance.

One characteristic that makes oversubscribed modules ideal for most data center servers is their ability to respond to line-rate bursts of data. The first-generation 32-port module (DS-X9032) for the Cisco MDS 9000 Family shared 2.5 Gbps of bandwidth per four-port group. This allowed any device to burst at high data rates and still have performance available for the other three ports in the group (Figure 1). The 32-port module also instituted a round robin mechanism that prevents any one device from taking all the bandwidth from another device. This allows for complete fairness if the aggregate of the four devices exceeds the available 2.5 Gbps of bandwidth available.

Figure 1. Data Burst and Fairness Capability in Oversubscribed Modules



The 32-port module's effectiveness in allowing burst traffic and also helping ensure fairness coupled with a significant cost per port savings made the oversubscribed module attractive within the data center. If applications with different I/O performances are mixed in the same oversubscribed port group, 2.5 Gbps of total bandwidth will be efficiently shared and utilized. However, depending on business requirements, sometimes setting bandwidth limits on certain interfaces to maximize or minimize I/O burst throughput is better, and this in turn somewhat limits designing the most effective possible large-scale oversubscribed storage network with a first-generation module.

The second-generation Cisco MDS 9000 Family modules include two additional modules capable of oversubscription to allow more flexibility and cost effectiveness within the SAN: the DS-X9124 (24-port) and the DS-X9148 (48-port). Each module can be used in an oversubscribed fashion, as shown in Table 1.

Table 1. Oversubscription Ratios per Module

Part Number	1 Gbps Fibre Channel	2 Gbps Fibre Channel	4 Gbps Fibre Channel
DS-X9032	1.6:1	3.2:1	—
DS-X9124	1:1	1:1	2:1
DS-X9148	1:1	2:1	4:1

In addition to supporting data burst capability and round robin fairness, the second-generation Cisco MDS modules also support a new traffic management feature called Bandwidth Allocation.

BANDWIDTH ALLOCATION

One of the most significant enhancements to the second-generation modules is their ability to do Bandwidth Allocation. Bandwidth Allocation allows a port to guarantee a certain rate of throughput to the end device. This is extremely useful on oversubscribed modules because it allows any port to perform like a line-rate interface. Used in conjunction with round robin fairness and data bursting capabilities, bandwidth allocation provides the capability to completely manage end-device performance.

Allocation of bandwidth is defined at the port level within a port group. A port group is defined by a series of ports that share back-end bandwidth. The 32-port module has a four-port port group. That is to say, four ports share the back-end bandwidth to the backplane of the chassis. Port groups in the second-generation modules are no longer composed of four ports. The increase in the number of ports in a port group, as show in Table 2, allows the most flexibility in allocating dedicated bandwidth while still maintaining burst capability for other ports in the group.

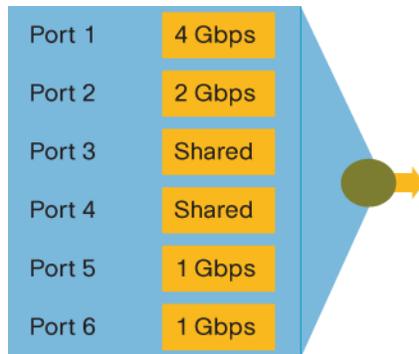
Table 2. Port Group Size per Module

Part Number	Port Group Size	Group Bandwidth
DS-X9032	4	2.5 Gbps
DS-X9124	6	12 Gbps
DS-X9148	12	12 Gbps

Within a port group, port speed can be set to dedicated bandwidth or shared bandwidth. An individual port can be configured as either 1 Gbps dedicated, 2 Gbps dedicated, 4 Gbps dedicated, or shared bandwidth. Bandwidth allocation is independent of the configured speed of the interface.

In Figure 2 the Fibre Channel interface speeds can be configured to 1 Gbps, 2 Gbps, or 4 Gbps, yet the bandwidth dedicated to them might be above or below that. With 12 Gbps of bandwidth available to a port group, the example explicitly reserves the amount of bandwidth required for ports 1, 2, 5, and 6. This allows the remaining 4 Gbps of bandwidth to be shared on ports 3 and 4 and allows either port to burst to 4 Gbps.

Figure 2. Bandwidth Allocation in a Six-Port Port Group



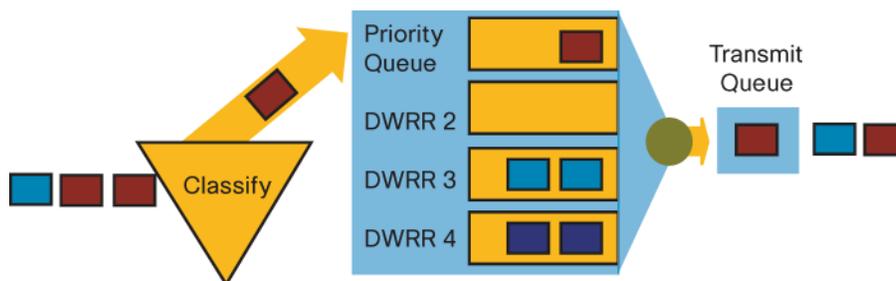
Bandwidth Allocation also allows ISLs to be used on an oversubscribed module. Because full line rate can be reserved for any given port, a port within the port group can be set to full 4-Gbps speed, guaranteeing full performance for saturated ISLs in the network. This feature makes a chassis with all oversubscribed modules a viable configuration in many data centers.

QUALITY OF SERVICE

The adoption of a service provider model within the enterprise data center has added the requirement of tiered service levels among storage arrays. Adding tiers enables the possibility of adding service-level agreements (SLAs). An important component to an SLA is the priority of one application or storage service over another. If a true quality-of-service (QoS) mechanism is enabled in the Cisco MDS 9000 Family module, applications can be distinguished by traffic priority.

The flexibility of QoS on the Cisco MDS 9000 Family module enables a simple mechanism to prioritize applications both within a switch and across the entire network. Using the proven mechanism of Deficit Weighted Round Robin (DWRR), users can determine not only application priority but also the weight of that priority. The Cisco MDS 9000 Family supports four QoS queues, the first of which is absolute priority. The remaining three queues are user definable (see Figure 3).

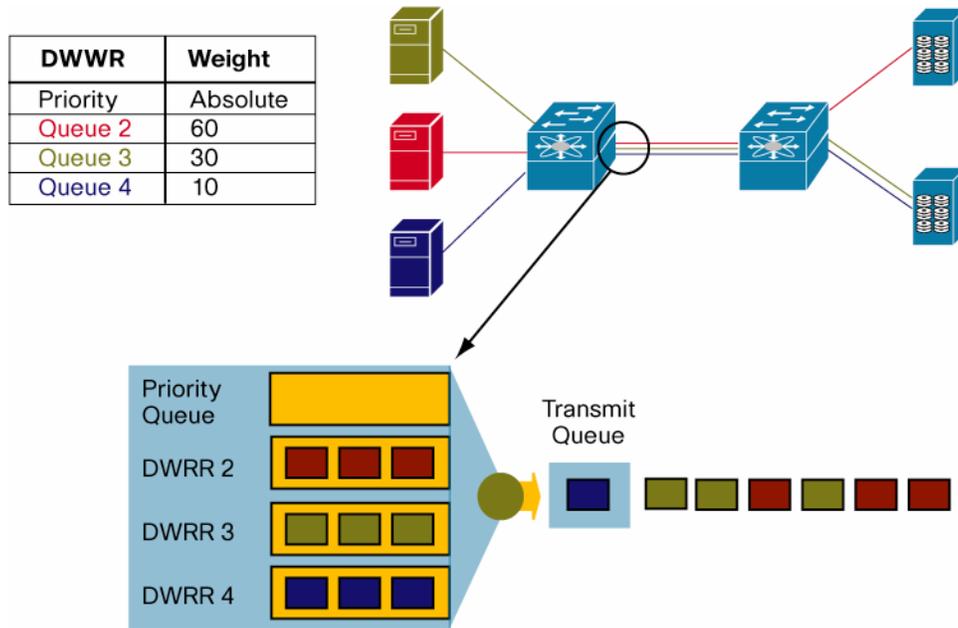
Figure 3. Fibre Channel Port Ingress QoS



Each of the three user-definable queues maintains a separate weight, with the total weight of the three queues equaling 100 percent. Although the fourth queue (absolute priority) is used for network control traffic such as Fabric Shortest Path First (FSPF) updates or zoning changes, the three user-definable queues are used for end devices. To help ensure simple management and easy operation, QoS is enabled on a per-zone basis. This simple method allows for easy QoS configuration either when a zone is created or when QoS is enabled.

Figure 4 shows a typical scenario in which QoS is applied within a switch and across ISLs, helping ensure critical application performance during network congestion. Several devices share the same ISL to communicate across the network. If the ISL becomes congested, QoS begins to prioritize traffic. Whether servers are talking to the same storage or different subsystems, QoS prioritizes output order, thereby in this case helping ensure that the traffic for queue 2 and queue 3 receives preferential treatment to bandwidth over queue 4.

Figure 4. QoS Applied in Switch and Across ISLs



QoS is useful in many different scenarios. One important scenario is when a network failure occurs. A failure could be the result of a network outage or simply a storage subsystem interface failure that forces failover traffic to another interface. In both situations bandwidth in the network can become constrained at crucial points. Enabling QoS prior to a failure allows users to determine which applications should have priority.

CONCLUSION

The evolving storage network demands a feature set that enables users to manage traffic from end to end. The Cisco MDS 9000 Family of switches and directors offers the features to support a dynamic and responsive data center. The second-generation modules for the Cisco MDS 9000 Family platform empower users to take control of the network and offer true traffic engineering.

With the swift adoption of oversubscription in the switch as a cost reducer, Cisco is continuing its industry-leading flexibility between cost and performance. Bandwidth allocation makes the choice between cost and performance simple by offering the flexibility to allow line-rate performance at oversubscribed cost. Completing the traffic-management strategy by enabling QoS within the network helps ensure that SLAs and application requirements are met during normal operation and in times of heavy network load.

Together, bandwidth allocation and QoS, coupled with the rest of the award-winning Cisco feature set, offer a feature-rich switch while providing low-cost connectivity.

**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 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)

