Administrator's Guide

48-Port Gigabit Ethernet Top-of-Rack Switch

AS4600-54T

DCSS Software v2.1.10.0

Administrator's Guide

AS4600-54T 52-Port Gigabit Ethernet Top-of-Rack Switch

with 48 10/100/1000BASE-T (RJ-45) Ports,
4 10GBASE SFP+ Ports,
2 Power Supply Units,
and 1 Fan Tray (F2B and B2F Airflow)

E092013/ST-R02
149100000218A

Contents

Contents

Contents

Figures

Figures

Tables

Tables

How to Use This Guide

This guide describes the switch software and provides configuration examples for many of its features. This switch is ideal for Layer 2/3 switching solutions in the data center. To deploy it effectively and ensure trouble-free operation, first read the relevant sections in this guide so that you are familiar with all of its software features.

Who Should Read This Guide?

This guide is for network administrators who are responsible for operating and maintaining network equipment. The guide assumes a basic working knowledge of LANs (Local Area Networks), the Internet Protocol (IP), and Simple Network Management Protocol (SNMP).

How This Guide is Organized

This guide contains the following chapters:

- Chapter 1: "DCSS Features," on page 19 provides an overview of the features that DCSS software supports.
- Chapter 2: "Getting Started," on page 31 contains information about the boot process, initial system configuration and user interface access.
- Chapter 3: "Configuring Switch Management," on page 41 describes how to perform typical system maintenance tasks, such as installing a new image.
- Chapter 4: "Security Features," on page 57 provides information about management security and network security features.
- Chapter 5: "Configuring Switching," on page 69 describes how to manage and monitor some of the key Layer 2 switching features, such as VLANs and spanning tree protocol (STP).
- Chapter 6: "Configuring Data Center Features," on page 105 contains information about the Data Center Bridging Exchange protocol (DCBX), Priority Flow Control (PFC), and other DCSS features that are of particular importance in data center environments.
- Chapter 7: "Configuring Quality of Service," on page 123 describes how to manage the DCSS software ACLs, and how to configure the Differentiated Services and Class of Service features.

Related Documentation

This guide focuses on general administrative tasks.

- *CLI Reference* — For information on configuring the switch through the Command Line Interface (CLI).
- *MIB Reference* — For information on the Management Information Base (MIB).
- *Release Notes* — For information on key changes to the software in the latest release.
- *Installation Guide* — For information on how to install the switch.
- *Quick Start Guide* — For a brief description of setting up the system.

- *Safety and Regulatory Information* — For all safety information and regulatory statements.

Document Conventions

The following conventions may be used in this document:

Symbol	*Description*	*Example*
Blue Text	Hyperlinked text.	See Chapter 1: "DCSS Features," on page 19.
`courier font`	Command or command-line text	`show network`
`italic courier font`	Variable value. You must replace the italicized text with an appropriate value, which might be a name or number.	*`value`*
[] square brackets	Optional parameter.	`[value]`
`{}` curly braces	Required parameter values. You must select a parameter from the list or range of choices.	`{choice1 | choice2}`
| Vertical bar	Separates the mutually exclusive choices.	`choice1 | choice2`
`[{}]` Braces within square brackets	Optional parameter values. Indicates a choice within an optional element.	`[{choice1 | choice2}]`

Revision History

This section summarizes the changes in each revision of this guide.

Revision	*Date*	*Change Description*
DCSS Software v2.1.10.0	09/01/13	**Added:** • "FTP File Update" on page 19 • "Malicious Code Detection" on page 20 • "Core Dump" on page 21 • "IP Source Address Guard" on page 23 • "Track LAG Member Port Flaps" on page 24 • "Source Specific Multicasting (SSM)" on page 27 • "Control Packet Flooding" on page 27 • "Flooding to mRouter Ports" on page 27 • "IGMP Snooping Querier" on page 27 • "Remote Switched Port Analyzer (RSPAN)" on page 28 • "Downloading a Core Dump" on page 53 • "Configuring RSPAN" on page 84 • "Configuring VLAN Based Mirroring" on page 86 • "Configuring Flow Based Mirroring" on page 86 • "IGMPv3/SSM Snooping" on page 98
DCSS Software v2.0.2.0	3/25/13	Initial release

1 DCSS Features

Management

This section describes the management features DCSS software supports. For additional information and configuration examples for some of these features, see Chapter 3: “Configuring Switch Management,” on page 41.

Multiple Management Options

You can use the following methods to manage the switch:

- Use a telnet client, SSH client, or a direct console connection to access the CLI. The CLI syntax and semantics conform as much as possible to common industry practice.
- Use a network management system (NMS) to manage and monitor the system through SNMP. The switch supports SNMP v1/v2c/v3 over the UDP/IP transport protocol.

Management of Basic Network Information

The DHCP client on the switch allows the switch to acquire information such as the IP address and default gateway from a network DHCP server. You can also disable the DHCP client and configure static network information. Other configurable network information includes a Domain Name Server (DNS), hostname to IP address mapping, and a default domain name.

The switch also includes a DHCPv6 client for acquiring IPv6 addresses, prefixes, and other IPv6 network configuration information.

Dual Software Images

The switch can store up to two software images. The dual image feature allows you to upgrade the switch without deleting the older software image. You designate one image as the active image and the other image as the backup image.

File Management

You can upload and download files such as configuration files and system images by using FTP, TFTP, Secure FTP (SFTP), or Secure Copy (SCP). Configuration file uploads from the switch to a server are a good way to back up the switch configuration. You can also download a configuration file from a server to the switch to restore the switch to the configuration in the downloaded file.

FTP File Update

This feature adds support for file transfers using FTP protocol. FTP Transfers are supported over both IPv4 and IPv6. Upon failure of a FTP transfer operation, a LOG message is sent to the logging component, the initiating application is notified of the failure, and any partial or temporary files for the transfer are removed from persistent memory.

Malicious Code Detection

This feature provides a mechanism to detect the integrity of the image, if the software binary is corrupted or tampered with while the end user attempts to download the software image to the switch. This release addresses this problem by using digital signatures to verify the integrity of the binary image. It also provides flexibility to download a digitally signed configuration script and verify the digital signature to ensure the integrity of the downloaded configuration file.

Automatic Installation of Firmware and Configuration

The Auto Install feature allows the switch to upgrade to a newer software image and update the configuration file automatically during device initialization with limited administrative configuration on the device. The switch can obtain the necessary information from a DHCP server on the network.

Warm Reboot

The Warm Reboot feature reduces the time it takes to reboot the switch thereby reducing the traffic disruption in the network during a switch reboot. For a typical switch, the traffic disruption is reduced from about two minutes for a cold reboot to about 20 seconds for a warm reboot.

SNMP Alarms and Trap Logs

The system logs events with severity codes and timestamps. The events are sent as SNMP traps to a trap recipient list.

CDP Interoperability through ISDP

Industry Standard Discovery Protocol (ISDP) allows the switch to interoperate with Cisco® devices running the Cisco Discovery Protocol (CDP). ISDP is a proprietary Layer 2 network protocol which inter-operates with Cisco network equipment and is used to share information between neighboring devices (routers, bridges, access servers, and switches).

Remote Monitoring (RMON)

RMON is a standard Management Information Base (MIB) that defines current and historical MAC-layer statistics and control objects, allowing real-time information to be captured across the entire network. The data collected is defined in the RMON MIB, RFC 2819 (32-bit counters), RFC 3273 (64-bit counters), and
RFC 3434 (High Capacity Alarm Table).

Statistics Application

The statistics application collects the statistics at a configurable time interval. The user can specify the port number(s) or a range of ports for statistics to be displayed. The configured time interval applies to all ports. Detailed statistics are collected between the specified time range in date and time format. The time range can be defined as having an absolute time entry and/or a periodic time. For example, a user can specify the statistics to be collected and displayed between 9:00 12 NOV 2011 (START) and 21:00 12 NOV 2011 (END) or schedule it on every MON, WED and FRI 9:00 (START) to 21:00 (END).

The user receives these statistics in a number of ways as listed below:

- User requests through CLI for a set of counters.
- User can configure the device to display statistics using syslog or email alert. The syslog or email alert messages are sent by statistics application at END time.

The statistics are presented on the console at END time.

Log Messages

The switch maintains in-memory log messages as well as persistent logs. You can also configure remote logging so that the switch sends log messages to a remote log server. You can also configure the switch to send log messages to a configured SMTP server. This allows you to receive the log message in an e-mail account of your choice. Switch auditing messages, CLI command logging, and SNMP logging can be enabled or disabled.

System Time Management

You can configure the switch to obtain the system time and date through a remote Simple Network Time Protocol (SNTP) server, or you can set the time and date locally on the switch. You can also configure the time zone and information about time shifts that might occur during summer months.

Note: The manually-configured local clock settings are not retained across a system reset.

Source IP Address Configuration

Syslog, TACACS, SNTP, sFlow, SNMP Trap, RADIUS, and DNS Clients allow the IP Stack to select the source IP address while generating the packet. This feature provides an option for the user to select an interface for the source IP address while the management protocol transmits packets to management stations. The source address is specified for each protocol.

Multiple Routing Tables

On this system, local and default IPv4 routes for the service port and network port are installed in routing tables dedicated to each management interface. Locally-originated IPv4 packets use these routing tables when the source IP address of the packet matches an address on one of these interfaces. This feature allows the IP stack to use default routes for different interfaces simultaneously.

Core Dump

The core dump feature provides the ability to retrieve the state from a crashed box such that it can be then loaded into a debugger and have that state re-created there.

Security

This section describes the security features DCSS software supports. For additional information and configuration examples for some of these features, see Chapter 4: “Security Features,” on page 57.

Configurable Access and Authentication Profiles

You can configure rules to limit access to the switch management interface based on criteria such as access type and source IP address of the management host. You can also require the user to be authenticated locally or by an external server, such as a RADIUS server.

Password-Protected Management Access

Access to the CLI and SNMP management interfaces is password protected, and there are no default users on the system.

Strong Password Enforcement

The Strong Password feature enforces a baseline password strength for all locally administered users. Password strength is a measure of the effectiveness of a password in resisting guessing and brute-force attacks. The strength of a password is a function of length, complexity and randomness. Using strong passwords lowers overall risk of a security breach.

MAC-Based Port Security

The port security feature limits access on a port to users with specific MAC addresses. These addresses are manually defined or learned on that port. When a frame is seen on a locked port, and the frame source MAC address is not tied to that port, the protection mechanism is invoked.

RADIUS Client

The switch has a Remote Authentication Dial In User Service (RADIUS) client and can support up to 32 authentication and accounting RADIUS servers.

TACACS+ Client

The switch has a TACACS+ client. TACACS+ provides centralized security for validation of users accessing the switch. TACACS+ provides a centralized user management system while still retaining consistency with RADIUS and other authentication processes.

Dot1x Authentication (IEEE 802.1X)

Dot1x authentication enables the authentication of system users through a local internal server or an external server. Only authenticated and approved system users can transmit and receive data. Supplicants are authenticated using the Extensible Authentication Protocol (EAP). Also supported are PEAP, EAP-TTL, EAP-TTLS, and EAP-TLS.

DCSS software supports RADIUS-based assignment (via 802.1X) of VLANs, including guest and unauthenticated VLANs. The Dot1X feature also supports RADIUS-based assignment of filter IDs as well as MAC-based authentication, which allows multiple supplicants connected to the same port to each authenticate individually.

Denial of Service

The switch supports configurable Denial of Service (DoS) attack protection for many different types of attacks.

DHCP Snooping

DHCP Snooping is a security feature that monitors DHCP messages between a DHCP client and DHCP server. It filters harmful DHCP messages and builds a bindings database of (MAC address, IP address, VLAN ID, port) tuples that are specified as authorized. DHCP snooping can be enabled globally and on specific VLANs. Ports within the VLAN can be configured to be trusted or untrusted. DHCP servers must be reached through trusted ports. This feature is supported for both IPv4 and IPv6 packets.

Dynamic ARP Inspection

Dynamic ARP Inspection (DAI) is a security feature that rejects invalid and malicious ARP packets. The feature prevents a class of man-in-the-middle attacks, where an unfriendly station intercepts traffic for other stations by poisoning the ARP caches of its unsuspecting neighbors. The malicious station sends ARP requests or responses mapping another station's IP address to its own MAC address.

IP Source Address Guard

IP Source Guard and Dynamic ARP Inspection use the DHCP snooping bindings database. When IP Source Guard is enabled, the switch drops incoming packets that do not match a binding in the bindings database. IP Source Guard can be configured to enforce just the source IP address or both the source IP address and source MAC address. Dynamic ARP Inspection uses the bindings database to validate ARP packets. This feature is supported for both IPv4 and IPv6 packets.

Switching

This section describes the Layer 2 switching features DCSS software supports. For additional information and configuration examples for some of these features, see Chapter 5: "Configuring Switching," on page 69.

VLAN Support

VLANs are collections of switching ports that comprise a single broadcast domain. Packets are classified as belonging to a VLAN based on either the VLAN tag or a combination of the ingress port and packet contents. Packets sharing common attributes can be groups in the same VLAN. DCSS software is in full compliance with IEEE 802.1Q VLAN tagging.

Double VLANs

The Double VLAN feature (IEEE 802.1QinQ) allows the use of a second tag on network traffic. The additional tag helps differentiate between customers in the Metropolitan Area Networks (MAN) while preserving individual customer's VLAN identification when they enter their own 802.1Q domain.

Spanning Tree Protocol (STP)

Spanning Tree Protocol (IEEE 802.1D) is a standard requirement of Layer 2 switches that allows bridges to automatically prevent and resolve L2 forwarding loops. The STP feature supports a variety of per-port settings including path cost, priority settings, Port Fast mode, STP Root Guard, Loop Guard, TCN Guard, and Auto Edge. These settings are also configurable per-LAG.

Rapid Spanning Tree

Rapid Spanning Tree Protocol (RSTP) detects and uses network topologies to enable faster spanning tree convergence after a topology change, without creating forwarding loops. The port settings supported by STP are also supported by RSTP.

Multiple Spanning Tree

Multiple Spanning Tree (MSTP) operation maps VLANs to spanning tree instances. Packets assigned to various VLANs are transmitted along different paths within MSTP Regions (MST Regions). Regions are one or more interconnected MSTP bridges with identical MSTP settings. The MSTP standard lets administrators assign VLAN traffic to unique paths.

The switch supports IEEE 802.1Q-2005, which is a version that corrects problems associated with the previous version, provides for faster transition-to-forwarding, and incorporates new features for a port (restricted role and restricted TCN).

Bridge Protocol Data Unit (BPDU) Guard

Spanning Tree BPDU Guard is used to disable the port in case a new device tries to enter the already existing topology of STP. Thus devices, which were originally not a part of STP, are not allowed to influence the STP topology.

BPDU Filtering

When spanning tree is disabled on a port, the BPDU Filtering feature allows BPDU packets received on that port to be dropped. Additionally, the BPDU Filtering feature prevents a port in Port Fast mode from sending and receiving BPDUs. A port in Port Fast mode is automatically placed in the forwarding state when the link is up to increase convergence time.

Link Aggregation

Up to eight ports can combine to form a single Link Aggregated Group (LAG). This enables fault tolerance protection from physical link disruption, higher bandwidth connections and improved bandwidth granularity.

A LAG is composed of ports of the same speed, set to full-duplex operation.

Track LAG Member Port Flaps

This feature enables a user to track how many times a LAG member has flapped. The member flap counter show the number of times a port member is INACTIVE, either because the link is down, or the admin state is disabled. The Link Down counter shows the number of times the port-channel is down because all its member ports are INACTIVE.

Link Aggregate Control Protocol (LACP)

Link Aggregate Control Protocol (LACP) uses peer exchanges across links to determine, on an ongoing basis, the aggregation capability of various links, and continuously provides the maximum level of aggregation capability achievable between a given pair of systems. LACP automatically determines, configures, binds, and monitors the binding of ports to aggregators within the system.

Flow Control Support (IEEE 802.3x)

Flow control enables lower speed switches to communicate with higher speed switches by requesting that the higher speed switch refrains from sending packets. Transmissions are temporarily halted to prevent buffer overflows.

Asymmetric Flow Control

Asymmetric Flow Control can only be configured globally for all ports on the switch.

When in asymmetric flow control mode, the switch responds to PAUSE frames received from peers by stopping packet transmission, but the switch does not initiate MAC control PAUSE frames.

When the switch is configured in asymmetric flow control (or no flow control mode), the device is placed in egress drop mode. Egress drop mode maximizes the throughput of the system at the expense of packet loss in a heavily congested system, and this mode avoids head of line blocking.

Asymmetric flow control is NOT supported on Fast Ethernet platforms as the support was introduced to the physical layer with the Gigabit PHY specifications.

Note: In asymmetric flow control mode, the switch advertises the symmetric flow control capability, but forces the Tx Pause to OFF in the MAC layer. At PHY level, Pause bit = 1, and ASM_DIR =1 have to be advertised to the peer. At Driver level, Tx Pause = 0, and Rx Pause = 1, as described in IEEE 802.3-2005 Table 28B-2. The operational state (MAC layer) of receive Flow Control (Rx) is based on the pause resolution in IEEE 802.3-2005 Table 28B-3. The operational state (MAC layer) of Flow Control on Send side (Tx) is always Off.

Alternate Store and Forward (ASF)

The Alternate Store and Forward (ASF) feature, which is also known as cut-through mode, reduces latency for large packets. When ASF is enabled, the memory management unit (MMU) can forward a packet to the egress port before it has been entirely received on the Cell Buffer Pool (CBP) memory.

Jumbo Frames Support

Jumbo frames enable transporting data in fewer frames to ensure less overhead, lower processing time, and fewer interrupts. The maximum transmission unit (MTU) size is configurable per-port.

Auto-MDI/MDIX Support

Your switch supports auto-detection between crossed and straight-through cables. Media-Dependent Interface (MDI) is the standard wiring for end stations, and the standard wiring for hubs and switches is known as Media-Dependent Interface with Crossover (MDIX).

Unidirectional Link Detection (UDLD)

The UDLD feature detects unidirectionally linked physical ports by exchanging packets containing information about neighboring devices. The purpose of the UDLD feature is to detect and avoid unidirectional links. A unidirectional link is a forwarding anomaly in a Layer 2 communication channel in which a bidirectional link stops passing traffic in one direction.

Expandable Port Configuration

Expandable ports allow the administrator to configure a 40G port in either 4×10G mode or 1×40G mode. When the 40G port is operating in 4×10G mode, the port operates as four 10G ports, each on a separate lane. This mode requires the use of a suitable 4×10G to 1×40G pigtail cable. The mode of the expandable port takes place when the system boots, so if the mode is changed during switch operation, the change does not take effect until the next boot cycle.

VLAN-Aware MAC-based Switching

Packets arriving from an unknown source address are sent to the CPU and added to the Hardware Table. Future packets addressed to or from this address are more efficiently forwarded.

Back Pressure Support

On half-duplex links, a receiver may prevent buffer overflows by jamming the link so that it is unavailable for additional traffic. On full duplex links, a receiver may send a PAUSE frame indicating that the transmitter should cease transmission of frames for a specified period.

When flow control is enabled, the switch will observe received PAUSE frames or jamming signals, and will issue them when congested.

Auto Negotiation

Auto negotiation allows the switch to advertise modes of operation. The auto negotiation function provides the means to exchange information between two switches that share a point-to-point link segment, and to automatically configure both switches to take maximum advantage of their transmission capabilities.

The switch enhances auto negotiation by providing configuration of port advertisement. Port advertisement allows the system administrator to configure the port speeds that are advertised.

Storm Control

When Layer 2 frames are forwarded, broadcast, unknown unicast, and multicast frames are flooded to all ports on the relevant virtual local area network (VLAN). The flooding occupies bandwidth, and loads all nodes connected on all ports. Storm control limits the amount of broadcast, unknown unicast, and multicast frames accepted and forwarded by the switch.

Port Mirroring

Port mirroring monitors and mirrors network traffic by forwarding copies of incoming and outgoing packets from up to four source ports to a monitoring port. The switch also supports flow-based mirroring, which allows you to copy certain types of traffic to a single destination port. This provides flexibility—instead of mirroring all ingress or egress traffic on a port, the switch can mirror a subset of that traffic. You can configure the switch to mirror flows based on certain kinds of Layer 2, Layer 3, and Layer 4 information.

sFlow

sFlow is the standard for monitoring high-speed switched and routed networks. sFlow technology is built into network equipment and gives complete visibility into network activity, enabling effective management and control of network resources. The switch supports sFlow version 5.

Static and Dynamic MAC Address Tables

You can add static entries to the switch's MAC address table and configure the aging time for entries in the dynamic MAC address table. You can also search for entries in the dynamic table based on several different criteria.

Link Layer Discovery Protocol (LLDP)

The IEEE 802.1AB defined standard, Link Layer Discovery Protocol (LLDP), allows the switch to advertise major capabilities and physical descriptions. This information can help you identify system topology and detect bad configurations on the LAN.

Link Layer Discovery Protocol (LLDP) for Media Endpoint Devices

The Link Layer Discovery Protocol for Media Endpoint Devices (LLDP-MED) provides an extension to the LLDP standard for network configuration and policy, device location, Power over Ethernet management, and inventory management.

DHCP Layer 2 Relay

This feature permits Layer 3 Relay agent functionality in Layer 2 switched networks. The switch supports L2 DHCP relay configuration on individual ports, link aggregation groups (LAGs) and VLANs.

MAC Multicast Support

Multicast service is a limited broadcast service that allows one-to-many and many-to-many connections. In Layer 2 multicast services, a single frame addressed to a specific multicast address is received, and copies of the frame to be transmitted on each relevant port are created.

IGMP Snooping

Internet Group Management Protocol (IGMP) Snooping is a feature that allows a switch to forward multicast traffic intelligently on the switch. Multicast IP traffic is traffic that is destined to a host group. Host groups are identified by class D IP addresses, which range from 224.0.0.0 to 239.255.255.255. Based on the IGMP query and report messages, the switch forwards traffic only to the ports that request the multicast traffic. This prevents the switch from broadcasting the traffic to all ports and possibly affecting network performance.

Source Specific Multicasting (SSM)

This mechanism provides the ability for a host to report interest in receiving a particular multicast stream only from among a set of specific source addresses, or its interest in receiving a multicast stream from any source other than a set of specific source addresses.

Control Packet Flooding

This feature enhances the MGMD Snooping functionality to flood multicast packets with DIP=224.0.0.x to ALL members of the incoming VLAN irrespective of the configured filtering behavior. This enhancement depends on the ability of the underlying switching silicon to flood packets with DIP=224.0.0.x irrespective of the entries in the L2 Multicast Forwarding Tables. In platforms that do not have the said hardware capability, 2 ACLs (one for IPv4 and another for IPv6) would be consumed in the switching silicon to accomplish the flooding using software.

Flooding to mRouter Ports

This feature enhances the MGMD Snooping functionality to flood unregistered multicast streams to ALL mRouter ports in the VLAN irrespective of the configured filtering behavior. This enhancement depends on the ability of the underlying switching silicon to flood packets to specific ports in the incoming VLAN when there are no entries in the L2 Multicast Forwarding Tables for the specific stream. In platforms that do not have the this hardware capability, incoming multicast streams will always be flooded in the ingress VLAN when there is a L2MC-MISS in the switching silicon.

IGMP Snooping Querier

When Protocol Independent Multicast (PIM) and IGMP are enabled in a network with IP multicast routing, the IP multicast router acts as the IGMP querier. However, if it is desirable to keep the multicast network Layer 2 switched only, the IGMP Snooping Querier can perform the query functions of a Layer 3 multicast router.

Management and Control Plane ACLs

This feature provides hardware-based filtering of traffic to the CPU. An optional 'management' feature is available to apply the ACL on the CPU port. Currently, control packets like BPDU are dropped because of the implicit 'deny all' rule added at the end of the list. To overcome this rule, you must add rules that allow the control packets.

Support for user-defined simple rate limiting rule attributes for inbound as well as outbound traffic is also available. This attribute is supported on all QoS capable interfaces - physical, lag, and control-plane. Outbound direction is only supported on platforms with an Egress Field Processor (EFP).

Remote Switched Port Analyzer (RSPAN)

Along with the physical source ports, the network traffic received/transmitted on a VLAN can be monitored. A port mirroring session is operationally active if and only if both a destination (probe) port and at least one source port or VLAN is configured. If neither is true, the session is inactive. DCSS supports remote port mirroring. DCSS also supports VLAN mirroring. Traffic from/to all the physical ports which are members of that particular VLAN is mirrored.

Note: The source for a port mirroring session can be either physical ports or VLAN.

For Flow based mirroring, ACLs are attached to the mirroring session. The network traffic that matches the ACL is only sent to the destination port. This feature is supported for remote monitoring also. IP/MAC access-list can be attached to the mirroring session.

Note: Flow based mirroring is supported only if QoS feature exists in the package.

Data Center

This section describes the data center features DCSS software supports. For additional information and configuration examples for some of these features, see Section 6: “Configuring Data Center Features,” on page 105.

Priority-based Flow Control

The Priority-based Flow Control (PFC) feature allows the user to pause or inhibit transmission of individual priorities within a single physical link. By configuring PFC to pause a congested priority (priorities) independently, protocols that are highly loss sensitive can share the same link with traffic that has different loss tolerances. Priorities are differentiated by the priority field of the 802.1Q VLAN header.

An interface that is configured for PFC is automatically disabled for 802.3x flow control.

Data Center Bridging Exchange Protocol

The Data Center Bridging Exchange Protocol (DCBX) is used by data center bridge devices to exchange configuration information with directly-connected peers. The protocol is also used to detect misconfiguration of the peer DCBX devices and optionally, for configuration of peer DCBX devices.

Quantized Congestion Notification

Quantized Congestion Notification (QCN) supports congestion management of long-lived data flows within a network domain by enabling bridges to signal congestion information to end stations capable of transmission rate limiting to avoid frame loss. This mechanism enables support for higher-layer protocols that are highly loss or latency sensitive. QCN helps to allow network storage traffic, high performance computing traffic, and internet traffic to coexist within the same network.

QCN allows the flow of traffic to increase or decrease based on the behavior of the reaction point.

CoS Queuing and Enhanced Transmission Selection

The CoS Queuing feature allows the switch administrator to directly configure certain aspects of the device hardware queuing to provide the desired QoS behavior for different types of network traffic. The priority of a packet arriving at an interface can be used to steer the packet to the appropriate outbound CoS queue through a mapping table. CoS queue characteristics such as minimum guaranteed bandwidth, transmission rate shaping, etc. are user configurable at the queue (or port) level.

Enhanced Transmission Selection (ETS) allows Class of Service (CoS) configuration settings to be advertised to other devices in a data center network through DCBX ETS TLVs. CoS information is exchanged with peer DCBX devices using ETS TLVs.

OpenFlow

The OpenFlow feature enables the switch to be managed by a centralized OpenFlow Controller using the OpenFlow protocol. DCSS supports two variants of the OpenFlow protocol: Data Center Tenant Networking mode and OpenFlow 1.0 modeThe mode can be configured through the user interface.

Controllers

OpenFlow 1.0 mode enables the switch to inter-operate with standard OpenFlow controllers such as NOX, Beacon, and Big Switch.

Quality of Service

This section describes the Quality of Service (QoS) features DCSS software supports. For additional information and configuration examples for some of these features, see Chapter 7: “Configuring Quality of Service,” on page 123.

Access Control Lists (ACL)

Access Control Lists (ACLs) ensure that only authorized users have access to specific resources while blocking off any unwarranted attempts to reach network resources. ACLs are used to provide traffic flow control, restrict the contents of routing updates, decide which types of traffic are forwarded or blocked, and above all provide security for the network. The switch supports the following ALC types:

- IPv4 ACLs
- IPv6 ACLs
- MAC ACLs

For all ACL types, you can apply the ACL rule when the packet enters or exits the physical port, LAG, or VLAN interface.

Differentiated Services (DiffServ)

The QoS Differentiated Services (DiffServ) feature allows traffic to be classified into streams and given certain QoS treatment in accordance with defined per-hop behaviors. DCSS software supports both IPv4 and IPv6 packet classification.

Class Of Service (CoS)

The Class Of Service (CoS) queueing feature lets you directly configure certain aspects of switch queuing. This provides the desired QoS behavior for different types of network traffic when the complexities of DiffServ are not required. CoS queue characteristics, such as minimum guaranteed bandwidth and transmission rate shaping, are configurable at the queue (or port) level.

2 Getting Started

Accessing the Switch Command-Line Interface

The command-line interface (CLI) provides a text-based way to manage and monitor the switch features. You can access the CLI by using a direct connection to the console port or by using a Telnet or SSH client.

To access the switch by using Telnet or Secure Shell (SSH), the switch must have an IP address configured on either the service port or the network interface, and the management station you use to access the device must be able to ping the switch IP address. DHCP is enabled by default on the service port. It is disabled on the network interface.

Note: By default, entry into Privileged EXEC mode requires a password for Telnet and SSH access methods, and if the correct password is not supplied access is denied. Because no password is configured by default, access is always denied. For information about changing the default settings for Telnet and SSH access methods, see "Configuring and Applying Authentication Profiles" on page 59

Connecting to the Switch Console

To connect to the switch and configure or view network information, use the following steps:

1. Using a straight-through modem cable, connect a VT100/ANSI terminal or a workstation to the console (serial) port.

   If you attached a PC, Apple®, or UNIX® workstation, start a terminal-emulation program, such as HyperTerminal or TeraTerm.
2. Configure the terminal-emulation program to use the following settings:
   - Baud rate: 115200 bps
   - Data bits: 8
   - Parity: none
   - Stop bit: 1
   - Flow control: none
3. Power on the switch.

   For information about the boot process, including how to access the boot menu, see "Booting the Switch" on page 34.

   After the system completes the boot cycle, the `User:` prompt appears.
4. At the `User:` prompt, type `admin` and press ENTER. The `Password:` prompt appears.
5. There is no default password. Press ENTER at the password prompt if you did not change the default password.

   After a successful login, the screen shows the system prompt, for example `(DCSS Switching) >`.
6. At the `(DCSS Switching) >` prompt, enter `enable` to enter the Privileged EXEC command mode.
7. There is no default password to enter Privileged EXEC mode. Press ENTER at the password prompt if you did not change the default password.

The command prompt changes to `(DCSS Switching) #`.

8. To view service port network information, type `show serviceport` and press ENTER.

```
(DCSS Switching) #show serviceport

Interface Status.............................. Up
IP Address.................................... 10.27.21.33
Subnet Mask................................... 255.255.252.0
Default Gateway............................... 10.27.20.1
IPv6 Administrative Mode...................... Enabled
IPv6 Prefix is ............................... fe80::210:18ff:fe82:157c/64
Configured IPv4 Protocol...................... DHCP
Configured IPv6 Protocol...................... None
IPv6 AutoConfig Mode.......................... Disabled
Burned In MAC Address......................... 00:10:18:82:15:7C
```

By default, the DHCP client on the service port is enabled. If your network has a DHCP server, then you need only to connect the switch service port to your management network to allow the switch to acquire basic network information.

Accessing the Switch CLI Through the Network

Remote management of the switch is available through the service port or through the network interface. To use telnet, SSH, or SNMP for switch management, the switch must be connected to the network, and you must know the IP or IPv6 address of the management interface. The switch has no IP address by default. The DHCP client on the service port is enabled, and the DHCP client on the network interface is disabled.

After you configure or view network information, configure the authentication profile for telnet or SSH (see "Configuring and Applying Authentication Profiles" on page 59) and physically and logically connect the switch to the network, you can manage and monitor the switch remotely. You can also continue to manage the switch through the terminal interface via the console port.

Using the Service Port or Network Interface for Remote Management

The service port is a dedicated Ethernet port for out-of-band management. Using the service port to manage the switch is recommended. Traffic on this port is segregated from operational network traffic on the switch ports and cannot be switched or routed to the operational network. Additionally, if the production network is experiencing problems, the service port still allows you to access the switch management interface and troubleshoot issues. Configuration options on the service port are limited, which makes it difficult to accidentally cut off management access to the switch.

Alternatively, you can choose to manage the switch through the production network, which is known as in-band management. Because in-band management traffic is mixed in with production network traffic, it is subject to all of the filtering rules usually applied on a switched/routed port such as ACLs and VLAN tagging. You can access the in-band network management interface through a connection to any front-panel port.

Configuring Service Port Information

To disable DHCP/BootP and manually assign an IPv4 address, enter:

```
serviceport protocol none
serviceport ip ipaddress netmask [gateway]
```

For example, `serviceport ip 192.168.2.23 255.255.255.0 192.168.2.1`

To disable DHCP/BootP and manually assign an IPv6 address and (optionally) default gateway, enter:

```
serviceport protocol none
serviceport ipv6 address address/prefix-length [eui64]
serviceport ipv6 gateway gateway
```

To view the assigned or configured network address, enter:

```
show serviceport
```

To enable the DHCP client on the service port, enter:

```
serviceport protocol dhcp
```

To enable the BootP client on the service port, enter:

```
serviceport protocol bootp
```

Configuring the In-Band Network Interface

To use a DHCP server to obtain the IP address, subnet mask, and default gateway information, enter:

```
network protocol dhcp.
```

To use a BootP server to obtain the IP address, subnet mask, and default gateway information, enter:

```
network protocol bootp.
```

To manually configure the IPv4 address, subnet mask, and (optionally) default gateway, enter:

```
network parms ipaddress netmask [gateway],
```

For example, `network parms 192.168.2.23 255.255.255.0 192.168.2.1`

To manually configure the IPv6 address, subnet mask, and (optionally) default gateway, enter:

```
network ipv6 address address/prefix-length [eui64]
network ipv6 gateway gateway
```

To view the network information, enter:

```
show network.
```

To save these changes so they are retained during a switch reset, enter the following command:

```
copy system:running-config nvram:startup-config
```

DHCP Option 61

DHCP Option 61 (client Identifier) allows the DHCP server to be configured to provide an IP address to a switch based on its Media Access Control (MAC) Address or an ID entered into the system. DHCP servers use this value to index their database of address bindings. This value is expected to be unique for all clients in an administrative domain. This option allows the system to move from one part of the network to another while maintaining the same IP address.

DHCP client Identifier (Option 61) is used by DHCP clients to specify their unique identifier. The client identifier option is optional and can be specified while configuring the DHCP on the interfaces. DHCP Option 61 is enabled by default.

Configuring DHCP Option 61

Configuring the DHCP with client-id (option 61) differs depending on the port or interface. Refer to the information below:

Service Port:

To enable DHCP with client-id (option 61) on from the service port:

1. Issue the command `serviceport protocol dhcp client-id` at privilege mode.
   `(DCSS Switching) #serviceport protocol dhcp client-id`

Network Port:

To enable DHCP with client-id (option 61) on from the network port:

1. Issue the command `network protocol dhcp client-id` at privilege mode.
   `(DCSS Switching) #network protocol dhcp client-id`

Booting the Switch

When the power is turned on with the local terminal already connected, the switch goes through Power-On Self-Test (POST). POST runs every time the switch is initialized and checks hardware components to determine if the switch is fully operational before completely booting.

If a critical problem is detected, the program flow stops. If POST passes successfully, a valid executable image is loaded into RAM.

POST messages are displayed on the terminal and indicate test success or failure.

To view the text that prints to the screen during the boot process, perform the following steps:

1. Make sure that the serial cable is connected to the terminal.
2. Connect the power supply to the switch.
3. Power on the switch.

   As the switch boots, the bootup test first counts the switch memory availability and then continues to boot.
4. During boot, you can use the Utility menu, if necessary, to run special procedures. To enter the Boot menu, press **2** within the first five seconds after the following message appears.

   ```
   Select startup mode.  If no selection is made within 5 seconds,
   the DCSS Application will start automatically...

   DCSS Startup -- Main Menu

   1 - Start DCSS Application
   2 - Display Utility Menu
   Select (1, 2):
   ```

   For information about the Boot menu, see “Utility Menu Functions” on page 35.
5. If you do not start the boot menu, the operational code continues to load.

After the switch boots successfully, the User login prompt appears and you can use the local terminal to begin configuring the switch. However, before configuring the switch, make sure that the software version installed on the switch is the latest version.

Utility Menu Functions

Note: Utility menu functions vary on different platforms. The following example might not represent the options available on your platform.

You can perform many configuration tasks through the Utility menu, which can be invoked after the first part of the POST is completed.

To display the Utility menu, boot the switch observe the output that prints to the screen. After various system initialization information displays, the following message appears:

```
DCSS Startup Rev: 7.0

Select startup mode.  If no selection is made within 5 seconds,
the DCSS Application will start automatically...

DCSS Startup -- Main Menu

1 - Start DCSS Application
2 - Display Utility Menu
Select (1, 2):
```

Press press **2** within five seconds to start the Utility menu. If you do not press 2, the system loads the operational code.

After you press 2 the following information appears:

```
DCSS Startup -- Utility Menu

 1  - Start DCSS Application
 2  - Load Code Update Package
 3  - Load Configuration
 4  - Select Serial Speed
 5  - Retrieve Error Log
 6  - Erase Current Configuration
 7  - Erase Permanent Storage
 8  - Select Boot Method
 9  - Activate Backup Image
10  - Start Diagnostic Application
11  - Reboot
12  - Erase All Configuration Files

 Q  - Quit from DCSS Startup

Select option (1-12 or Q):
```

The following sections describe the Utility menu options.

Start DCSS Application

Use option 1 to resume loading the operational code. After you enter 1, the switch exits the Startup Utility menu and the switch continues the boot process.

Load Code Update Package

Use option 2 to download a new software image to the switch to replace a corrupted image or to update, or upgrade the system software.

The switch is preloaded with DCSS software, so these procedures are needed only for upgrading or downgrading to a different image.

You can use any of the following methods to download the image:

- TFTP
- XMODEM
- YMODEM
- ZMODEM

If you use TFTP to download the code, the switch must be connected to the network, and the code to download must be located on the TFTP server.

When you use XMODEM, YMODEM, or ZMODEM to download the code, the code must be located on an administrative system that has a console connection to the switch.

Use the following procedures to download an image to the switch by using TFTP:

1. From the Utility menu, select **2** and press ENTER.

   The switch creates a temporary directory and prompts you to select the download method:

   ```
   Creating tmpfs filesystem on tmpfs for download...done.
   Select Mode of Transfer (Press T/X/Y/Z for TFTP/XMODEM/YMODEM/ZMODEM) []:
   ```

2. Enter **T** to download the image from a TFTP server to the switch.
3. Enter the IP address of the TFTP server where the new image is located, for example:

   ```
   Enter Server IP []:192.168.1.115
   ```

4. Enter the desired IP address of the switch management interface, for example:

   ```
   Enter Host IP []192.168.1.23
   ```

Note: The switch uses the IP address, subnet mask, and default gateway information you specify for the TFTP download process only. The switch automatically reboots after the process completes, and this information is not saved.

5. Enter the subnet mask associated with the management interface IP address or press ENTER to accept the default value, which is 255.255.255.0.
6. Optionally, enter the IP address of the default gateway for the switch management interface, for example:
   `Enter Gateway IP []192.168.1.1`
7. Enter the filename, including the file path (if it is not in the TFTP root directory), of the image to download, for example:

   ```
   Enter Filename[]images/image0630.stk
   ```

8. Confirm the information you entered and enter **y** to allow the switch to contact the TFTP server.

   After the download completes, you are prompted to reboot the switch. The switch loads the image during the next boot cycle.

Use the following procedures to download an image to the switch by using XMODEM, YMODEM, or ZMODEM.

1. From the Utility menu, select **2** and press ENTER.

   The switch creates a temporary directory and prompts you to select the download method:

```
Creating tmpfs filesystem on tmpfs for download...done.
Select Mode of Transfer (Press T/X/Y/Z for TFTP/XMODEM/YMODEM/ZMODEM) []:
```

2. Specify the protocol to use for the download.
   - Enter **X** to download the image by using the XMODEM file transfer protocol.
   - Enter **Y** to download the image by using the YMODEM file transfer protocol.
   - Enter **Z** to download the image by using the ZMODEM file transfer protocol.
3. When you are ready to transfer the file from the administrative system, enter **y** to continue.

   ```
   Do you want to continue? Press(Y/N): y
   ```
4. From the terminal or terminal emulation application on the administrative system, initiate the file transfer.

   For example, if you use HyperTerminal, use the following procedures:
   a. From the **HyperTerminal** menu bar, click **Transfer** > **Send File**.

      The **Send File** window displays.
   b. Browse to the file to download and click **Open** to select it.
   c. From the **Protocol:** field, select the protocol to use for the file transfer.
   d. Click **Send**.

   After you start the file transfer, the software is downloaded to the switch, which can take several minutes. The terminal emulation application might display the loading process progress.
5. After software downloads, you are prompted to reboot the switch. The switch loads the image during the next boot cycle.

Load Configuration

Use option 3 to download a new configuration that will replace the saved system configuration file. You can use any of the following methods to download the configuration file:

- TFTP
- XMODEM
- YMODEM
- ZMODEM

Use the following procedures to download a configuration file to the switch.

1. From the Utility menu, select **3** and press ENTER.
2. Enter T to download the text-based configuration file to the switch.
3. Specify the protocol to use for the download.
4. Respond to the prompts to begin the file transfer.

   The configuration file download procedures are very similar to the software image download procedures. For more information about the prompts and how to respond, see "Load Code Update Package" on page 35.

Select Serial Speed

Use option 4 to change the baud rate of the serial interface (console port) on the switch. When you select option **4**, the following information displays:

```
1 - 2400
2 - 4800
```

```
3 - 9600
4 - 19200
5 - 38400
6 - 57600
7 - 115200
8 - Exit without change
Select option (1-8):
```

To set the serial speed, enter the number that corresponds to the desired speed.

Note: The selected baud rate takes effect immediately.

Retrieve Error Log

Use option 5 to retrieve the error log that is stored in nonvolatile memory and upload it from the switch to your ASCII terminal or administrative system. You can use any of the following methods to copy the error log to the system:

- TFTP
- XMODEM
- YMODEM
- ZMODEM

Use the following procedures to upload the error log from the switch:

1. From the Utility menu, select **5** and press ENTER.
2. Specify the protocol to use for the download.
3. Respond to the prompts to begin the file transfer.

   If you use TFTP to upload the file from the switch to the TFTP server, the prompts and procedures are very similar to the steps described for the TFTP software image download. For more information about the prompts and how to respond, see "Load Code Update Package" on page 35.

   If you use XMODEM, YMODEM, or ZMODEM to transfer the file, configure the terminal or terminal emulation application with the appropriate settings to receive the file. For example, if you use HyperTerminal, click **Transfer** > **Receive File**, and then specify where to put the file and which protocol to use.

Erase Current Configuration

Use option 6 to clear changes to the startup-config file and reset the system to its factory default setting. This option is the same as executing the `clear config` command from Privileged EXEC mode. You are not prompted to confirm the selection.

Erase Permanent Storage

Use option 7 to completely erase the switch software application, any log files, and any configurations. The boot loader and operating system are not erased. Use this option only if a file has become corrupt and your are unable to use option 2, Load Code Update Package, to load a new image onto the switch. After you erase permanent storage, you must download an image to the switch; otherwise, the switch will not be functional.

Select Boot Method

Use option 8 to specify whether the system should boot from the image stored on the internal flash, from an image over the network, or from an image over the serial port. By default, the switch boots from the flash image.

To boot over the network, the image must be located on a TFTP server that can be accessed by the switch. To boot from the serial port, the switch must be connected through the console port to a terminal or system with a terminal emulator. The image must be located on the connected device.

If you select option 8, the following menu appears:

```
Current boot method: FLASH
1 - Flash Boot
2 - Network Boot
3 - Serial Boot
4 - Exit without change
Select option (1-4):
```

If you select a new boot method, the switch uses the selected method for the next boot cycle.

Activate Backup Image

Use option 9 to activate the backup image. The active image becomes the backup when you select this option. When you exit the Startup Utility and resume the boot process, the switch loads the image that you activated, but reloading the switch is recommended so it can perform an entire boot cycle with the newly active image.

After you active the backup image, the following information appears.

```
Image image1 is now active.
Code update instructions found!
Extracting kernel and rootfs from image1
Copying kernel/rootfs uimage to boot flash area
Activation complete
image1 activated -- system reboot recommended!
Reboot? (Y/N):
```

Enter **y** to reload the switch.

Start Diagnostic Application

Option 10 is for field support personnel only. Access to the diagnostic application is password protected.

Reboot

Use option 11 to restart the boot process.

Erase All Configuration Files

Use option 12 to clear changes to the startup-config file and the factory-defaults file and reset the system to its factory default (compile-time) setting. You are not prompted to confirm the selection.

Understanding the User Interfaces

DCSS software includes a set of comprehensive management functions for configuring and monitoring the system by using one of the following two methods:

- Command-Line Interface (CLI)
- Simple Network Management Protocol (SNMP)

These standards-based management methods allows you to configure and monitor the components of the DCSS software. The method you use to manage the system depends on your network size and requirements, and on your preference.

Note: Not all features are supported on all hardware platforms, so some CLI commands and object identifiers (OIDs) might not be available on your platform.

Using the Command-Line Interface

The command-line interface (CLI) is a text-based way to manage and monitor the system. You can access the CLI by using a direct serial connection or by using a remote logical connection with telnet or SSH.

The CLI groups commands into modes according to the command function. Each of the command modes supports specific software commands. The commands in one mode are not available until you switch to that particular mode, with the exception of the User EXEC mode commands. You can also execute the User EXEC mode commands in the Privileged EXEC mode.

To display the commands available in the current mode, enter a question mark (?) at the command prompt. To display the available command keywords or parameters, enter a question mark (?) after each word you type at the command prompt. If there are no additional command keywords or parameters, or if additional parameters are optional, the following message appears in the output:

```
<cr>                    Press Enter to execute the command
```

For more information about the CLI, see the *CLI Command Reference.*

The *CLI Command Reference* lists each command available from the CLI by the command name and provides a brief description of the command. Each command reference also contains the following information:

- The command keywords and the required and optional parameters.
- The command mode you must be in to access the command.
- The default value, if any, of a configurable setting on the device.

The `show` commands in the document also include a description of the information that the command shows.

Using SNMP

SNMP is enabled by default. The `show sysinfo` command displays the information you need to configure an SNMP manager to access the switch. You can configure SNMP groups and users that can manage traps that the SNMP agent generates.

DCSS uses both standard public MIBs for standard functionality and private MIBs that support additional switch functionality. All private MIBs begin with a "AS4600-54T-" prefix. The main object for interface configuration is in AS4600-54T-SWITCHING-MIB, which is a private MIB. Some interface configurations also involve objects in the public MIB, IF-MIB.

SNMPv3

SNMP version 3 (SNMPv3) adds security and remote configuration enhancements to SNMP. DCSS has the ability to configure SNMP server, users, and traps for SNMPv3. Any user can connect to the switch using the SNMPv3 protocol, but for authentication and encryption, you need to configure a new user profile. To configure a profile by using the CLI, see the SNMP section in the *CLI Command Reference*.

3 Configuring Switch Management

Managing Images and Files

DCSS-based switches maintain several different types of files on the flash file system. Table 1 describes the files that you can manage. You use the copy command to copy a source file to a destination file. The copy command may permit the following actions (depending on the file type):

- Copy a file from the switch to a remote server
- Copy a file from a remote server to the switch
- Overwrite the contents of the destination file with the contents of the source file.

Table 1: Files to Manage

File	*Description*
active	The switch software image that has been loaded and is currently running on the switch.
backup	A second software image that is currently not running on the switch.
startup-config	Contains the software configuration that loads during the boot process.
running-config	Contains the current switch configuration.
factory-defaults	Contains the software configuration that can be used to load during the boot process or after “clear config”.
backup-config	An additional configuration file that serves as a backup. You can copy the startup-config file to the backup-config file.
dcss.cfg	A binary configuration file.
Configuration script	Text file with CLI commands. When you apply a script on the switch, the commands are executed and added to the running-config.
CLI Banner	Text file containing the message that displays upon connecting to the switch or logging on to the switch by using the CLI.
Log files	Trap, error, or other log files that provide various information about events that occur on the switch.
SSH key files	Contains information to authenticate SSH sessions. The switch supports the following files for SSH: • SSH-1 RSA Key File • SSH-2 RSA Key File (PEM Encoded) • SSH-2 Digital Signature Algorithm (DSA) Key File (PEM Encoded) ***Note:*** If you use the CLI to manage the switch over an SSH connection, you must copy the appropriate key files to the switch.
IAS Users	List of Internal Authentication Server (IAS) users for IEEE 802.1X authentication. You can configure the switch to use the local IAS user database for port-based authentication instead of using a remote server, such as a RADIUS server.

Supported File Management Methods

For most file types, you can use any of the following protocols to download files from a remote system to the switch or to upload files from the switch to a remote system:

- FTP
- TFTP
- SFTP
- SCP
- XMODEM
- YMODEM
- ZMODEM

Note: The IAS Users file can be copied from a remote server to the switch only by using FTP, TFTP, SFTP, or SCP.

Uploading and Downloading Files

To use FTP, TFTP, SFTP, or SCP for file management, you must provide the IP address of the remote system that is running the appropriate server (FTP, TFTP, SFTP, or SCP). Make sure there is a route from the switch to the remote system. You can use the `ping` command from the CLI to verify that a route exists between the switch and the remote system.

If you are copying a file from the remote system to the switch, be sure to provide the correct path to the file (if the file is **not** in the root directory) and the correct file name.

Managing Switch Software (Images)

The switch can maintain two software images: the active image and the backup image. When you copy a new image from a remote system to the switch, you can specify whether to save it as the active or backup image. The downloaded image overwrites the image that you specify. If you save the new image as the active image, the switch continues to operate using the current (old) image until you reload the switch. Once the switch reboots, it loads with the new image. If you download the new image as the backup image, the file overwrites the current backup image, if it exists. To load the switch with the backup image, you must first set it as the active image and then reload the switch. The image that was previously the active image becomes the backup image after the switch reloads.

If you activate a new image and reload the switch, and the switch is unable to complete the boot process due to a corrupt image or other problem, you can use the boot menu to activate the backup image. You must be connected to the switch through the console port to access the boot menu.

To create a backup copy of the firmware on the switch, copy the active image to the backup image. You can also copy an image to a file on a remote server.

Managing Configuration Files

Configuration files contain the CLI commands that change the switch from its default configuration. The switch can maintain three separate configuration files: startup-config, running-config, and backup-config. The switch loads the startup-config file when the switch boots. Any configuration changes that take place after the boot process completes are written to the running-config file. The backup-config file does not exist until you explicitly create one by copying an existing configuration file to the backup-config file or downloading a backup-config file to the switch.

You can also create configuration scripts, which are text files that contains CLI commands.

When you apply (run) a configuration script on the switch, the commands in the script are executed in the order in which they are written as if you were typing them into the CLI. The commands that are executed in the configuration script are added to the running-config file.

You might upload a configuration file from the switch to a remote server for the following reasons:

- To create a backup copy
- To use the configuration file on another switch
- To manually edit the file

You might download a configuration file from a remote server to the switch for the following reasons:

- To restore a previous configuration
- To load the configuration copied from another switch
- To load the same configuration file on multiple switches

Use a text editor to open a configuration file and view or change its contents.

Editing and Downloading Configuration Files

Each configuration file contains a list of executable CLI commands. The commands must be complete and in a logical order, as if you were entering them by using the switch CLI.

When you download a startup-config or backup-config file to the switch, the new file replaces the previous version. To change the running-config file, you execute CLI commands either by typing them into the CLI or by applying a configuration script with the `script apply` command.

Creating and Applying Configuration Scripts

When you use configuration scripting, keep the following considerations and rules in mind:

Note: If your switch is currently at DCSS software version 2.1 and you plan to downgrade the switch to a version previous to DCSS 2.1, you must uncompress the scripts so they will operate. See "Uncompressing Configuration Scripts" on page 44

- The application of scripts is partial if the script fails. For example, if the script executes four of ten commands and the script fails, the script stops at four, and the final six commands are not executed.
- Scripts cannot be modified or deleted while being applied.
- Validation of scripts checks for syntax errors only. It does not validate that the script will run.
- The file extension must be `.scr`.
- A maximum of 10 scripts are allowed on the switch.
- There is no limit on the maximum number of scripts files that can be stored on the switch within a given storage space limit.
- The combined size of all script files on the switch cannot exceed 2048 Kbytes. Zlib compression technique is applied to script files to decrease script file size.

You can type single-line annotations in the configuration file to improve script readability. The exclamation point (!) character flags the beginning of a comment. Any line in the file that begins with the "!" character is recognized as a comment line and ignored by the parser. Do not use a comment character anywhere in a line that contains a command.

The following example shows annotations within a file (commands are bold):

```
!Configuration script for mapping lab hosts to IP addresses
!Enter Global Config mode and map host name to address
configure
    ip host labpc1 192.168.3.56
    ip host labpc2 192.168.3.57
    ip host labpc3 192.168.3.58
    exit
! End of the script file
```

Uncompressing Configuration Scripts

If you plan to downgrade your switch from DCSS 2.1, you must use the following procedure to uncompress the scripts.

1. Upload the scripts from the switch to an external server via FTP/TFTP. During the upload process from the switch, the scripts are uncompressed.
2. Downgrade the software image on the switch.
3. Download the uncompressed script files to the switch.

Saving the Running Configuration

Changes you make to the switch configuration while the switch is operating are written to the running-config. These changes are not automatically written to the startup-config. When you reload the switch, the startup-config file is loaded. If you reload the switch (or if the switch resets unexpectedly), any settings in the running-config that were not explicitly saved to the startup-config are lost. You must save the running-config to the startup-config to ensure that the settings you configure on the switch are saved across a switch reset.

To save the running-config to the startup-config from the CLI, use the `write memory` command.

File and Image Management Configuration Examples

This section contains the following examples:

- Upgrading the Firmware
- Managing Configuration Scripts

Upgrading the Firmware

This example shows how to download a firmware image to the switch and activate it. The TFTP server in this example is PumpKIN, an open source TFTP server running on a Windows system.

- TFTP server IP address: 10.27.65.112
- File path: \image
- File name: dcss_2100.stk

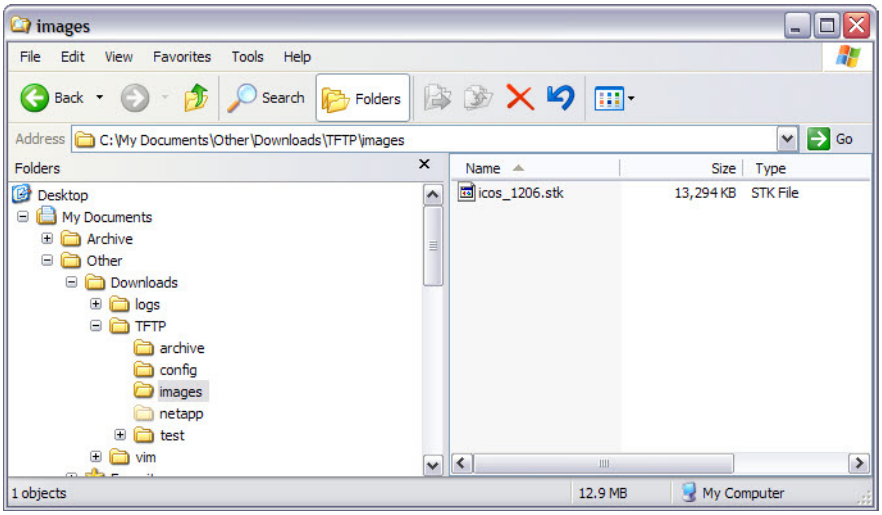
Use the following steps to prepare the download, and then download and upgrade the switch image.

1. Check the connectivity between the switch and the TFTP server.

```
(DCSS Switching) #ping 10.27.65.112
 Pinging 10.27.65.112 with 0 bytes of data:

Reply From 10.27.65.112: icmp_seq = 0. time= 5095 usec.
```

```
----10.27.65.112 PING statistics----
1 packets transmitted, 1 packets received, 0% packet loss
round-trip (msec) min/avg/max = 5/5/5
```

2. Copy the image file to the appropriate directory on the TFTP server. In this example, the TFTP root directory is `C:\My Documents\Other\Downloads\TFTP`, so the file path is `images`.



3. View information about the current image.

```
(DCSS Switching) #show bootvar
Image Descriptions
active : default image
 backup :

 Images currently available on Flash
----------------------------------------------------------------------
 unit      active      backup     current-active        next-active
----------------------------------------------------------------------
    1    I.12.5.1 11.21.16.52          I.12.5.1           I.12.5.1
```

4. Download the image to the switch. After you execute the `copy` command, you must verify that you want to start the download. The image is downloaded as the backup image.

```
(DCSS Switching) #copy tftp://10.27.65.112/images/dcss_2100.stk backup

Mode........................................... TFTP
Set Server IP.................................. 10.27.65.112
Path........................................... images/
Filename....................................... dcss_2100.stk
Data Type...................................... Code
Destination Filename........................... backup

Management access will be blocked for the duration of the transfer
Are you sure you want to start? (y/n)y
```

5. After the transfer completes, activate the new image so that it becomes the active image after the switch resets.

```
(DCSS Switching) #boot system backup
Activating image backup ..
```

6. View information about the current image.

```
(DCSS Switching) #show bootvar
Image Descriptions

 active : default image
 backup :

 Images currently available on Flash
----------------------------------------------------------------------
 unit      active     backup         current-active    next-active
----------------------------------------------------------------------
    1     I.12.5.1    11.21.16.52          I.12.5.1       I.12.6.2
```

7. Copy the running configuration to the startup configuration to save the current configuration to NVRAM.

```
(DCSS Switching) #write memory

This operation may take a few minutes.
Management interfaces will not be available during this time.

Are you sure you want to save? (y/n)y

Configuration Saved!
```

8. Reset the switch to boot the system with the new image.

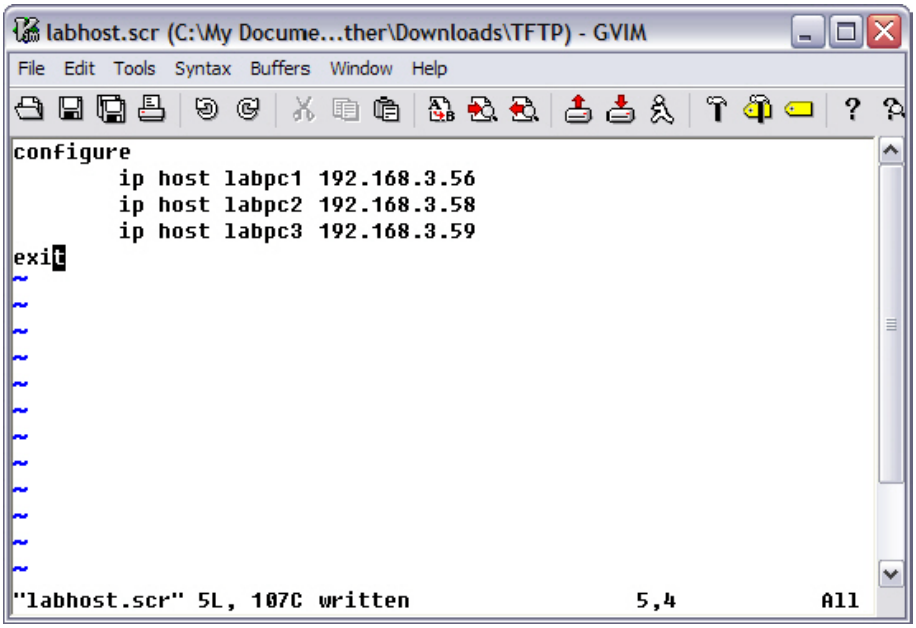
```
(DCSS Switching) #reload
Are you sure you want to continue? (y/n)y
Reloading all switches...
```

Managing Configuration Scripts

This example shows how to create a configuration script that adds three hostname-to-IP address mappings to the host table.

To configure the switch:

1. Open a text editor on an administrative computer and type the commands as if you were entering them by using the CLI.



2. Save the file with an *.scr extension and copy it to the appropriate directory on your TFTP server.
3. Download the file from the TFTP server to the switch.

```
(DCSS Switching) #copy tftp://10.27.65.112/labhost.scr nvram:script labhost.scr

Mode........................................... TFTP
Set Server IP.................................. 10.27.65.112
Path........................................... ./
Filename....................................... labhost.scr
Data Type...................................... Config Script
Destination Filename........................... labhost.scr

Management access will be blocked for the duration of the transfer
Are you sure you want to start? (y/n)
```

4. After you confirm the download information and the script successfully downloads, it is automatically validated for correct syntax.

```
Are you sure you want to start? (y/n) y

135 bytes transferred

Validating configuration script...
configure
exit
configure
ip host labpc1 192.168.3.56

ip host labpc2 192.168.3.58

ip host labpc3 192.168.3.59

Configuration script validated.
File transfer operation completed successfully.
```

5. Run the script to execute the commands.

```
(DCSS Switching) #script apply labhost.scr

Are you sure you want to apply the configuration script? (y/n)y

configure
exit
configure
ip host labpc1 192.168.3.56

ip host labpc2 192.168.3.58sj

ip host labpc3 192.168.3.59

Configuration script 'labhost.scr' applied.
```

6. Verify that the script was successfully applied.

```
(DCSS Switching) #show hosts
.
.
.
```

```
Configured host name-to-address mapping:
Host                      Addresses
------------------------  ------------------------
labpc1                    192.168.3.56
labpc2                    192.168.3.58
labpc3                    192.168.3.59
```

Enabling Automatic Image Installation and System Configuration

The Auto Install feature can automatically update the firmware image and obtain configuration information when the switch boots. Auto Install begins the automatic download and installation process when the switch boots and loads a saved configuration that has the persistent Auto Install mode enabled. Additionally, the switch supports a non-persistent Auto Install mode so that Auto Install can be stopped or restarted at any time during switch operation.

DHCP Auto Install Process

The switch can use a DHCP server to obtain configuration information from a TFTP server.

DHCP Auto Install is accomplished in three phases:

1. Assignment or configuration of an IP address for the switch
2. Assignment of a TFTP server
3. Obtaining a configuration file for the switch from the TFTP server

Auto Install is successful when an image or configuration file is downloaded to the switch from a TFTP server.

Note: The downloaded configuration file is not automatically saved to startup-config. You must explicitly issue a save request (`write memory`) in order to save the configuration.

Obtaining IP Address Information

DHCP is enabled by default on the service port. If an IP address has not been assigned, the switch issues requests for an IP address assignment.

A network DHCP server returns the following information:

- IP address and subnet mask to be assigned to the interface
- IP address of a default gateway, if needed for IP communication

Obtaining Other Dynamic Information

The following information is also processed and may be returned by a BOOTP or DHCP server:

- Name of configuration file (the *file* field in the DHCP header or option 67) to be downloaded from the TFTP server.
- Identification of the TFTP server providing the file. The TFTP server can be identified by name or by IP address as follows:
  - hostname: DHCP option 66 or the *sname* field in the DHCP header
  - IP address: DHCP option 150 or the *siaddr* field in the DHCP header

When a DHCP OFFER identifies the TFTP server more than once, the DHCP client selects one of the options in the following order: *sname*, option 66, option 150, *siaddr*. If the TFTP server is identified by hostname, a DNS server is required to translate the name to an IP address.

The DHCP client on the switch also processes the name of the text file (option 125, the V-I vendor-specific Information option) which contains the path to the image file.

Obtaining the Image

Auto Install attempts to download an image file from a TFTP server only if the switch loads with a saved configuration file that has Auto Install enabled (the `boot host dhcp` command) or if Auto Install has been administratively activated by issuing the `boot autoinstall start` command during switch operation.

The network DHCP server returns a DHCP OFFER message with option 125. When configuring the network DHCP server for image downloads, you must include Option 125 and specify the Broadcom Enterprise Number, 4413. Within the Broadcom section of option 125, sub option 5 must specify the path and name of a file on the TFTP server. This file is not the image file itself, but rather a text file that contains the path and name of the image file. Upon receipt of option 125, the switch downloads the text file from the TFTP server, reads the name of the image file, and downloads the image file from the TFTP server.

After the switch successfully downloads and installs the new image, it automatically reboots. The download or installation might fail for one of the following reasons:

- The path or filename of the image on the TFTP server does not match the information specified in DHCP option 125.
- The downloaded image is the same as the current image.
- The validation checks, such as valid CRC Checksum, fails.

If the download or installation was unsuccessful, a message is logged.

Obtaining the Configuration File

If the DHCP OFFER identifies a configuration file, either as option 67 or in the *file* field of the DHCP header, the switch attempts to download the configuration file.

Note: The configuration file is required to have a file type of *.cfg.

The TFTP client makes three unicast requests. If the unicast attempts fail, or if the DHCP OFFER did not specify a TFTP server address, the TFTP client makes three broadcast requests.

If the DHCP server does not specify a configuration file or download of the configuration file fails, the Auto Install process attempts to download a configuration file with the name fp-net.cfg. The switch unicasts or broadcasts TFTP requests for a network configuration file in the same manner as it attempts to download a host-specific configuration file.

The default network configuration file consists of a set of IP address-to-hostname mappings, using the command `ip host hostname` *`address`*. The switch finds its own IP address, as learned from the DHCP server, in the configuration file and extracts its hostname from the matching command. If the default network configuration file does not contain the switch's IP address, the switch attempts a reverse DNS lookup to resolve its hostname.

A sample `fp-net.cfg` file follows:

```
config
    ...
    ip host switch1 192.168.1.10
    ip host switch2 192.168.1.11
    ... <other hostname definitions>
exit
```

Once a hostname has been determined, the switch issues a TFTP request for a file named *hostname*.cfg, where *hostname* is the first thirty-two characters of the switch's hostname.

If the switch is unable to map its IP address to a hostname, Auto Install sends TFTP requests for the default configuration file host.cfg.

Table 2 summarizes the config files that may be downloaded and the order in which they are sought.

Table 2: Configuration File Possibilities

Order Sought	*File Name*	*Description*	*Final File Sought*
1	bootfile.cfg	Host-specific config file, ending in a *.cfg file extension	Yes
2	fp-net.cfg	Default network config file	No
3	hostname.cfg	Host-specific config file, associated with hostname.	Yes
4	host.cfg	Default config file	Yes

Table 3 displays the determining factors for issuing unicast or broadcast TFTP requests.

Table 3: TFTP Request Types

TFTP Server Address Available	*Host-specific Switch Config Filename Available*	*TFTP Request Method*
Yes	Yes	Issue a unicast request for the host-specific router config file to the TFTP server
Yes	No	Issue a unicast request for a default network or router config file to the TFTP server
No	Yes	Issue a broadcast request for the host-specific router config file to any available TFTP server
No	No	Issue a broadcast request for the default network or router config file to any available TFTP server

Monitoring and Completing the DHCP Auto Install Process

When the switch boots and triggers an Auto Install, a message is written to the buffered log. After the process completes, the Auto Install process writes a log message. You can use the `show logging buffered` command to view information about the process. The following log message indicates that the switch has broadcast a request to download the fp-net.cfg file from any TFTP server on the network.

```
<14> Jan  1 00:00:42 10.27.22.157-1 AUTO_INST[310234388]: auto_install_control.c(2427) 202 %%
AutoInstall<->TFTP : Downloading tftp://255.255.255.255/fp-net.cfg (via eth0)
```

Additionally, while the Auto Install is running, you can issue the `show autoinstall` command to view information about the current Auto Install state.

When Auto Install has successfully completed, you can execute a `show running-config` command to validate the contents of configuration.

Saving a Configuration

The Auto Install feature includes an AutoSave feature that allows the downloaded configuration to be automatically saved; however, AutoSave is disabled by default. If AutoSave has not been enabled, you must explicitly save the downloaded configuration in non-volatile memory. This makes the configuration available for the next reboot. In the CLI, this is performed by issuing a `write memory` command or `copy system:running-config nvram:startup-config` command and should be done after validating the contents of saved configuration.

Stopping and Restarting the Auto Install Process

You can terminate the Auto Install process at any time before the image or configuration file is downloaded. This is useful when the switch is disconnected from the network. Termination of the Auto Install process ends further periodic requests for a host-specific file.

Managing Downloaded Config Files

The configuration files downloaded to the switch by Auto Install are stored in the nonvolatile memory as `.scr` files. The files may be managed (viewed or deleted) along with files downloaded by the configuration scripting utility. If the Auto Install persistent mode is enabled (`boot dhcp host`) and the switch reboots, the `.scr` configuration file created by the switch in the non-volatile memory is overwritten during the Auto Install process.

To ensure that the downloaded configuration file is used during the next boot cycle, make sure that the Auto Install persistent mode is disabled (`no boot dhcp host`) and save the configuration (`write memory`).

DHCP Auto Install Dependencies

The Auto Install process from TFTP servers depends upon the following network services:

- A DHCP server must be configured on the network with appropriate services.
- An image file and a text file containing the image file name for the switch must be available from a TFTP server if DHCP image download is desired.
- A configuration file (either from bootfile (or) option 67 option) for the switch must be available from a TFTP server.
- The switch must be connected to the network and have a Layer 3 interface that is in an UP state.
- A DNS server must contain an IP address to hostname mapping for the TFTP server if the DHCP server response identifies the TFTP server by name.
- A DNS server must contain an IP address to hostname mapping for the switch if a `<hostname>.cfg` file is to be downloaded.
- If a default gateway is needed to forward TFTP requests, an IP helper address for TFTP needs to be configured on the default gateway.

Default Auto Install Values

Table 4 describes the Auto Install defaults.

Table 4: Auto Install Defaults

Feature	*Default*	*Description*
Retry Count	3	When the DHCP or BootP server returns information about the TFTP server and bootfile, the switch makes three unicast TFTP requests for the specified bootfile. If the unicast attempts fail or if a TFTP server address was not provided, the switch makes three broadcast requests to any available TFTP server for the specified bootfile.
AutoSave	Disabled	If the switch is successfully auto-configured, the running configuration is not saved to the startup configuration.
AutoReboot	Enabled	After an image is successfully downloaded during the Auto Install process, the switch automatically reboots and makes the downloaded image the active image.

Enabling DHCP Auto Install and Auto Image Download

A network administrator is deploying three switches and wants to quickly and automatically install the latest image and a common configuration file that configures basic settings such as VLAN creation and membership and RADIUS server settings. This example describes the procedures to complete the configuration. The DHCP and TFTP servers in this example are reachable from the service port on the switch.

To use DHCP Auto Install:

1. Log on to each switch and enable persistent Auto Install mode.

   ```
   (DCSS Switching) #boot host dhcp
   ```

2. Save the running configuration to the startup configuration file.

   ```
   (DCSS Switching) #write memory
   ```

3. Create a default config file for the switches named `host.cfg`. For information about creating configuration files, see “Managing Images and Files” on page 41.
4. Upload the `host.cfg` file to the TFTP server.
5. Upload the image file to the TFTP server.
6. Configure an address pool on the DHCP server that contains the following information:
   a. The IP address (*yiaddr*) and subnet mask (option 1) to be assigned to the interface
   b. The IP address of a default gateway (option 3)
   c. DNS server address (option 6)
   d. Name of config file for each host
   e. Identification of the TFTP server by hostname (DHCP option 66 or the *sname* field in the DHCP header) or IP address (DHCP option 150 or the *siaddr* field in the DHCP header)
   f. Name of the text file (option 125, the V-I vendor-specific Information option) that contains the path to the image file.
7. Connect the service port on each switch to the management network. This network must have a route to the DHCP server and TFTP server that are used for Auto Install process.

8. Reboot each switch.

```
(DCSS Switching) #reload
```

Downloading a Core Dump

~~There are two ways to download a core dump file:~~

- ~~NFS~~
- ~~TFTP~~

~~Using NFS to Download a Core Dump~~

~~Use the following commands to download a core dump file via NFS:~~

```
(DCSS Switching) (Config)#exception protocol nfs
(DCSS Switching) (Config)#exception dump nfs 192.168.1.10://home/nfs_test
(DCSS Switching) (Config)#show exception

Coredump file name.............................. ASDF
Coredump filename uses hostname................. TRUE
Coredump filename uses time-stamp............... TRUE
NFS mount point................................. 192.168.1.10://home/nfs_test
TFTP server IP.................................. 10.27.9.99
File path....................................... ./.
Protocol........................................ nfs
Switch chip register.......................... TRUE
(DCSS Switching) (Config)#

(DCSS Switching) #write core test

The configured protocol nfs test PASS
(DCSS Switching) #
```

Using TFTP to Download a Core Dump

Use the following commands to download a core dump file via TFTP:

```
(DCSS Switching) (Config)#exception protocol tftp
(DCSS Switching) (Config)#exception dump tftp-server 192.168.1.2
(DCSS Switching) (Config)#show exception

Coredump file name.............................. core
Coredump filename uses hostname................. FALSE
Coredump filename uses time-stamp............... TRUE
TFTP server IP.................................. 192.168.1.2
File path....................................... ./.
Protocol........................................ tftp
Switch-chip-register.......................... FALSE
(DCSS Switching) (Config)#

(DCSS Switching) #write core test
The configured protocol tftp test PASS
(DCSS Switching) #
```

Setting the System Time

The switch uses the system clock to provide time stamps on log messages. Additionally, some show commands include the time in the command output. For example, the `show users login-history` command includes a Login Time field. The system clock provides the information for the Login Time field.

You can configure the system time manually, or you can configure the switch to obtain the time by using a Simple Network Time Protocol (SNTP) server. A network SNTP server can provide more accurate switch clock time synchronization than manually-configured time.

Note: The manually-configured local clock settings are not retained across a system reset if the platform does not include a Real Time Clock (RTC).

The SNTP client on the switch can request the time from an SNTP server on the network (unicast), or you can allow the switch to receive SNTP broadcasts. Requesting the time from a unicast SNTP server is more secure. Use this method if you know the IP address of the SNTP server on your network. If you allow the switch to receive SNTP broadcasts, any clock synchronization information is accepted, even if it has not been requested by the device. This method is less secure than polling a specified SNTP server.

The switch also supports the following time configuration settings:

- Time Zone — Allows you to specify the offset from Coordinated Universal Time (UTC), which is also known as Greenwich Mean Time (GMT).
- Summer Time/Daylight Saving Time (DST) — In some regions, the time shifts by one hour in the fall and spring. The switch supports manual entry of one-time or recurring shifts in the time.

Manual Time Configuration

The example in this section shows how to manually configure the time, date, time zone, and summer time settings for a switch in Hyderabad, India.

1. Set the time. The system clock uses a 24-hour clock, so 6:23 PM is entered as 18:23:00.

```
(DCSS Switching) #configure
(DCSS Switching) (Config)#clock set 18:23:00
```

2. Set the date. In this example, the date is April 30, 2012.

```
(DCSS Switching) (Config)#clock set 04/30/2012
```

3. Configure the time zone. In this example, the time zone is India Standard Time (IST), which is UTC/GMT +5 hours and 30 minutes.

```
(DCSS Switching) (Config)#clock timezone 5 minutes 30 zone IST
```

4. Configure the offset for a hypothetical daylight saving time. In this example, the offset is one hour. It occurs every year on Sunday in the first week of April and ends the fourth Sunday in October. The start and end times are 2:30 AM, and the time zone is India Standard Summer Time (ISST).

```
(DCSS Switching) (Config)#clock summer-time recurring 1 sun apr 02:30 4 sun oct 02:30 offset 60 zone
ISST
(DCSS Switching) (Config)#exit
```

5. View the clock settings.

```
(DCSS Switching) #show clock detail

20:30:07 ISST(UTC+6:30) Apr 30 2012
No time source

 Time zone:
 Acronym is IST
 Offset is UTC+5:30

 Summertime:
 Acronym is ISST
 Recurring every year
 Begins at first Sunday of Apr at 02:30
 Ends at fourth Sunday of Oct at 02:30
 offset is 60 minutes
```

Configuring SNTP

This example shows how to configure the system clock for a switch in New York City, which has a UTC/GMT offset of –5 hours.

1. Specify the SNTP server the client on the switch should contact. You can configure the IP address or host name of the SNTP server.

```
(DCSS Switching) #configure
(DCSS Switching) (Config)#sntp server time1.rtp.dcss.com
```

2. Configure the UTC/GMT offset for the location.

```
(DCSS Switching) (Config)#clock timezone -5
```

3. Configure the time offset for DST.

```
(DCSS Switching) (Config)#clock summer-time recurring USA
```

4. Enable the SNTP client on the device in unicast mode.

```
(DCSS Switching) (Config)#sntp client mode unicast
```

5. View the time information.

```
(DCSS Switching) #show sntp

Last Update Time:                 Apr 27 16:42:23 2012
Last Unicast Attempt Time:        Apr 27 16:43:28 2012
Last Attempt Status:              Success

(DCSS Switching) #show clock

12:47:22 (UTC-4:00) Apr 27 2012
Time source is SNTP
```

4 Security Features

Controlling Management Access

A user can access the switch management interface only after providing a valid username and password combination that matches the user account information stored in the user database configured on the switch.

DCSS software includes several additional features to increase management security and help prevent unauthorized access to the switch configuration interfaces.

Using RADIUS Servers for Management Security

Many networks use a RADIUS server to maintain a centralized user database that contains per-user authentication information. RADIUS servers provide a centralized authentication method for:

- Telnet Access
- Console to Switch Access
- Access Control Port (802.1X)

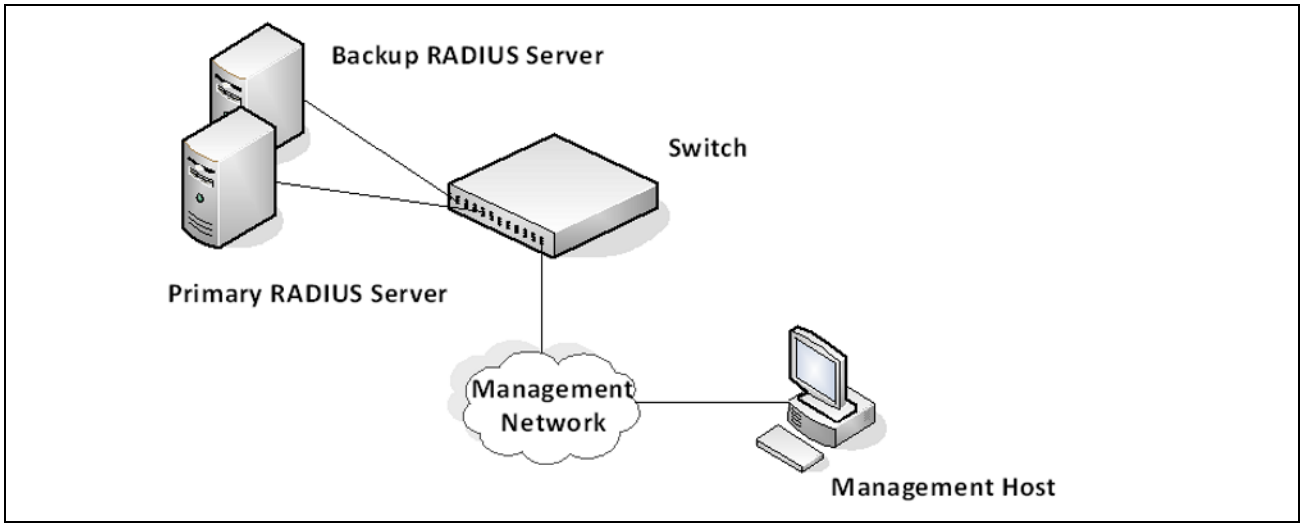
RADIUS access control utilizes a database of user information on a remote server. Making use of a single database of accessible information—as in an Authentication Server—can greatly simplify the authentication and management of users in a large network. One such type of Authentication Server supports the Remote Authentication Dial In User Service (RADIUS) protocol as defined by RFC 2865.

For authenticating users prior to access, the RADIUS standard has become the protocol of choice by administrators of large accessible networks. To accomplish the authentication in a secure manner, the RADIUS client and RADIUS server must both be configured with the same shared password or *secret*. This secret is used to generate one-way encrypted authenticators that are present in all RADIUS packets. The secret is never transmitted over the network.

RADIUS conforms to a secure communications client/server model using UDP as a transport protocol. It is extremely flexible, supporting a variety of methods to authenticate and statistically track users. RADIUS is also extensible, allowing for new methods of authentication to be added without disrupting existing functionality.

As a user attempts to connect to the switch management interface, the switch first detects the contact and prompts the user for a name and password. The switch encrypts the supplied information, and a RADIUS client transports the request to a pre-configured RADIUS server.

Figure 1: RADIUS Topology



The server can authenticate the user itself or make use of a back-end device to ascertain authenticity. In either case a response may or may not be forthcoming to the client. If the server accepts the user, it returns a positive result with attributes containing configuration information. If the server rejects the user, it returns a negative result. If the server rejects the client or the shared secrets differ, the server returns no result. If the server requires additional verification from the user, it returns a challenge, and the request process begins again.

If you use a RADIUS server to authenticate users, you must configure user attributes in the user database on the RADIUS server. The user attributes include the user name, password, and privilege level.

Using TACACS+ to Control Management Access

TACACS+ (Terminal Access Controller Access Control System) provides access control for networked devices via one or more centralized servers. TACACS+ simplifies authentication by making use of a single database that can be shared by many clients on a large network. TACACS+ uses TCP to ensure reliable delivery and a shared key configured on the client and daemon server to encrypt all messages.

If you configure TACACS+ as the authentication method for user login and a user attempts to access the user interface on the switch, the switch prompts for the user login credentials and requests services from the TACACS+ client. The client then uses the configured list of servers for authentication, and provides results back to the switch.

You can configure the TACACS+ server list with one or more hosts defined via their network IP address. You can also assign each a priority to determine the order in which the TACACS+ client will contact them. TACACS+ contacts the server when a connection attempt fails or times out for a higher priority server.

You can configure each server host with a specific connection type, port, timeout, and shared key, or you can use global configuration for the key and timeout.

The TACACS+ server can do the authentication itself, or redirect the request to another back-end device. All sensitive information is encrypted and the shared secret is never passed over the network; it is used only to encrypt the data.

Configuring and Applying Authentication Profiles

A user can access the switch management interface only after providing a valid username and password combination that matches the user account information stored in the user database configured on the switch.

DCSS software includes several additional features to increase management security and help prevent unauthorized access to the CLI.

An authentication profile specifies which authentication method or methods to use to authenticate a user who attempts to access the switch management interface. The profile includes a method list, which defines how authentication is to be performed, and in which order. The list specifies the authentication method to use first, and if the first method returns an error, the next method in the list is tried. This continues until all methods in the list have been attempted. If no method can perform the authentication, then the authentication fails. A method might return an error if, for example, the authentication server is unreachable or misconfigured.

The authentication method can be one or more of the following:

- enable—Uses the enable password for authentication. If there is no enable password defined, then the enable method returns an error.
- line—Uses the Line password for authentication. If there is no line password defined for the access line, then the line method returns an error.
- local— Uses the ID and password in the Local User Database for authentication. If the user ID is not in the local database, access is denied. This method never returns an error. It always permits or denies a user.
- radius—Sends the user's ID and password to a RADIUS server to be authenticated. The method returns an error if the switch is unable to contact the server.
- tacacs+— Sends the user's ID and password to a TACACS+ server to be authenticated. The method returns an error if the switch is unable to contact the server.
- none—No authentication is used. This method never returns an error.
- deny —Access is denied. This method never returns an error.

An authentication method might require a user name and password to be supplied, a password only, or no user information. Some methods return errors when authentication fails, while other methods do not. The following table summarizes the method user name/password requirements and error behavior.

Table 5: Authentication Method Summary

Method	*Username Required*	*Password Required*	*Error Returned*
Local	Yes	Yes	No
RADIUS	Yes	Yes	Yes
TACACS+	Yes	Yes	Yes
Enable	No	Yes	Yes
Line	No	Yes	Yes
None	No	No	No
Deny	No	No	No

You can use the same Authentication Profile for all access types, or select or create a variety of profiles based on how a user attempts to access the switch management interface. Profiles can be applied to each of the following access types:

- Login—Authenticates all attempts to login to the switch.
- Enable—Authenticates all attempts to enter Privileged EXEC mode.
- Console—Authenticates access through the console port.
- Telnet—Authenticates users accessing the CLI by using telnet
- SSH—Authenticates users accessing the CLI by using an SSH client.

The following authentication profiles are configured by default:

- defaultList—Method is LOCAL, which means the user credentials are verified against the information in the local user database.
- networkList—Method is LOCAL, which means the user credentials are verified against the information in the local user database.
- enableList—Method is ENABLE, followed by NONE, which means that if the *enable* password is not configured access is granted. If the enable password is configured and user fails to authenticate then access is not granted.
- enableNetList—Method is ENABLE, followed by DENY, which means that if the *enable* password is not configured access is denied. This list is applied by default for telnet and SSH. In DCSS the enable password is not configured by default. That means that, by default, telnet and SSH users will not get access to Privileged EXEC mode. However, a console user always enters the Privileged EXEC mode without entering the enable password in the default configuration.

The methods can be changed, but the preconfigured profiles cannot be deleted or renamed.

Configuring Authentication Profiles for Port-Based Authentication

In addition to authentication profiles to control access to the management interface, you can configure an authentication profile for IEEE 802.1X port-based access control to control access to the network through the switch ports. To configure a port-based authentication profile, you specify `dot1x` as the access type, and configure `ias`, `local`, `none`, or `radius` as the authentication method. The `ias` method specifies that the 802.1X feature should use the Internal Authentication Server (IAS) database for 801X port-based authentication. The IAS database is stored locally on the switch.

Configuring the Primary and Secondary RADIUS Servers

The commands in this example configure primary and secondary RADIUS servers that the switch will use to authenticate access. The RADIUS servers use the same RADIUS secret.

To configure the switch:

1. Configure the primary and secondary RADIUS servers.

```
(DCSS Switching)#configure
(DCSS Switching) (Config)#radius server host auth 10.27.65.103
(DCSS Switching) (Config)#radius server host auth 10.27.65.114
```

2. Specify which RADIUS server is the primary.

```
(DCSS Switching) (Config)#radius server primary 10.27.65.103
(DCSS Switching) (Config)#radius server key auth 10.27.65.103
```

3. Configure a shared secret that the switch will use to authenticate with the RADIUS servers.

```
Enter secret (64 characters max):********

Re-enter secret:********
```

4. View the configured RADIUS servers.

```
(DCSS Switching) (Config)#exit
(DCSS Switching) #show radius servers

Cur
rent Host Address             Server Name                     Port  Type
---- ---------------------- ------------------------------- ----- ---------
     10.27.65.114             Default-RADIUS-Server           1812  Secondary
  *  10.27.65.103             Default-RADIUS-Server           1812  Primary
```

Configuring an Authentication Profile

The commands in this example create a new authentication profile named myList that uses the RADIUS server configured in the previous example to authenticate users who attempt to access the switch management interface by using SSH or Telnet. If the RADIUS authentication is unsuccessful, the switch uses the local user database to attempt to authenticate the users.

To configure the switch:

1. Create an access profile list that uses RADIUS as the first access method and the local user database as the second login method.

```
(DCSS Switching) #configure
(DCSS Switching) (Config)#aaa authentication login myList radius local
```

Note: The switch attempts to contact the primary RADIUS server that has been configured on the switch. To see an example of how to configure a RADIUS server on the switch, see "Configuring the Primary and Secondary RADIUS Servers" on page 60

2. Enter line configuration mode for Telnet and specify that any attempt to access the switch by using Telnet are authenticated using the methods defined in the profile created in the previous step.

```
(DCSS Switching) (Config)#line telnet
(DCSS Switching) (Config-telnet)#login authentication myList
(DCSS Switching) (Config-telnet)#exit
```

3. Enter line configuration mode for SSH and specify that any attempt to access the switch by using SSH are authenticated using the methods defined in the myList profile.

```
(DCSS Switching) (Config)#line ssh
(DCSS Switching) (Config-ssh)#login authentication myList
(DCSS Switching) (Config-ssh)#end
```

4. View the current authentication methods and profiles.

```
(DCSS Switching) #show authentication methods

Login Authentication Method Lists
---------------------------------
defaultList         :  local
networkList         :  local
```

```
myList              :  radius   local

Enable Authentication Method Lists
----------------------------------
enableList          :  enable   none
enableNetList       :  enable   deny

Line     Login Method List    Enable Method List
-------  -----------------    ------------------
Console  defaultList          enableList
Telnet   myList               enableList
SSH      myList               enableList
```

Configuring DHCP Snooping, DAI, and IPSG

Dynamic Host Configuration Protocol (DHCP) Snooping, IP Source Guard (IPSG), and Dynamic ARP Inspection (DAI) are layer 2 security features that examine traffic to help prevent accidental and malicious attacks on the switch or network.

DHCP Snooping monitors DHCP messages between a DHCP client and DHCP server to filter harmful DHCP messages and to build a bindings database. The IPSG and DAI features use the DHCP Snooping bindings database to help enforce switch and network security.

IP Source Guard allows the switch to drop incoming packets that do not match a binding in the bindings database. Dynamic ARP Inspection allows the switch to drop ARP packets whose sender MAC address and sender IP address do not match an entry in the DHCP snooping bindings database.

DHCP Snooping Overview

Dynamic Host Configuration Protocol (DHCP) Snooping is a security feature that monitors DHCP messages between a DHCP client and DHCP server to accomplish the following tasks:

- Filter harmful DHCP messages
- Build a bindings database with entries that consist of the following information:
  - MAC address
  - IP address
  - VLAN ID
  - Client port

Entries in the bindings database are considered to be authorized network clients.

DHCP snooping can be enabled on VLANs, and the trust status (trusted or untrusted) is specified on individual physical ports or LAGS that are members of a VLAN. When a port or LAG is configured as untrusted, it could potentially be used to launch a network attack. DHCP servers must be reached through trusted ports.

DHCP snooping enforces the following security rules:

- DHCP packets from a DHCP server (DHCPOFFER, DHCPACK, DHCPNAK, DHCPRELEASEQUERY) are dropped if they are received on an untrusted port.
- DHCPRELEASE and DHCPDECLINE messages are dropped if the MAC addresses is in the snooping database, but the binding's interface is other than the interface where the message was received.
- On untrusted interfaces, the switch drops DHCP packets with a source MAC address that does not match the client hardware address. This is a configurable option.

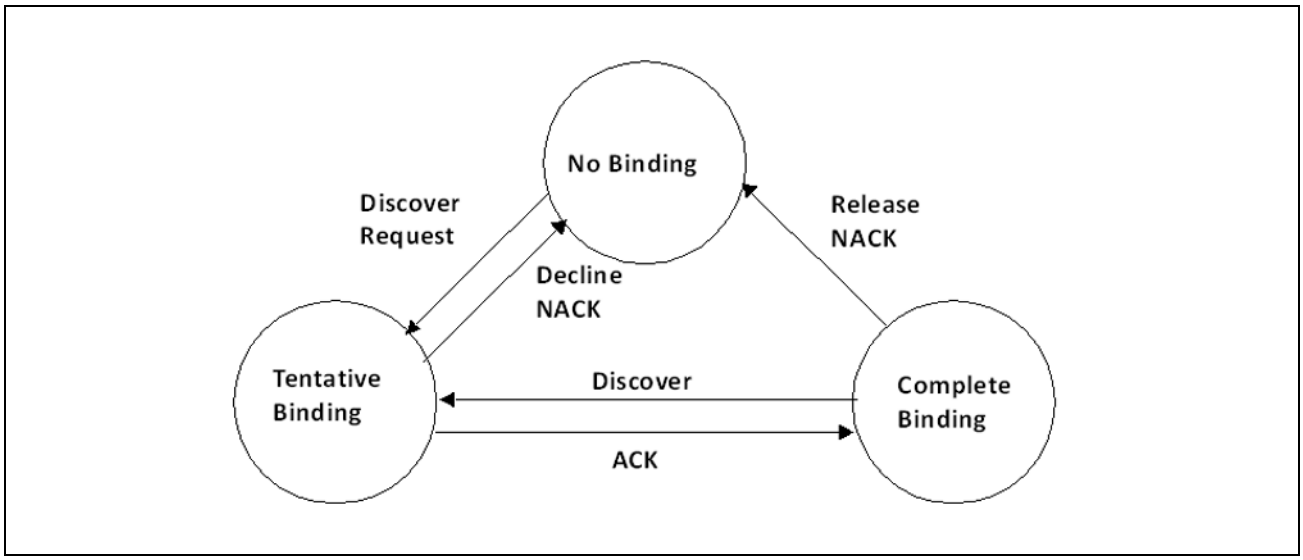
Populating the DHCP Snooping Bindings Database

The DHCP snooping application uses DHCP messages to build and maintain the binding's database. DHCP snooping creates a tentative binding from DHCP DISCOVER and REQUEST messages. Tentative bindings tie a client to a port (the port where the DHCP client message was received). Tentative bindings are completed when DHCP snooping learns the client's IP address from a DHCP ACK message on a trusted port. DHCP snooping removes bindings in response to DECLINE, RELEASE, and NACK messages. The DHCP snooping application ignores the ACK messages as a reply to the DHCP Inform messages received on trusted ports. You can also enter static bindings into the binding database.

When a switch learns of new bindings or loses bindings, the switch immediately updates the entries in the database. The switch also updates the entries in the binding file. The frequency at which the file is updated is based on a configurable delay, and the updates are batched.

If the absolute lease time of the snooping database entry expires, that entry is removed. Make sure the system time is consistent across the reboots. Otherwise, the snooping entries will not expire properly. If a host sends a DHCP release while the switch is rebooting, when the switch receives the DHCP discovery or request, the client's binding goes to the tentative binding as shown in Figure 2 on page 63.

Figure 2: DHCP Binding



The binding database includes data for clients only on untrusted ports.

DHCP Snooping and VLANs

DHCP snooping forwards valid DHCP client messages received on non-routing VLANs. The message is forwarded on all trusted interfaces in the VLAN.

DHCP snooping can be configured on switching VLANs and routing VLANs. When a DHCP packet is received on a routing VLAN, the DHCP snooping application applies its filtering rules and updates the bindings database. If a client message passes filtering rules, the message is placed into the software forwarding path where it may be processed by the DHCP relay agent, the local DHCP server, or forwarded as an IP packet.

DHCP Snooping Logging and Rate Limits

The DHCP snooping application processes incoming DHCP messages. For DHCPRELEASE and DHCPDECLINE messages, the application compares the receive interface and VLAN with the client interface and VLAN in the bindings database. If the interfaces do not match, the application logs the event and drops the message. For valid client messages, DHCP snooping compares the source MAC address to the DHCP client hardware address. When there is a mismatch, DHCP snooping drops the packet and generates a log message if logging of invalid packets is enabled.

If DHCP relay co-exists with DHCP snooping, DHCP client messages are sent to DHCP relay for further processing.

To prevent DHCP packets from being used as a DoS attack when DHCP snooping is enabled, the snooping application enforces a rate limit for DHCP packets received on interfaces. DHCP snooping monitors the receive rate on each interface separately. If the receive rate exceeds a configurable limit, DHCP snooping brings down the interface. Administrative intervention is necessary to enable the port, by using the **no shutdown** command in Interface Config mode.

IP Source Guard Overview

IPSG is a security feature that filters IP packets based on source ID. This feature helps protect the network from attacks that use IP address spoofing to compromise or overwhelm the network.

The source ID may be either the source IP address or a {source IP address, source MAC address} pair. You can configure:

- Whether enforcement includes the source MAC address
- Static authorized source IDs

The DHCP snooping bindings database and static IPSG entries identify authorized source IDs. IPSG can be enabled on physical and LAG ports.

If you enable IPSG on a port where DHCP snooping is disabled or where DHCP snooping is enabled but the port is trusted, all IP traffic received on that port is dropped depending on the admin-configured IPSG entries.

IPSG and Port Security

IPSG interacts with port security, also known as port MAC locking to enforce the source MAC address. Port security controls source MAC address learning in the layer 2 forwarding database (MAC address table). When a frame is received with a previously unlearned source MAC address, port security queries the IPSG feature to determine whether the MAC address belongs to a valid binding.

If IPSG is disabled on the ingress port, IPSG replies that the MAC is valid. If IPSG is enabled on the ingress port, IPSG checks the bindings database. If the MAC address is in the bindings database and the binding matches the VLAN the frame was received on, IPSG replies that the MAC is valid. If the MAC is not in the bindings database, IPSG informs port security that the frame is a security violation.

In the case of an IPSG violation, port security takes whatever action it normally takes upon receipt of an unauthorized frame. Port security limits the number of MAC addresses to a configured maximum. If the limit *n* is less than the number of stations *m* in the bindings database, port security allows only *n* stations to use the port. If $n > m$, port security allows only the stations in the bindings database.

Dynamic ARP Inspection Overview

Dynamic ARP Inspection (DAI) is a security feature that rejects invalid and malicious ARP packets. DAI prevents a class of man-in-the-middle attacks where an unfriendly station intercepts traffic for other stations by poisoning the ARP caches of its unsuspecting neighbors. The malicious attacker sends ARP requests or responses mapping another station's IP address to its own MAC address.

When DAI is enabled, the switch drops ARP packets whose sender MAC address and sender IP address do not match an entry in the DHCP snooping bindings database. You can optionally configure additional ARP packet validation.

When DAI is enabled on a VLAN, DAI is enabled on the interfaces (physical ports or LAGs) that are members of that VLAN. Individual interfaces are configured as trusted or untrusted. The trust configuration for DAI is independent of the trust configuration for DHCP snooping.

Optional DAI Features

If you configure the MAC address validation option, DAI verifies that the sender MAC address equals the source MAC address in the Ethernet header. There is a configurable option to verify that the target MAC address equals the destination MAC address in the Ethernet header. This check applies only to ARP responses, since the target MAC address is unspecified in ARP requests. You can also enable IP address checking. When this option is enabled, DAI drops ARP packets with an invalid IP address. The following IP addresses are considered invalid:

- 0.0.0.0
- 255.255.255.255
- all IP multicast addresses
- all class E addresses (240.0.0.0/4)
- loopback addresses (in the range 127.0.0.0/8)

The valid IP check is applied only on the sender IP address in ARP packets. In ARP response packets, the check is applied only on the target IP address.

Increasing Security with DHCP Snooping, DAI, and IPSG

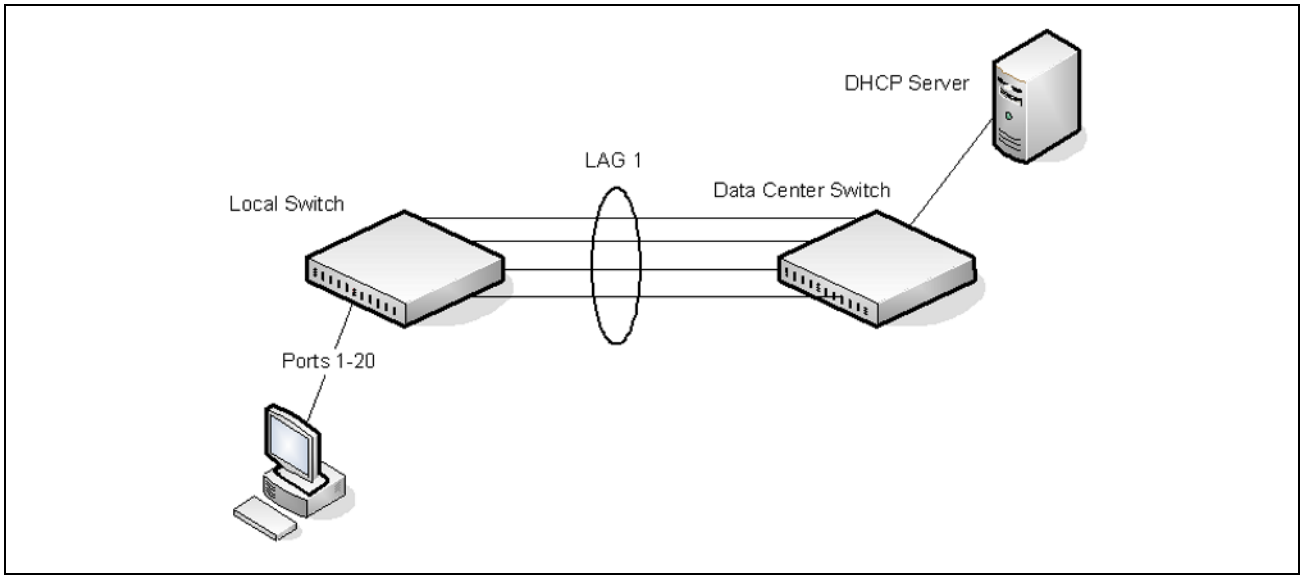
DHCP Snooping, IPSG, and DAI are security features that can help protect the switch and the network against various types of accidental or malicious attacks. It might be a good idea to enable these features on ports that provide network access to hosts that are in physically unsecured locations or if network users connect nonstandard hosts to the network.

For example, if an employee unknowingly connects a workstation to the network that has a DHCP server, and the DHCP server is enabled, hosts that attempt to acquire network information from the legitimate network DHCP server might obtain incorrect information from the rogue DHCP server. However, if the workstation with the rogue DHCP server is connected to a port that is configured as untrusted and is a member of a DHCP Snooping-enabled VLAN, the port discards the DHCP server messages.

Configuring DHCP Snooping

In this example, DHCP snooping is enabled on VLAN 100. Ports 1-20 connect end users to the network and are members of VLAN 100. These ports are configured to limit the maximum number of DHCP packets with a rate limit of 100 packets per second. LAG 1, which is also a member of VLAN 100 and contains ports 21-24, is the trunk port that connects the switch to the data center, so it is configured as a trusted port.

Figure 3: DHCP Snooping Configuration Topology



The commands in this example also enforce rate limiting and remote storage of the bindings database. The switch has a limited amount of storage space in NVRAM and flash memory, so the administrator specifies that the DHCP snooping bindings database is stored on an external TFTP server.

To configure the switch:

1. Enable DHCP snooping on VLAN 100.

```
(DCSS Switching) #config
(DCSS Switching) (Config)#ip dhcp snooping vlan 100
```

2. Configure LAG 1, which includes ports 21-24, as a trusted port. All other interfaces are untrusted by default.

```
(DCSS Switching) (Config)#interface 3/1
(DCSS Switching) (Interface 3/1))#ip dhcp snooping trust
(DCSS Switching) (Interface 3/1)#exit
```

3. Enter interface configuration mode for all untrusted interfaces (ports 1-20) and limit the number of DHCP packets that an interface can receive to 100 packets per second. LAG 1 is a trusted port and keeps the default value for rate limiting (unlimited).

```
(DCSS Switching) (Config)#interface 0/1-0/20
(DCSS Switching) (Interface 0/1-0/20)#ip dhcp snooping limit rate 100
(DCSS Switching) (Interface 0/1-0/20)#exit
```

4. Specify that the DHCP snooping database is to be stored remotely in a file called dsDb.txt on a TFTP server with and IP address of 10.131.11.1.

```
(DCSS Switching) (Config)#ip dhcp snooping database tftp://10.131.11.1/dsDb.txt
```

5. Enable DHCP snooping for the switch

```
(DCSS Switching) (Config)#ip dhcp snooping
(DCSS Switching) (Config)#exit
```

6. View DHCP snooping information.

```
(DCSS Switching) #show ip dhcp snooping

DHCP snooping is Enabled
DHCP snooping source MAC verification is enabled
DHCP snooping is enabled on the following VLANs:
100

 Interface    Trusted     Log Invalid Pkts
-----------  ----------   ----------------
```

Configuring IPSG

This example builds on the previous example and uses the same topology shown in Figure 3 on page 66. In this configuration example, IP source guard is enabled on ports 1-20. DHCP snooping must also be enabled on these ports. Additionally, because the ports use IP source guard with source IP and MAC address filtering, port security must be enabled on the ports as well.

To configure the switch:

1. Enter interface configuration mode for the host ports and enable IPSG.

```
(DCSS Switching) #config
(DCSS Switching) (Config)#interface 0/1-0/20
(DCSS Switching) (Interface 0/1-0/20)#ip verify source port-security
```

2. Enable port security on the ports.

```
(DCSS Switching) (Interface 0/1-0/20)#port-security
(DCSS Switching) (Interface 0/1-0/20)#end
```

3. View IPSG information.

```
(DCSS Switching) #show ip verify source

Interface      Filter Type    IP Address        MAC Address           VLAN
-----------    -----------    ---------------   -----------------     -----
0/1            ip-mac         192.168.3.45      00:1C:23:55:D4:8E     100
0/2            ip-mac         192.168.3.33      00:1C:23:AA:B8:01     100
0/3            ip-mac         192.168.3.18      00:1C:23:55:1B:6E     100
0/4            ip-mac         192.168.3.49      00:1C:23:67:D3:CC     100
--More-- or (q)uit
```

5 Configuring Switching

VLANs

By default, all switch ports on the switch are in the same broadcast domain. This means when one host connected to the switch broadcasts traffic, every device connected to the switch receives that broadcast. All ports in a broadcast domain also forward multicast and unknown unicast traffic to the connected host. Large broadcast domains can result in network congestion, and end users might complain that the network is slow. In addition to latency, large broadcast domains are a greater security risk since all hosts receive all broadcasts.

Virtual Local Area Networks (VLANs) allow you to divide a broadcast domain into smaller, logical networks. Like a bridge, a VLAN switch forwards traffic based on the Layer 2 header, which is fast, and like a router, it partitions the network into logical segments, which provides better administration, security, and management of multicast traffic.

Network administrators have many reasons for creating logical divisions, such as department or project membership. Because VLANs enable logical groupings, members do not need to be physically connected to the same switch or network segment. Some network administrators use VLANs to segregate traffic by type so that time-sensitive traffic, like voice traffic, has priority over other traffic, such as data. Administrators also use VLANs to protect network resources. Traffic sent by authenticated clients might be assigned to one VLAN, while traffic sent from unauthenticated clients might be assigned to a different VLAN that allows limited network access.

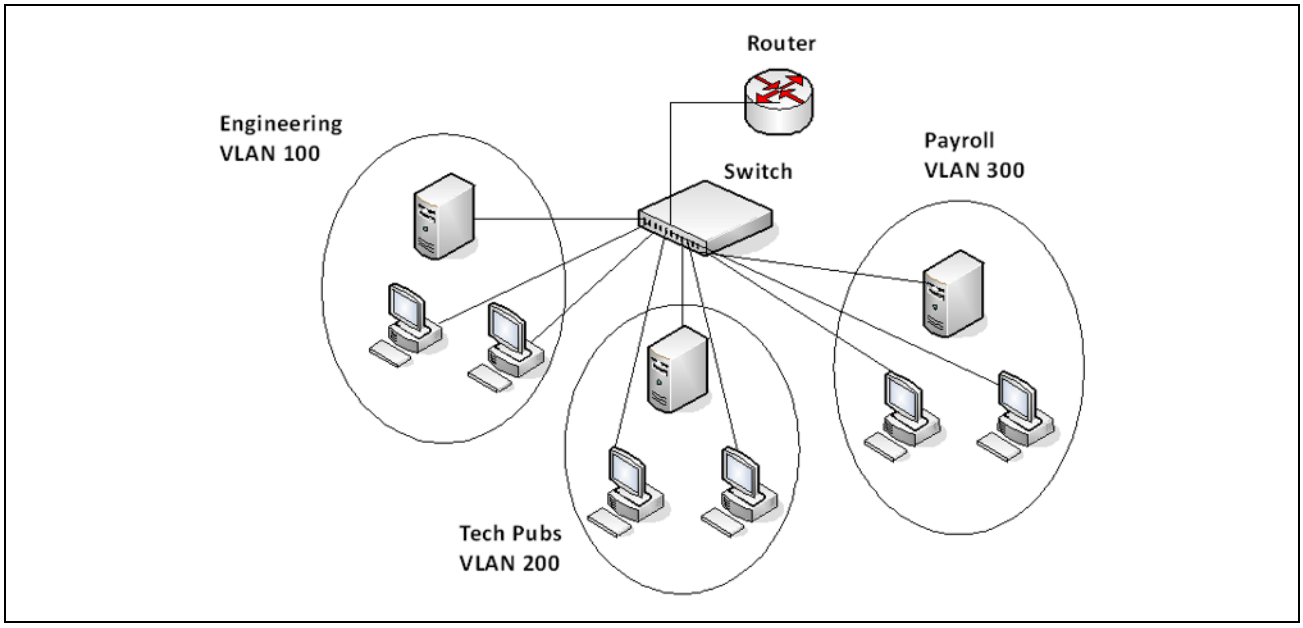
When one host in a VLAN sends a broadcast, the switch forwards traffic only to other members of that VLAN. For traffic to go from a host in one VLAN to a host in a different VLAN, the traffic must be forwarded by a layer 3 device, such as a router. VLANs work across multiple switches, so there is no requirement for the hosts to be located near each other to participate in the same VLAN.

Each VLAN has a unique number, called the VLAN ID. The DCSS supports a configurable VLAN ID range of 2–4093. A VLAN with VLAN ID 1 is configured on the switch by default. You can associate a name with the VLAN ID. In a tagged frame, the VLAN is identified by the VLAN ID in the tag. In an untagged frame, the VLAN identifier is the Port VLAN ID (PVID) specified for the port that received the frame. For information about tagged and untagged frames, see “VLAN Tagging” on page 70.

DCSS supports adding individual ports and Link Aggregation Groups (LAGs) as VLAN members.

Figure 4 on page 70 shows an example of a network with three VLANs that are department-based. The file server and end stations for the department are all members of the same VLAN.

Figure 4: Simple VLAN Topology



In this example, each port is manually configured so that the end station attached to the port is a member of the VLAN configured for the port. The VLAN membership for this network is port-based or static.

VLAN Tagging

DCSS supports IEEE 802.1Q tagging. Ethernet frames on a tagged VLAN have a 4-byte VLAN tag in the header. VLAN tagging is required when a VLAN spans multiple switches, which is why trunk ports transmit and receive only tagged frames.

Tagging may be required when a single port supports multiple devices that are members of different VLANs. For example, a single port might be connected to an IP phone, a PC, and a printer (the PC and printer are connected via ports on the IP phone). IP phones are typically configured to use a tagged VLAN for voice traffic, while the PC and printers typically use the untagged VLAN.

When a port is added to a VLAN as an untagged member, untagged packets entering the switch are tagged with the PVID (also called the *native VLAN*) of the port. If the port is added to a VLAN as an untagged member, the port does not add a tag to a packet in that VLAN when it exits the port. Configuring the PVID for an interface is useful when untagged and tagged packets will be sent and received on that port and a device connected to the interface does not support VLAN tagging.

When ingress filtering is on, the frame is dropped if the port is not a member of the VLAN identified by the VLAN ID in the tag. If ingress filtering is off, all tagged frames are forwarded. The port decides whether to forward or drop the frame when the port receives the frame.

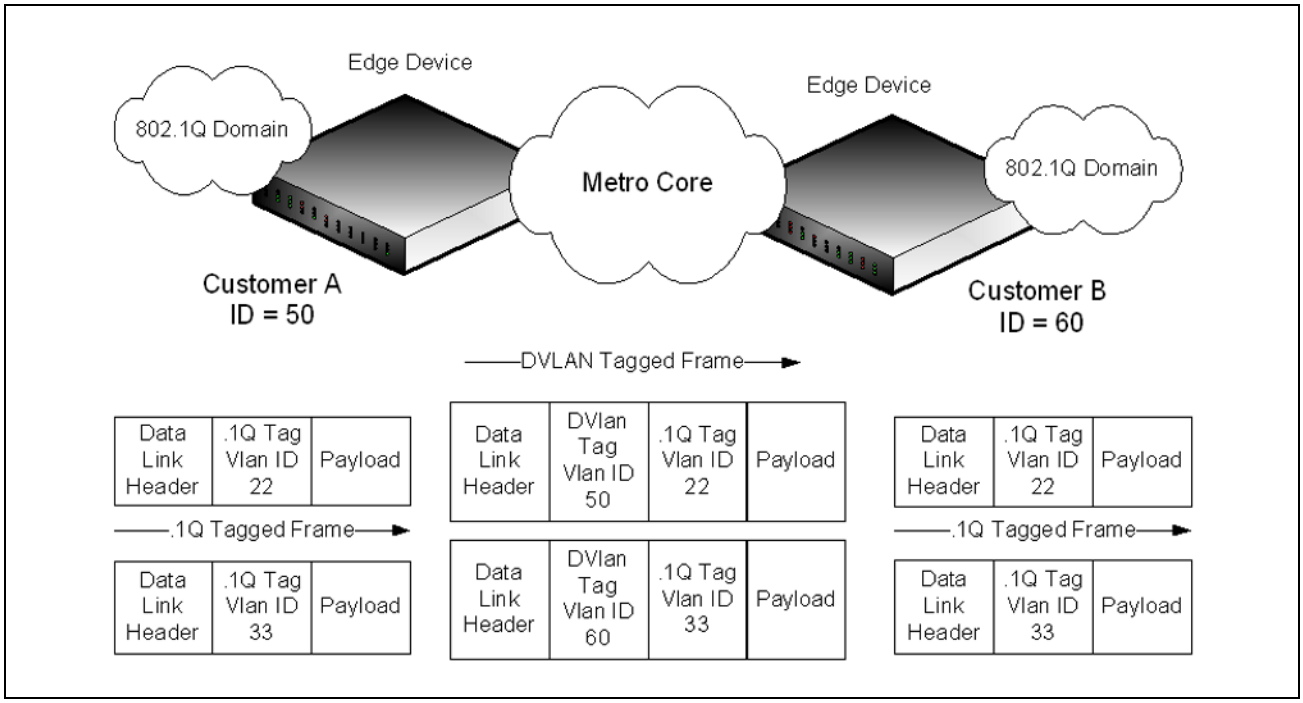
Double-VLAN Tagging

For trunk ports, which are ports that connect one switch to another switch, DCSS software supports double-VLAN tagging. This feature allows service providers to create Virtual Metropolitan Area Networks (VMANs). With double-VLAN tagging, service providers can pass VLAN traffic from one customer domain to another through a metro core in a simple and cost-effective manner. By using an additional tag on the traffic, the switch can differentiate between customers in the MAN while preserving an individual customer's VLAN identification when the traffic enters the customer's 802.1Q domain.

With the introduction of this second tag, customers are no longer required to divide the 4-byte VLAN ID space to send traffic on a Ethernet-based MAN. In short, every frame that is transmitted from an interface has a double-VLAN tag attached, while every packet that is received from an interface has a tag removed (if one or more tags are present).

In Figure 5, two customers share the same metro core. The service provider assigns each customer a unique ID so that the provider can distinguish between the two customers and apply different rules to each. When the configurable EtherType is assigned to something different than the 802.1Q (0x8100) EtherType, it allows the traffic to have added security from misconfiguration while exiting the metro core. For example, if the edge device on the other side of the metro core is not stripping the second tag, the packet would never be classified as a 802.1Q tag, so the packet would be dropped rather than forwarded in the incorrect VLAN.

Figure 5: Double VLAN Tagging Network Example



Default VLAN Behavior

One VLAN exists on the switch by default. The VLAN ID is 1, and all ports are included in the VLAN as access ports, which are untagged. This means when a device connects to any port on the switch, the port forwards the packets without inserting a VLAN tag. If a device sends a tagged frame to a port, the frame is dropped. Since all ports are members of this VLAN, all ports are in the same broadcast domain and receive all broadcast and multicast traffic received on any port.

When you add a new VLAN to the VLAN database, no ports are members. The configurable VLAN range is 2–4093. VLANs 4094 and 4095 are reserved.

Table 6 shows the default values or maximum values for VLAN features.

Table 6: VLAN Default and Maximum Values

Feature	*Value*
Default VLAN ID	1
VLAN Name	default
VLAN Range	2–4093
Frames accepted	Untagged Incoming untagged frames are classified into the VLAN whose VLAN ID is the currently configured PVID.
Frames sent	Untagged
Ingress Filtering	On
PVID	1
Double-VLAN tagging	Disabled If double-VLAN tagging is enabled, the default EtherType value is 802.1Q

VLAN Configuration Example

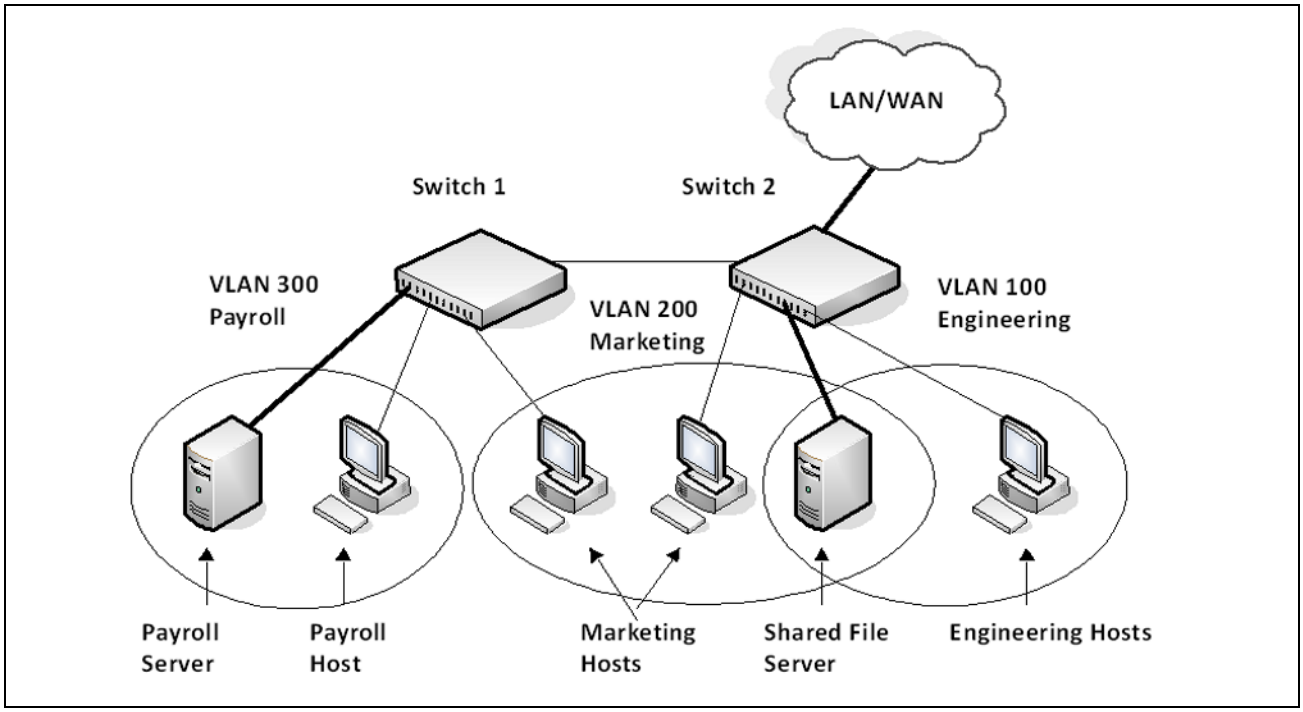
A network administrator wants to create the VLANs in Table 7:

Table 7: Example VLANs

VLAN ID	*VLAN Name*	*VLAN Type*	*Purpose*
100	Engineering	Port-based	All employees in the Engineering department use this VLAN. Confining this department's traffic to a single VLAN helps reduce the amount of traffic in the broadcast domain, which increases bandwidth.
200	Marketing	Port-based	All employees in the Marketing department use this VLAN.
300	Payroll	Port-based	The payroll department has sensitive traffic and needs its own VLAN to help keep that traffic private.

Figure 6 on page 73 shows the network topology for this example. As the figure shows, there are two switches, two file servers, and many hosts. One switch has an uplink port that connects it to a layer 3 device and the rest of the corporate network.

Figure 6: Network Topology for VLAN Configuration



The network in Figure 6 has the following characteristics:

- Each connection to a host represents multiple ports and hosts.
- The Payroll and File servers are connected to the switches through a LAG.
- Some of the Marketing hosts connect to Switch 1, and some connect to Switch 2.
- The Engineering and Marketing departments share the same file server.
- Because security is a concern for the Payroll VLAN, the ports and LAG that are members of this VLAN will accept and transmit only traffic tagged with VLAN 300.

Table 8 shows the port assignments on the switches.

Table 8: Switch Port Connections

Port/LAG	*Function*
Switch 1	
1	Connects to Switch 2
2–15	Host ports for Payroll
16–20	Host ports for Marketing
LAG1 (ports 21–24)	Connects to Payroll server
Switch 2	
1	Connects to Switch 1
2–10	Host ports for Marketing
11–30	Host ports for Engineering
LAG1 (ports 35–39)	Connects to file server
LAG2 (ports 40–44)	Uplink to router.

Configure the VLANs and Ports on Switch 1

Use the following steps to configure the VLANs and ports on Switch 1. None of the hosts that connect to Switch 1 use the Engineering VLAN (VLAN 100), so it is not necessary to create it on that switch.

To configure Switch 1:

1. Create VLANs 200 (Marketing), 300 (Payroll), and associate the VLAN ID with the appropriate name.

```
(DCSS Switching) #vlan database
(DCSS Switching) (Vlan)#vlan 200,300
(DCSS Switching) (Vlan)#vlan name 200 Marketing
(DCSS Switching) (Vlan)#vlan name 300 Payroll
(DCSS Switching) (Vlan)#exit
```

2. Assign ports 16–20 to the Marketing VLAN.

```
(DCSS Switching) #configure
(DCSS Switching) (Config)#interface 0/16-0/20
(DCSS Switching) (Interface 0/16-0/20)#vlan participation include 200
(DCSS Switching) (Interface 0/16-0/20)#exit
```

3. Assign ports 2–15 to the Payroll VLAN

```
(DCSS Switching) (Config)#interface 0/2-0/15
(DCSS Switching) (Interface 0/2-0/15)#vlan participation include 300
(DCSS Switching) (Interface 0/2-0/15)#exit
```

4. Assign LAG1 to the Payroll VLAN and specify that frames will always be transmitted tagged with a PVID of 300.

```
(DCSS Switching) (Config)#interface 3/1
(DCSS Switching) (Interface 3/1)#vlan participation include 300
(DCSS Switching) (Interface 3/1)#vlan tagging 300
(DCSS Switching) (Interface 3/1)#vlan pvid 300
(DCSS Switching) (Interface 3/1)#exit
```

5. Configure port 1 as a trunk port and add VLAN 200 and VLAN 300 as members. Trunk ports accept and transmits tagged frames only and have ingress filtering enabled.

```
(DCSS Switching) (Config)#interface 0/1
(DCSS Switching) (Interface 0/1)#vlan acceptframe vlanonly
(DCSS Switching) (Interface 0/1)#vlan participation include 200,300
(DCSS Switching) (Interface 0/1)#vlan participation exclude 1
(DCSS Switching) (Interface 0/1)#vlan tagging 200,300
(DCSS Switching) (Interface 0/1)#vlan ingressfilter
(DCSS Switching) (Interface 0/1)#end
```

6. To save the configuration so that it persists across a system reset, use the following command:

```
(DCSS Switching) #copy system:running-config nvram:startup-config
```

7. View the VLAN settings.

```
(DCSS Switching) #show vlan

VLAN ID VLAN Name                         VLAN Type
------- --------------------------------- -------------------
1       default                           Default
200     Marketing                         Static
300     Payroll                           Static

(DCSS Switching) #show vlan 300

VLAN ID: 300
VLAN Name: Payroll
VLAN Type: Static

Interface   Current   Configured   Tagging
----------  --------  -----------  --------
0/1         Include   Include      Tagged
0/2         Include   Include      Untagged
0/3         Include   Include      Untagged
0/4         Include   Include      Untagged
0/5         Include   Include      Untagged
--More-- or (q)uit
```

8. View the VLAN information for a port.

```
(DCSS Switching) #show vlan port 0/1

          Port       Port                 Ingress    Ingress
          VLAN ID    VLAN ID  Acceptable  Filtering  Filtering Default
Interface Configured Current  Frame Types Configured Current   Priority
--------- ---------- -------- ----------- ---------- --------- --------
0/1       1          1        VLAN Only   Enable     Enable

Protected Port ............................. False
```

Configure the VLANs and Ports on Switch 2

Use the following steps to configure the VLANs and ports on Switch 2. Many of the procedures in this section are the same as procedures used to configure Switch 1. For more information about specific procedures, see the details and figures in the previous section.

To configure Switch 2:

1. Create the Engineering, Marketing, and Payroll VLANs.

   Although the Payroll hosts do not connect to this switch, traffic from the Payroll department must use Switch 2 to reach the rest of the network and Internet through the uplink port. For that reason, Switch 2 must be aware of VLAN 300 so that traffic is not rejected by the trunk port.
2. Configure ports 2-10 to participate in VLAN 200.
3. Configure ports 11–30 to participate in VLAN 100.
4. Configure LAG 1 to participate in VLAN 100 and VLAN 200.
5. Configure port 1 and LAG 2 as participants in ports and add VLAN 100, VLAN 200, and VLAN 300 that accept and transit tagged frames only.
6. Enable ingress filtering on port 1 and LAG 2.
7. If desired, copy the running configuration to the startup configuration.
8. View VLAN information for the switch and ports.

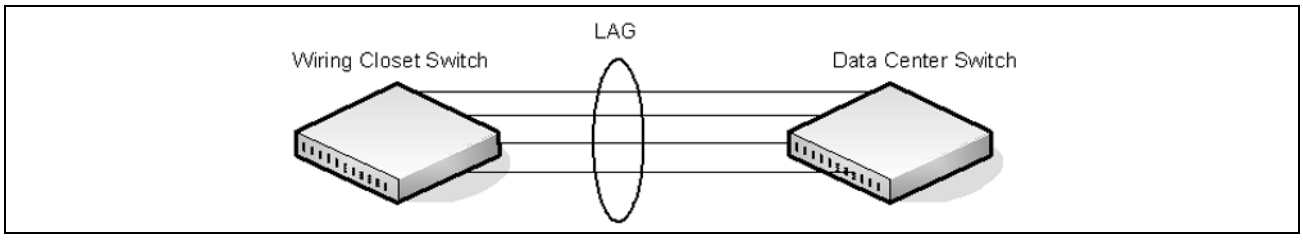
LAGs

Link Aggregation allows one or more full-duplex (FDX) Ethernet links of the same speed to be aggregated together to form a Link Aggregation Group (LAG). This allows the switch to treat the LAG as if it is a single link. The primary purpose of LAGs is to increase the overall bandwidth between two switches. This is accomplished by effectively aggregating multiple ports together that act as a single, logical connection between the two switches. LAGs also provide redundancy. If a link fails, traffic is automatically redistributed across the remaining links.

DCSS software supports industry-standard LAGs that adhere to the IEEE 802.3ad specification. Both static and dynamic LAGs are supported. Each LAG can have a maximum of 8 ports as members. You can configure LAGs until all switch ports are assigned to a LAG.

Figure 7 shows an example of a switch in the wiring closet connected to a switch in the data center by a LAG that consists of four physical 10 Gbps links. The LAG provides full-duplex bandwidth of 40 Gbps between the two switches.

Figure 7: LAG Configuration



Static and Dynamic Link Aggregation

Link aggregation can be configured as either dynamic or static. Dynamic configuration is supported using the IEEE 802.3ad standard, which is known as Link Aggregation Control Protocol (LACP). Static configuration is used when connecting the switch to an external Gigabit Ethernet switch that does not support LACP.

One advantage of LACP is that the protocol enables the switch to confirm that the external switch is also configured for link aggregation. When using static configuration, a cabling or configuration mistake involving the local switch or the external switch could go undetected and thus cause undesirable network behavior. Both static and dynamic LAGs (via LACP) can detect physical link failures within the LAG and continue forwarding traffic through the other connected links within that same LAG. LACP can also detect switch or port failures that do not result in loss of link. This provides a more resilient LAG. Best practices suggest using dynamic link aggregation instead of static link aggregation. When a port is added to a LAG as a static member, it neither transmits nor receives LACP PDUs.

LAG Hashing

DCSS software supports configuration of hashing algorithms for each LAG interface. The hashing algorithm is used to distribute the traffic load among the physical ports of the LAG while preserving the per-flow packet order.

The hashing algorithm uses various packet attributes to determine the outgoing physical port.

The switch supports the following set of packet attributes to be used for hash computation:

- Source MAC, VLAN, EtherType, and incoming port.
- Destination MAC, VLAN, EtherType, and incoming port.
- Source IP and Source TCP/UDP port numbers.
- Destination IP and Destination TCP/UDP port numbers.
- Source/Destination MAC, VLAN, EtherType, and incoming port.
- Source/Destination IP and Source/Destination TCP/UDP port numbers.
- Enhanced hashing mode

Enhanced hashing mode has following advantages:

- MODULO-N operation based on the number of ports in the LAG.
- Packet attribute selection based on the packet type. For L2 packets, Source and Destination MAC address are used for hash computation. For IP packets, Source IP, Destination IP address, TCP/UDP ports are used.
- Non-Unicast traffic and Unicast traffic is hashed using a common hash algorithm.
- Excellent load balancing performance.

LAG Interface Naming Convention

LAGs are logical interfaces and follow a slot/port naming convention. The slot number is always 3, and the port number ranges from 1 to the maximum number of LAGs the switch supports. The `show port-channel brief` command provides summary information about all LAGs available on the system. In the following output, LAG 3/1 has been configured as a dynamic LAG with five member ports. No other LAGs have been configured.

```
(DCSS Switching) #show port-channel brief

Logical   Port-Channel  Min Link State Trap     Type    Mbr Ports Active Ports
Interface Name                       Flag
--------- ------------- --- ---------- -------- ------- --------- ------------
3/1       ch1            1  Down       Disabled Dynamic 0/1,0/2,
                                                         0/3,0/6,
```

```
                                                          0/7
3/2       ch2             1  Down       Disabled Static
3/3       ch3             1  Down       Disabled Static
3/4       ch4             1  Down       Disabled Static
3/5       ch5             1  Down       Disabled Static
```

LAG Interaction with Other Features

From a system perspective, a LAG is treated just as a physical port, with the same configuration parameters for administrative enable/disable, spanning tree port priority, path cost as may used be for any other physical port.

VLAN

When members are added to a LAG, they are removed from all existing VLAN membership. When members are removed from a LAG they are added back to the VLANs that they were previously members of as per the configuration file. Note that a port's VLAN membership can still be configured when it's a member of a LAG. However this configuration is only actually applied when the port leaves the LAG.

The LAG interface can be a member of a VLAN complying with IEEE 802.1Q.

STP

Spanning tree does not maintain state for members of a LAG, but the Spanning Tree does maintain state for the LAG interface. As far as STP is concerned, members of a LAG do not exist. (Internally, the STP state of the LAG interface is replicated for the member links.)

When members are deleted from a LAG they become normal links, and spanning tree maintains their state information.

Statistics

Statistics are maintained for all LAG interfaces as they are done for the physical ports, besides statistics maintained for individual members as per the 802.3ad MIB statistics.

LAG Configuration Guidelines

Ports to be aggregated must be configured so that they are compatible with the link aggregation feature and with the partner switch to which they connect.

Ports to be added to a LAG must meet the following requirements:

- Interface must be a physical Ethernet link.
- Each member of the LAG must be running at the same speed and must be in full duplex mode.
- The port cannot be a mirrored port

The following are the interface restrictions

- The configured speed of a LAG member cannot be changed.
- An interface can be a member of only one LAG.

Link Aggregation Configuration Examples

This section contains the following examples:

- Configuring Dynamic LAGs
- Configuring Static LAGs

Note: The examples in this section show the configuration of only one switch. Because LAGs involve physical links between two switches, the LAG settings and member ports must be configured on both switches.

Configuring Dynamic LAGs

The commands in this example show how to configure a dynamic LAG on a switch. The LAG number is 1 (interface 3/1), and the member ports are 1, 2, 3, 6, and 7.

To configure the switch:

1. Enter interface configuration mode for the ports that are to be configured as LAG members.

```
(DCSS Switching) #config
(DCSS Switching) (Config)#interface 0/1-0/3,0/6-0/7
```

2. Add the ports to LAG 1 with LACP.

```
(DCSS Switching) (Interface 0/1-0/3,0/6-0/7)#addport 3/1
(DCSS Switching) (Interface 0/1-0/3,0/6-0/7)#exit
```

3. Configure LAG 1 as dynamic.

```
(DCSS Switching) (Config)#interface 3/1
(DCSS Switching) (Interface 3/1)#no port-channel static
(DCSS Switching) (Interface 3/1)#end
```

4. View information about LAG 1.

```
(DCSS Switching) #show port-channel 3/1

Local Interface.............................. 3/1
Channel Name................................. ch1
Link State................................... Down
Admin Mode................................... Enabled
Type......................................... Dynamic
Port-channel Min-links....................... 1
Load Balance Option.......................... 3
(Src/Dest MAC, VLAN, EType, incoming port)

Mbr    Device/       Port      Port
Ports  Timeout       Speed     Active
------ ------------- --------- -------
0/1    actor/long    Auto      False
       partner/long
0/2    actor/long    Auto      False
       partner/long
0/3    actor/long    Auto      False
       partner/long
0/6    actor/long    Auto      False
       partner/long
0/7    actor/long    Auto      False
       partner/long
```

Configuring Static LAGs

The commands in this example show how to configure a static LAG on a switch. The LAG number is 3 (interface 3/3), and the member ports are 10, 11, 14, and 17.

To configure the switch:

1. Enter interface configuration mode for the ports that are to be configured as LAG members.

```
(DCSS Switching) (Config)#interface 0/10-0/12,0/14,0/17
```

2. Add the ports to LAG 2 without LACP.

```
(DCSS Switching) (Interface 0/10-0/12,0/14,0/17)#addport 3/3
(DCSS Switching) (Interface 0/10-0/12,0/14,0/17)#end
```

3. View information about LAG 2.

```
(DCSS Switching) #show port-channel 3/3

Local Interface.............................. 3/3
Channel Name................................. ch3
Link State................................... Up
Admin Mode................................... Enabled
Type......................................... Static
Port-channel Min-links....................... 1
Load Balance Option.......................... 3
(Src/Dest MAC, VLAN, EType, incoming port)

Mbr    Device/       Port      Port
Ports  Timeout       Speed     Active
------ ------------- --------- -------
0/10   actor/long    Auto      True
       partner/long
0/11   actor/long    Auto      False
       partner/long
0/12   actor/long    Auto      False
       partner/long
0/14   actor/long    Auto      False
       partner/long
0/17   actor/long    Auto      False
       partner/long
--More-- or (q)uit
```

Unidirectional Link Detection (UDLD)

The UDLD feature detects unidirectional links on physical ports. UDLD must be enabled on the both sides of the link in order to detect a unidirectional link. The UDLD protocol operates by exchanging packets containing information about neighboring devices.

The purpose of UDLD feature is to detect and avoid unidirectional links. A unidirectional link is a forwarding anomaly in a Layer 2 communication channel in which a bidirectional link stops passing traffic in one direction.

UDLD Modes

The UDLD supports two modes: normal and aggressive.

In normal mode, a port's state is classified as *undetermined* if an anomaly exists. An anomaly might be the absence of its own information in received UDLD messages or the failure to receive UDLD messages. An *undetermined* state has no effect on the operation of the port. The port is not disabled and continues operating. When operating in UDLD normal mode, a port will be put into a disabled state (D-Disable) only in the following situations:

- The UDLD PDU received from a partner does not have its own details (echo).
- When there is a loopback, and information sent out on a port is received back exactly as it was sent.

When operating in UDLD aggressive mode, a port is put into a disabled state for the same reasons that it occurs in normal mode. Additionally, a port in UDLD aggressive mode can be disabled if the port does not receive any UDLD echo packets even *after* bidirectional connection was established. If a bidirectional link is established, and packets suddenly stop coming from partner device, the UDLD aggressive-mode port assumes that link has become unidirectional.

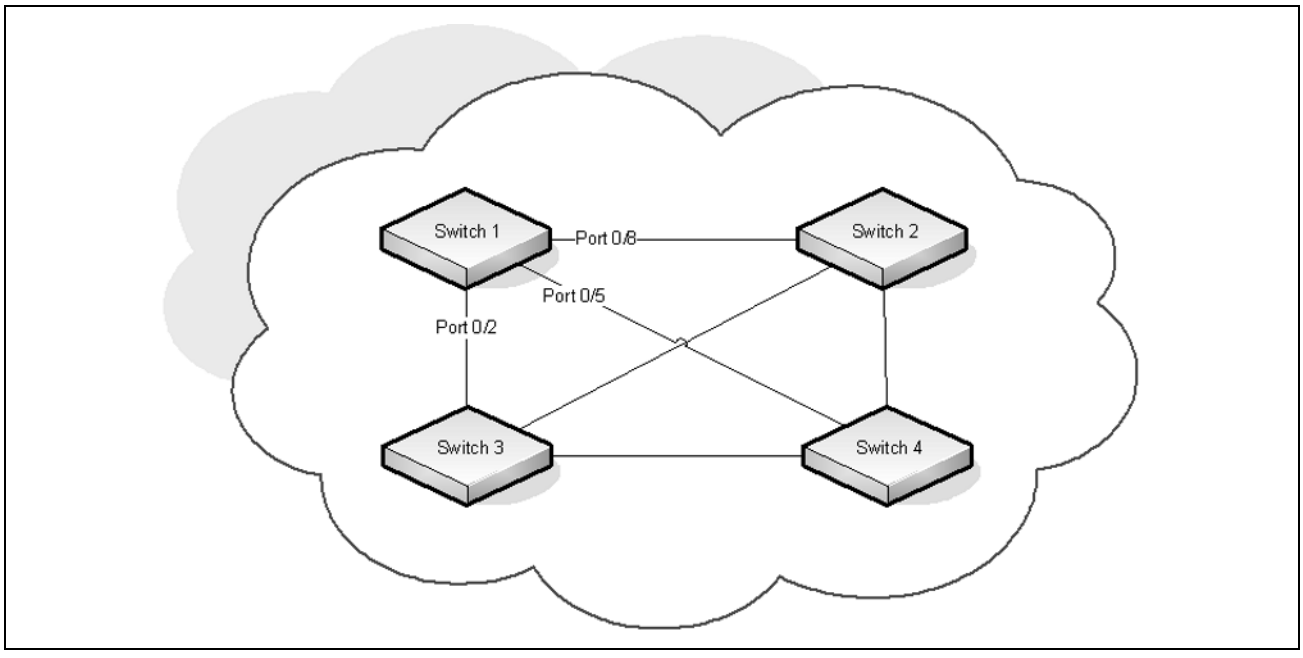
UDLD and LAG Interfaces

UDLD is supported on individual physical ports that are members of port channel interface. If any of the aggregated links becomes unidirectional, UDLD detects it and disables the individual link, but not the entire port channel. This improves fault tolerance of port-channel.

Configuring UDLD

A network administrator decides to use the UDLD feature while building a loop-free topology with the use of STP. The administrator configures the ports on both sides of the link to use UDLD in aggressive mode to ensure that ports with unidirectional links will be shut down, and no loops will be introduced into the topology. This example shows the steps to configure UDLD on Switch 1 only. The same configuration must be performed on all ports that form partner links with the ports on Switch 1.

Figure 8: UDLD Configuration Example



To configure the ports on Switch 1:

1. Globally enable UDLD on the switch.

```
(DCSS Switching) #configure
(DCSS Switching) (Config)#udld enable
```

2. Enter interface configuration mode for the ports that are connected to other switches and enable UDLD on the ports.

```
(DCSS Switching) (Config)#interface 0/8,0/11,0/20
(DCSS Switching) (Interface 0/8,0/11,0/20)#udld enable
```

3. Configure the UDLD mode on the ports to be aggressive.

```
(DCSS Switching) (Interface 0/8,0/11,0/20)#udld port aggressive
(DCSS Switching) (Interface 0/8,0/11,0/20)#end
```

4. After configuring UDLD on Switch 2, Switch, 3, and Switch 4, view the UDLD status for the ports.

```
(DCSS Switching) #show udld all

Port       Admin Mode   UDLD Mode      UDLD Status
-----      ----------   -----------    --------------
0/1        Disabled     Normal         Not Applicable
0/8        Enabled      Aggressive     Bidirectional
0/3        Disabled     Normal         Not Applicable
0/4        Disabled     Normal         Not Applicable
0/8        Enabled      Aggressive     Bidirectional
0/6        Disabled     Normal         Not Applicable
0/7        Disabled     Normal         Not Applicable
0/8        Enabled      Aggressive     Bidirectional
0/9        Disabled     Normal         Not Applicable
--More-- or (q)uit
```

Note: If a port has become disabled by the UDLD feature and you want to re-enable the port, use the `udld reset` command in Privileged EXEC mode.

Port Mirroring

Port mirroring is used to monitor the network traffic that a port sends and receives. The Port Mirroring feature creates a copy of the traffic that the source port handles and sends it to a destination port. The source port is the port that is being monitored. The destination port is monitoring the source port. The destination port is where you would connect a network protocol analyzer to learn more about the traffic that is handled by the source port.

A port monitoring session includes one or more source ports that mirror traffic to a single destination port. DCSS software supports a single port monitoring session. LAGs (port channels) cannot be used as the source or destination ports.

For each source port, you can specify whether to mirror ingress traffic (traffic the port receives, or RX), egress traffic (traffic the port sends, or TX), or both ingress and egress traffic.

The packet that is copied to the destination port is in the same format as the original packet on the wire. This means that if the mirror is copying a received packet, the copied packet is VLAN tagged or untagged as it was received on the source port. If the mirror is copying a transmitted packet, the copied packet is VLAN tagged or untagged as it is being transmitted on the source port.

After you configure the port mirroring session, you can enable or disable the administrative mode of the session to start or stop the probe port from receiving mirrored traffic.

Configuring Port Mirroring

In this example, traffic from ports 1 and 4 is mirrored to probe port 10.

1. Configure the source ports. Traffic received and transmitted on by these ports will be mirrored.

```
(DCSS Switching) #configure
(DCSS Switching) (Config)#monitor session 1 source interface 0/1
(DCSS Switching) (Config)#monitor session 1 source interface 0/4
```

2. Configure the destination (probe) port.

```
(DCSS Switching) (Config)#monitor session 1 destination interface 0/10
```

3. Enable port mirroring on the switch.

```
(DCSS Switching) (Config)#monitor session 1 mode
(DCSS Switching) (Config)#exit
```

4. View summary information about the port mirroring configuration.

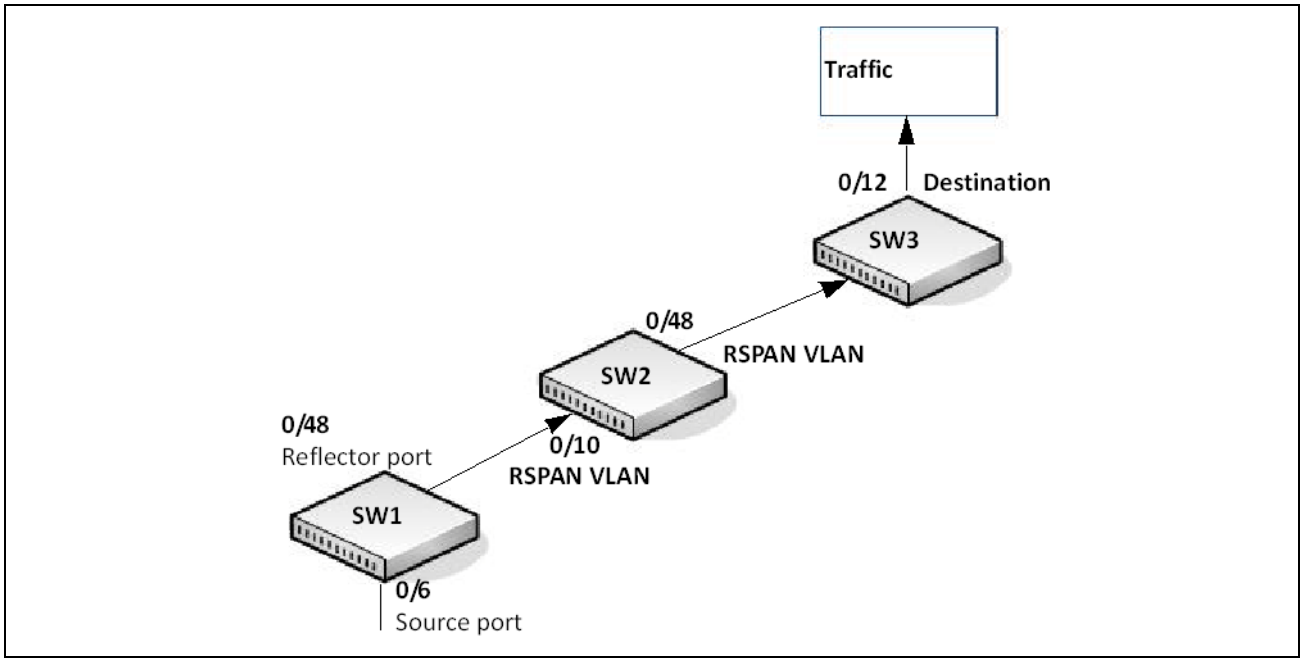
```
(DCSS Switching) #show monitor session 1

Session Admin   Probe   Src  Mirrored Ref.    Src   Dst   Type  IP      MAC
ID      Mode    Port    VLAN Port     Port    RVLAN RVLAN       ACL     ACL
------- ------- ------- ---- -------- ------- ----- ----- ----  ------- -------
1       Enable  0/10         0/1                          Rx,Tx
                             0/4                          Rx,Tx
```

Configuring RSPAN

This example mirrors traffic from port 6 on a source switch (SW1) to a probe port on a remote switch (port 12 on SW3). The mirrored traffic is carried in the RSPAN VLAN and VLAN 100, which traverses an intermediate switch (SW2). The commands in this example show how to configure port mirroring on the source, intermediate, and destination switches.

Figure 9 provides a visual overview of the RSPAN configuration example.

Figure 9: RSPAN Configuration Example



Configuration on the Source Switch (SW1)

To configure the source switch:

1. Access the VLAN configuration mode and create VLAN 100, which will be the RSPAN VLAN.

```
(DCSS Switching) #vlan database
(DCSS Switching) (Vlan) #vlan 100
(DCSS Switching) (Vlan) #exit
```

2. Configure VLAN 100 as the RSPAN VLAN.

```
(DCSS Switching) #configure
(DCSS Switching) (Config) #vlan 100
(DCSS Switching) (Config) (vlan 100) #remote-span
(DCSS Switching) (Config) (vlan 100) #exit
```

3. Configure the RSPAN VLAN as the destination port and the reflector port as port 0/48.

```
(DCSS Switching) #configure
(DCSS Switching) (Config) #monitor session 1 destination remote vlan 100 reflector-port 0/48
```

4. Configure the source interface port as port 0/6.

```
(DCSS Switching) (Config) #monitor session 1 source interface 0/6
```

5. Enable the port mirroring session on the switch.

```
(DCSS Switching) (Config) #monitor session 1 mode
(DCSS Switching) #exit
```

Configuration on the Intermediate Switch (SW2)

To configure the intermediate switch:

1. Access the VLAN configuration mode and create VLAN 100.

```
(DCSS Switching) #vlan database
(DCSS Switching) (Vlan) #vlan 100
(DCSS Switching) (Vlan) #exit
```

2. Enable RSPAN on vlan 100.

```
(DCSS Switching) #configure
(DCSS Switching) (Config) #vlan 100
(DCSS Switching) (Config) (vlan 100) #remote-span
(DCSS Switching) (Config) (vlan 100) #exit
```

3. Configure VLAN participation so the interface is always a member of the VLAN.

```
(DCSS Switching) (Config) #vlan participation include 100
(DCSS Switching) (Config) #interface 0/10
```

4. Enable VLAN tagging on the interface.

```
(DCSS Switching) (Config) #vlan tagging 100
(DCSS Switching) (Config) #exit
```

5. Configure VLAN participation so the interface is always a member of the VLAN.

```
(DCSS Switching) (Config) #vlan participation include 100
(DCSS Switching) (Config) #interface 0/48
(DCSS Switching) (Config) #exit
```

Configuration on the Destination Switch (SW3)

To configure the destination switch:

1. Access the VLAN configuration mode and create VLAN 100.

```
(DCSS Switching) #vlan database
(DCSS Switching) (Vlan) #vlan 100
(DCSS Switching) (Vlan) #exit
```

2. Enable RSPAN on vlan 100.

```
(DCSS Switching) #configure
(DCSS Switching) (Config) #vlan 100
(DCSS Switching) (Config) (vlan 100) #remote-span
(DCSS Switching) (Config) (vlan 100) #exit
```

3. Configure the RSPAN VLAN as the source interface for the port mirroring session.

```
(DCSS Switching) #configure
(DCSS Switching) (Config) #monitor session 1 source remote vlan 100
```

4. Configure the destination port as port 0/12. This is the probe port that is attached to a network traffic analyzer.

```
(DCSS Switching) (Config) #monitor session 1 destination interface 0/12
```

5. Enable the port mirroring session on the switch.

```
(DCSS Switching) (Config) #monitor session 1 mode
(DCSS Switching) (Config) #exit
```

Configuring VLAN Based Mirroring

In this example, traffic from all ports that are members of VLAN 10 is mirrored to port 0/18.

To configure VLAN based mirroring:

1. Access VLAN Config mode and create VLAN 10.

```
(DCSS Switching) #vlan database
(DCSS Switching) (Vlan) #vlan 10
(DCSS Switching) (Vlan) #exit
```

2. Configure the destination interface port as port 0/18.

```
(DCSS Switching) #configure
(DCSS Switching) (Config) #monitor session 1 destination interface 0/18
```

3. Configure VLAN 10 as the source interface for the port mirroring session.

```
(DCSS Switching) (Config) #monitor session 1 source vlan 10
```

4. Enable the port mirroring session on the switch.

```
(DCSS Switching) (Config) #monitor session 1 mode
(DCSS Switching) (Config) #exit
```

Configuring Flow Based Mirroring

In this example, traffic from port 1 is mirrored to port 18 if it matches the criteria defined in the IP ACL or MAC ACL that are associated with the port mirroring session.

To configure flow based mirroring:

1. Create the extended IP access list IPACL.

```
(DCSS Switching) #configure
(DCSS Switching) (Config) #ip access-list IPACL
(DCSS Switching) (Config) #permit ip 1.1.1.1 0.0.0.0 any
(DCSS Switching) (Config) #exit
```

2. Create the mac access list MACL.

```
(DCSS Switching) #configure
(DCSS Switching) (Config) #mac access-list extended MACL
(DCSS Switching) (Config) #permit 00:00:00:00:00:11 00:00:00:00:00:00 any
(DCSS Switching) (Config) #exit
```

3. Configure the destination port as port 0/18.

```
(DCSS Switching) #monitor session 1 destination interface 0/18
```

4. Configure the source port as port 0/2.

```
(DCSS Switching) #monitor session 1 source interface 0/2
```

5. Enable the port mirroring session.

```
(DCSS Switching) #monitor session 1 mode
```

6. To filter L3 traffic so only flows that match the rules in the IP ACL called IPACL are mirrored to the destination port, add the IPACL ACL.

```
(DCSS Switching) #monitor session 1 filter ip access-group IPACL
```

7. To filter L2 traffic so only flows that match the rules in the MAC-based ACL called MACL are mirrored to the destination port, add the MACL ACL.

```
(DCSS Switching) #monitor session 1 filter mac access-group MACL
(DCSS Switching) #exit
```

Spanning Tree Protocol

Spanning Tree Protocol (STP) is a layer 2 protocol that provides a tree topology for switches on a bridged LAN. STP allows a network to have redundant paths without the risk of network loops. STP uses the spanning-tree algorithm to provide a single path between end stations on a network.

DCSS software supports Classic STP, Multiple STP, and Rapid STP.

Classic STP, Multiple STP, and Rapid STP

Classic STP provides a single path between end stations, avoiding and eliminating loops. Multiple Spanning Tree Protocol (MSTP) is specified in IEEE 802.1s and supports multiple instances of Spanning Tree to efficiently channel VLAN traffic over different interfaces. Each instance of the Multiple Spanning Tree behaves in the manner specified in IEEE 802.1w, Rapid Spanning Tree (RSTP), with slight modifications in the working but not the end effect (chief among the effects, is the rapid transitioning of the port to Forwarding). The difference between the RSTP and the traditional STP (IEEE 802.1d) is the ability to configure and recognize full-duplex connectivity and ports which are connected to end stations, resulting in rapid transitioning of the port to the Forwarding state and the suppression of Topology Change Notifications.

MSTP is compatible to both RSTP and STP. It behaves appropriately to STP and RSTP bridges. A MSTP bridge can be configured to behave entirely as a RSTP bridge or a STP bridge.

STP Operation

The switches (bridges) that participate in the spanning tree elect a switch to be the root bridge for the spanning tree. The root bridge is the switch with the lowest bridge ID, which is computed from the unique identifier of the bridge and its configurable priority number. When two switches have an equal bridge ID value, the switch with the lowest MAC address is the root bridge.

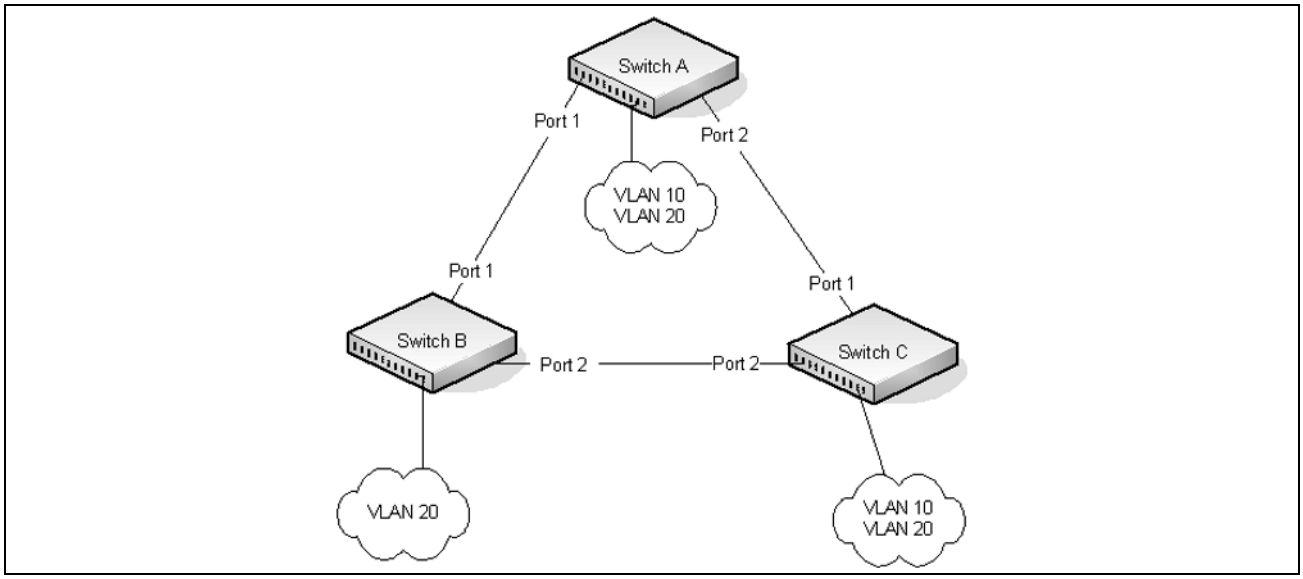
After the root bridge is elected, each switch finds the lowest-cost path to the root bridge. The port that connects the switch to the lowest-cost path is the root port on the switch. The switches in the spanning tree also determine which ports have the lowest-path cost for each segment. These ports are the designated ports. Only the root ports and designated ports are placed in a forwarding state to send and receive traffic. All other ports are put into a blocked state to prevent redundant paths that might cause loops.

To determine the root path costs and maintain topology information, switches that participate in the spanning tree use Bridge Protocol Data Units (BPDUs) to exchange information.

MSTP in the Network

In the following diagram of a small 802.1d bridged network, STP is necessary to create an environment with full connectivity and without loops.

Figure 10: STP in a Small Bridged Network



Assume that Switch A is elected to be the Root Bridge, and Port 1 on Switch B and Switch C are calculated to be the root ports for those bridges, Port 2 on Switch B and Switch C would be placed into the Blocking state. This creates a loop-free topology. End stations in VLAN 10 can talk to other devices in VLAN 10, and end stations in VLAN 20 have a single path to communicate with other VLAN 20 devices.

Figure 11 shows the logical single STP network topology.

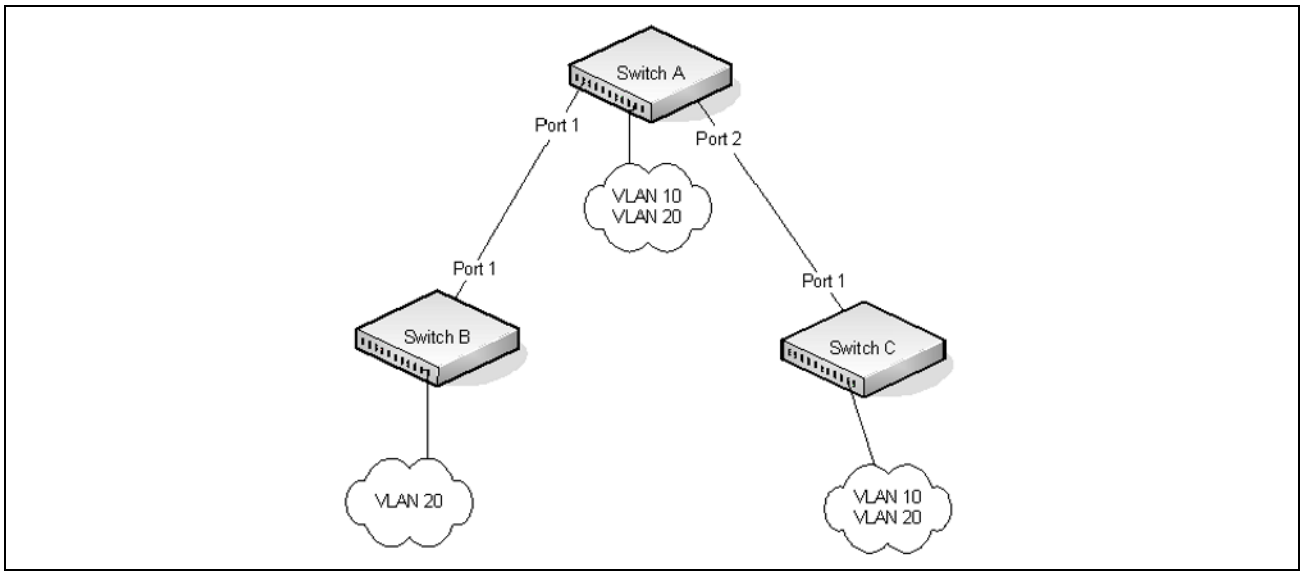
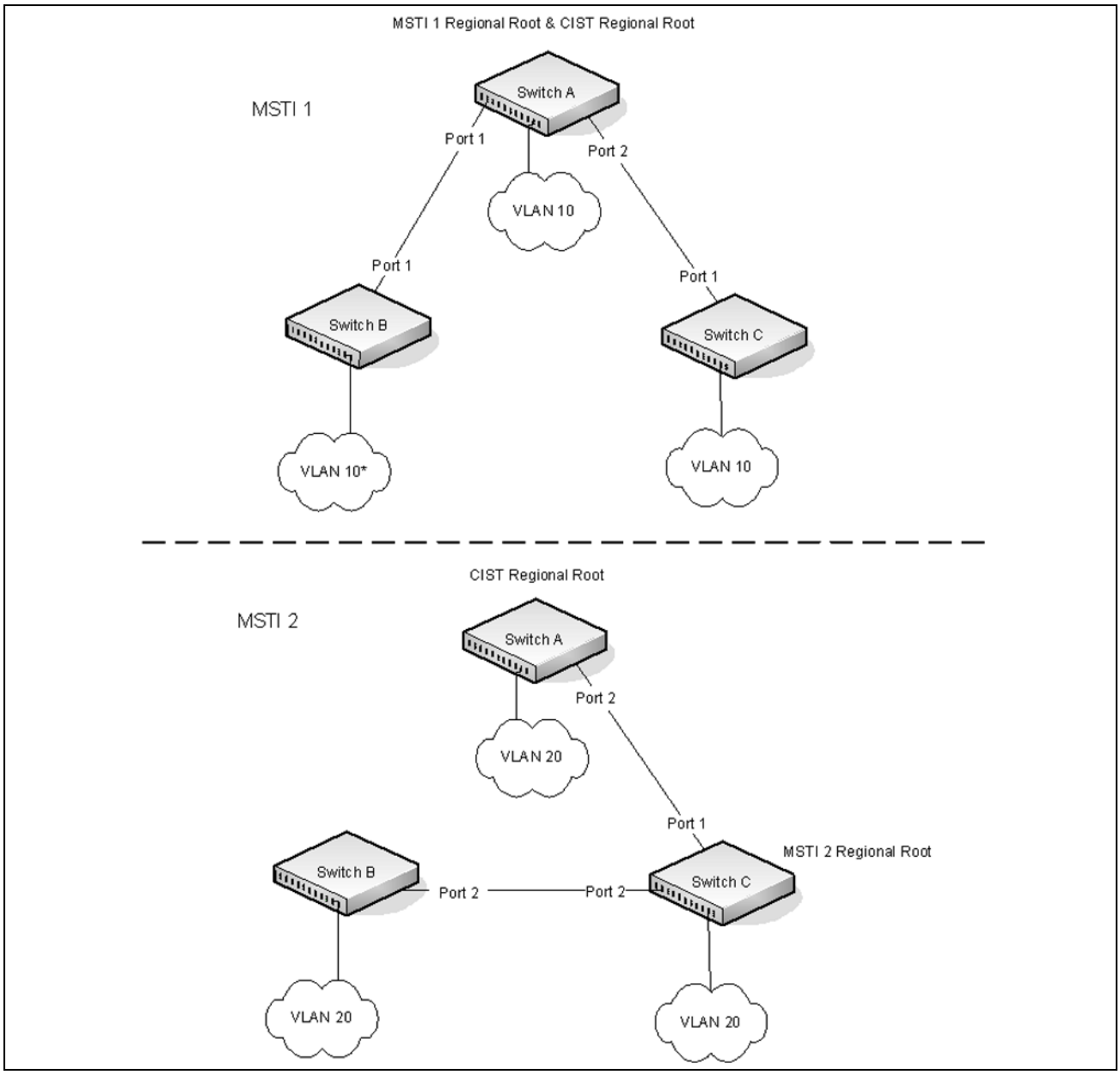


Figure 11: Single STP Topology

For VLAN 10 this single STP topology is fine and presents no limitations or inefficiencies. On the other hand, VLAN 20's traffic pattern is inefficient. All frames from Switch B will have to traverse a path through Switch A before arriving at Switch C. If the Port 2 on Switch B and Switch C could be used, these inefficiencies could be eliminated. MSTP does just that, by allowing the configuration of MSTIs based upon a VLAN or groups of VLANs. In this simple case, VLAN 10 could be associated with Multiple Spanning Tree Instance (MSTI)1 with an active topology similar to Figure 12 and VLAN 20 could be associated with MSTI 2 where Port 1 on both Switch A and Switch B begin discarding and all others forwarding. This simple modification creates an active topology with a better distribution of network traffic and an increase in available bandwidth.

The logical representation of the MSTP environment for these three switches is shown in Figure 12.

Figure 12: Logical MSTP Environment



For MSTP to correctly establish the different MSTIs as above, some additional changes are required. For example, the configuration would have to be the same on each and every bridge. That means that Switch B would have to add VLAN 10 to its list of supported VLANs (shown in Figure 12 on page 90 with a *). This is necessary with MSTP to allow the formation of Regions made up of all switches that exchange the same MST Configuration Identifier. It is within only these MST Regions that multiple instances can exist. It will also allow the election of Regional Root Bridges for each instance. One common and internal spanning tree (CIST) Regional Root for the CIST and an MSTI Regional Root Bridge per instance will enable the possibility of alternate paths through each Region. Above Switch A is elected as both the MSTI 1 Regional Root and the CIST Regional Root Bridge, and after adjusting the Bridge Priority on Switch C in MSTI 2, it would be elected as the MSTI 2 Regional Root.

To further illustrate the full connectivity in an MSTP active topology, the following rules apply:

1. Each Bridge or LAN is in only one Region.
2. Every frame is associated with only one VID.
3. Frames are allocated either to the IST or MSTI within any given Region.
4. The internal spanning tree (IST) and each MSTI provides full and simple connectivity between all LANs and Bridges in a Region.
5. All Bridges within a Region reach a consistent agreement as to which ports interconnect that Region to a different Region and label those as Boundary Ports.
6. At the Boundary Ports, frames allocated to the CIST or MSTIs are forwarded or not forwarded alike.
7. The CIST provides full and simple connectivity between all LANs and Bridges in the network.

Optional STP Features

DCSS software supports the following optional STP features:

- BPDU flooding
- Edge Port
- BPDU filtering
- Root guard
- Loop guard
- BPDU protection

BPDU Flooding

The BPDU flooding feature determines the behavior of the switch when it receives a BPDU on a port that is disabled for spanning tree. If BPDU flooding is configured, the switch will flood the received BPDU to all the ports on the switch which are similarly disabled for spanning tree.

Edge Port

The Edge Port feature reduces the STP convergence time by allowing ports that are connected to end devices (such as a desktop computer, printer, or file server) to transition to the forwarding state without going through the listening and learning states.

BPDU Filtering

Ports that have the Edge Port feature enabled continue to transmit BPDUs. The BPDU filtering feature prevents ports configured as edge ports from sending BPDUs.

If BPDU filtering is configured globally on the switch, the feature is automatically enabled on all operational ports where the Edge Port feature is enabled. These ports are typically connected to hosts that drop BPDUs. However, if an operational edge port receives a BPDU, the BPDU filtering feature disables the Edge Port feature and allows the port to participate in the spanning-tree calculation.

Enabling BPDU filtering on a specific port prevents the port from sending BPDUs and allows the port to drop any BPDUs it receives.

Root Guard

Enabling root guard on a port ensures that the port does not become a root port or a blocked port. When a switch is elected as the root bridge, all ports are designated ports unless two or more ports of the root bridge are connected together. If the switch receives superior STP BPDUs on a root-guard enabled port, the root guard feature moves this port to a root-inconsistent STP state, which is effectively equal to a listening state. No traffic is forwarded across this port. In this way, the root guard feature enforces the position of the root bridge.

When the STP mode is MSTP, the port may be a designated port in one MSTI and an alternate port in the CIST, etc. Root guard is a per port (not a per port per instance command) configuration, so all the MSTP instances this port participates in should not be in a root role.

Loop Guard

Loop guard protects a network from forwarding loops induced by BPDU packet loss. The reasons for failing to receive packets are numerous, including heavy traffic, software problems, incorrect configuration, and unidirectional link failure. When a non-designated port no longer receives BPDUs, the spanning-tree algorithm considers that this link is loop free and begins transitioning the link from blocking to forwarding. Once in forwarding state, the link may create a loop in the network.

Enabling loop guard prevents such accidental loops. When a port is no longer receiving BPDUs and the max age timer expires, the port is moved to a *loop-inconsistent blocking state*. In the loop-inconsistent blocking state, traffic is not forwarded so the port behaves as if it is in the blocking state. The port will remain in this state until it receives a BPDU. It will then transition through the normal spanning tree states based on the information in the received BPDU.

Note: Loop Guard should be configured only on non-designated ports. These include ports in alternate or backup roles. Root ports and designated ports should not have loop guard enabled so that they can forward traffic

BPDU Protection

When the switch is used as an access layer device, most ports function as edge ports that connect to a device such as a desktop computer or file server. The port has a single, direct connection and is configured as an edge port to implement the fast transition to a forwarding state. When the port receives a BPDU packet, the system sets it to non-edge port and recalculates the spanning tree, which causes network topology flapping. In normal cases, these ports do not receive any BPDU packets. However, someone may forge a BPDU to maliciously attack the switch and cause network flapping.

BPDU protection can be enabled in RSTP to prevent such attacks. When BPDU protection is enabled, the switch disables an edge port that has received a BPDU and notifies the network manager about it.

STP Configuration Examples

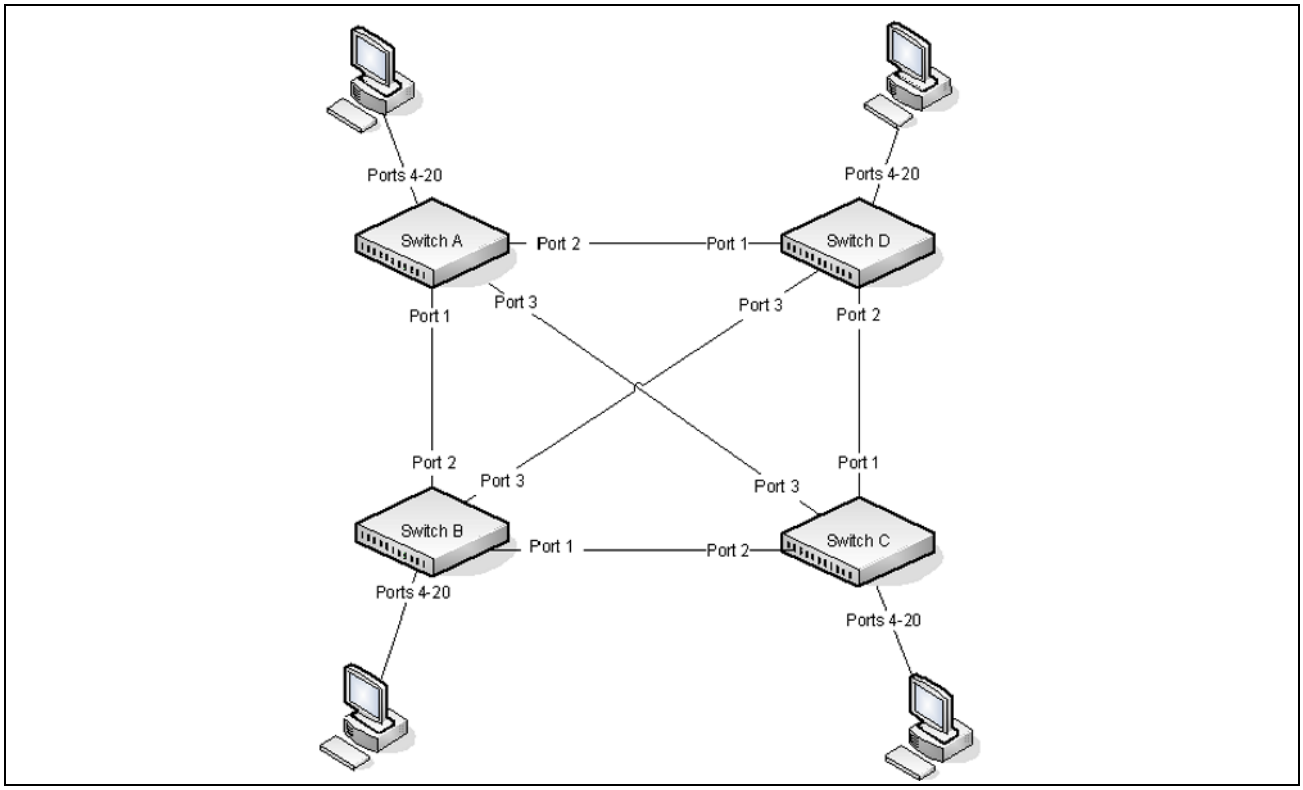
This section contains the following examples:

- Configuring STP
- Configuring MSTP

Configuring STP

This example shows a LAN with four switches. On each switch, ports 1, 2, and 3 connect to other switches, and ports 4–20 connect to hosts (in Figure 13, each PC represents 17 host systems).

Figure 13: STP Example Network Diagram



Of the four switches in Figure 13 on page 93, the administrator decides that Switch A is the most centrally located in the network and is the least likely to be moved or redeployed. For these reasons, the administrator selects it as the root bridge for the spanning tree. The administrator configures Switch A with the highest priority and uses the default priority values for Switch B, Switch C, and Switch D.

For all switches, the administrator also configures ports 4–17 in Port Fast mode because these ports are connected to hosts and can transition directly to the Forwarding state to speed up the connection time between the hosts and the network.

The administrator also configures Port Fast BPDU filtering and Loop Guard to extend STP's capability to prevent network loops. For all other STP settings, the administrator uses the default STP values.

To configure the switch:

1. Connect to Switch A and configure the priority to be higher (a lower value) than the other switches, which use the default value of 32768.

```
(DCSS Switching)#config
(DCSS Switching) (Config)#spanning-tree mst priority 0 8192
```

2. Configure ports 4–20 to be in Edge Port mode.

```
(DCSS Switching) (Config)#interface 0/4-0/20
(DCSS Switching)(Interface 0/4-0/20)#spanning-tree edgeport
(DCSS Switching)(Interface 0/4-0/20)#exit
```

3. Enable Loop Guard on ports 1–3 to help prevent network loops that might be caused if a port quits receiving BPDUs.

```
(DCSS Switching) (Config)#interface 0/1-0/3
(DCSS Switching)(Interface 0/1-0/3)#spanning-tree guard loop
(DCSS Switching)(Interface 0/1-0/3)#exit
```

4. Enable Port Fast BPDU Filter. This feature is configured globally, but it affects only access ports that have the Edge Port feature enabled.

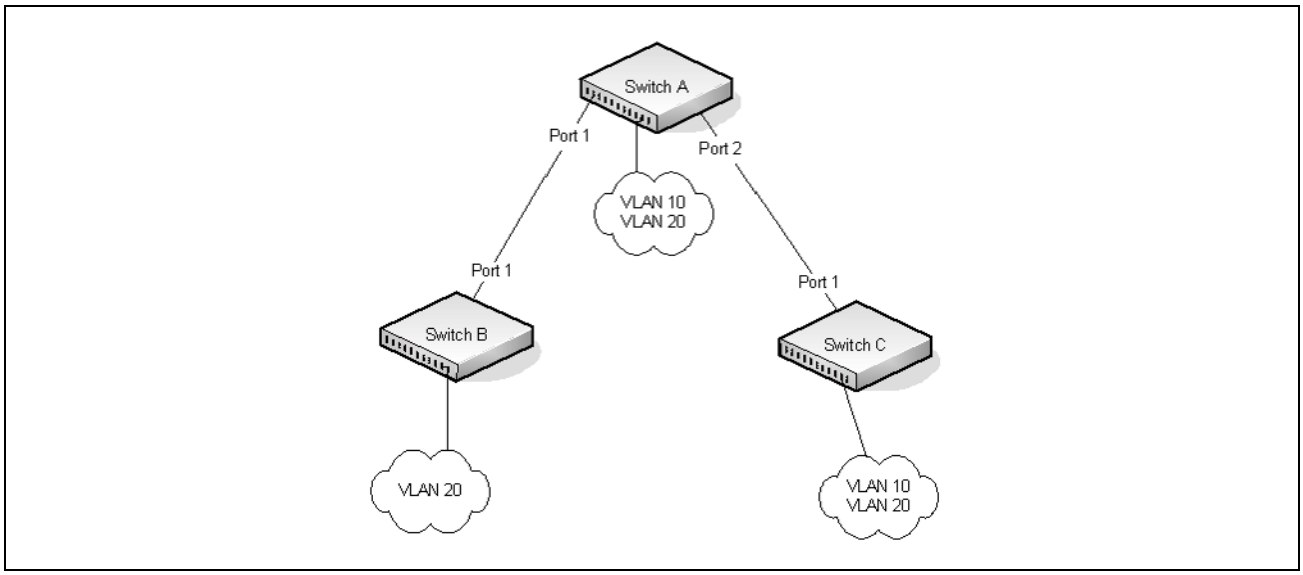
```
(DCSS Switching) (Config)#spanning-tree bpdufilter default
```

5. Repeat Step 2 through Step 4 on Switch B, Switch C, and Switch D to complete the configuration.

Configuring MSTP

This example shows how to configure IEEE 802.1s Multiple Spanning Tree (MST) protocol on the switches shown in Figure 14.

Figure 14: MSTP Configuration Example



To make multiple switches be part of the same MSTP region, make sure the STP operational mode for all switches is MSTP. Also, make sure the MST region name and revision level are the same for all switches in the region.

To configure the switches:

1. Create VLAN 10 (Switch A and Switch B) and VLAN 20 (all switches).

Note: Even Switch B does not have any ports that are members of VLAN 10, this VLAN must be created to allow the formation of MST regions made up of all bridges that exchange the same MST Configuration Identifier. It is only within these MST Regions that multiple instances can exist.

```
(DCSS Switching)#vlan database
(DCSS Switching)(Vlan)#vlan 10,20
(DCSS Switching)(Vlan)#exit
```

2. Set the STP operational mode to MSTP.

```
(DCSS Switching)#config
(DCSS Switching) (Config)#spanning-tree forceversion 802.1s
```

3. Create MST instance 10 and associate it to VLAN 10.

```
(DCSS Switching) (Config)#spanning-tree mst instance 10
(DCSS Switching) (Config)#spanning-tree mst vlan 10 10
```

4. Create MST instance 20 and associate it to VLAN 20.

```
(DCSS Switching) (Config)#spanning-tree mst instance 20
(DCSS Switching) (Config)#spanning-tree mst vlan 20 20
```

5. Change the region name so that all the bridges that want to be part of the same region can form the region.

```
(DCSS Switching) (Config)#spanning-tree configuration name dcss
```

6. (Switch A only) Make Switch A the Regional Root for MSTI 1 by configuring a higher priority for MST ID 10.

```
(DCSS Switching) (Config)#spanning-tree mst priority 10 12288
```

7. (Switch A only) Change the priority of MST ID 20 to ensure Switch C is the Regional Root bridge for this MSTI.

```
(DCSS Switching) (Config)#spanning-tree mst priority 20 61440
```

8. (Switch C only) Change the priority of port 1 to force it to be the root port for MST 20.

```
(DCSS Switching) (Config)#interface 0/1
(DCSS Switching)(Interface 0/1)#spanning-tree mst 20 port-priority 64
(DCSS Switching)(Interface 0/1)#end
```

IGMP Snooping

IGMP Snooping is a layer 2 feature that allows the switch to dynamically add or remove ports from IP multicast groups by listening to IGMP join and leave requests. By “snooping” the IGMP packets transmitted between hosts and routers, the IGMP Snooping feature enables the switch to forward IP multicast traffic more intelligently and help conserve bandwidth.

Based on the IGMP query and report messages, the switch forwards traffic only to the ports that request the multicast traffic. This prevents the switch from broadcasting the traffic to all ports and possibly affecting network performance. The switch uses the information in the IGMP packets as they are being forwarded throughout the network to determine which segments should receive packets directed to the group address.

IGMP Snooping Querier

When PIM and IGMP are enabled in a network with IP multicast routing, the IP multicast router acts as the IGMP querier. However, if the IP-multicast traffic in a VLAN needs to be Layer 2 switched only, an IP-multicast router is not required. The IGMP Snooping Querier can perform the IGMP snooping functions on the VLAN.

Without an IP-multicast router on a VLAN, you must configure another switch as the IGMP querier so that it can send queries.

When the IGMP snooping querier is enabled, the IGMP snooping querier sends out periodic IGMP queries that trigger IGMP report messages from the switch that wants to receive IP multicast traffic. The IGMP snooping feature listens to these IGMP reports to establish appropriate forwarding.

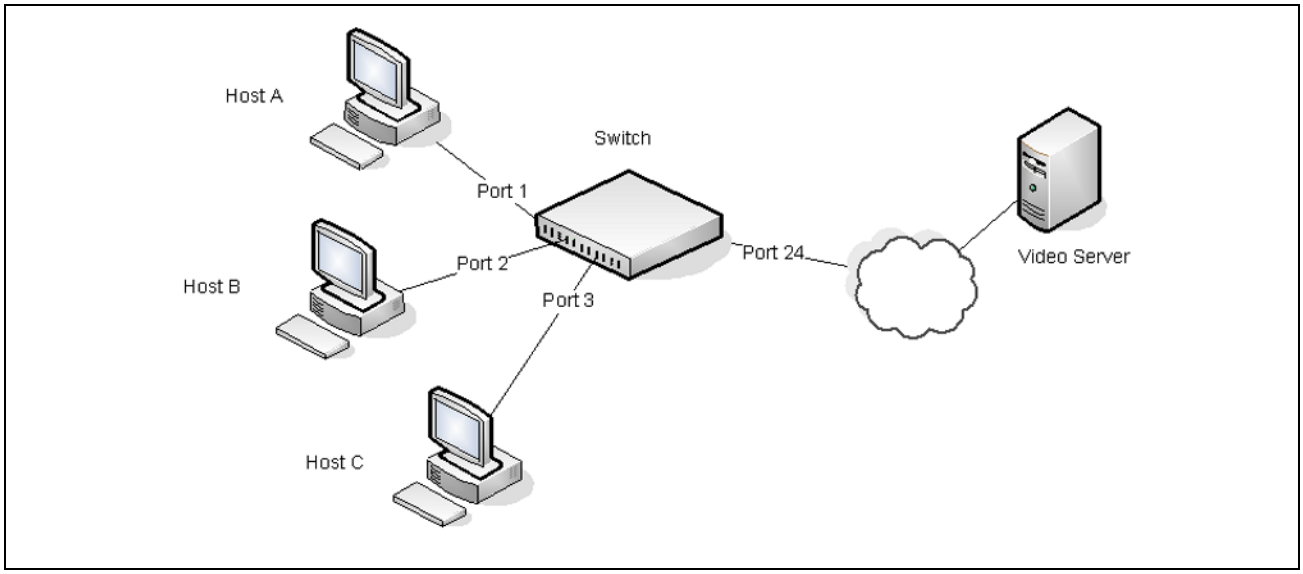
Configuring IGMP Snooping

This example configures IGMP snooping on the switch to limit multicast traffic and to allow L2 multicast forwarding on a single VLAN. The IP-multicast traffic in VLAN 100 needs to be Layer 2 switched only, so the IGMP snooping querier is enabled on the switch to perform the IGMP snooping functions on the VLAN, if necessary. The switch can send queries even if it is not the IGMP snooping querier and will use 0.0.0.0 as the source IP address. This will not cause any disruption to the operation of external querier.

In this configuration, an IP-multicast router is not required.

The three hosts in Figure 15 are connected to ports that are enabled for IGMP snooping and are members of VLAN 100. Port 24 is a trunk port and connects the switch to the data center, where the L3 multicast router is located.

Figure 15: Switch with IGMP Snooping



To configure the switch:

1. Enable IGMP snooping globally.

```
(DCSS Switching) #configure
(DCSS Switching) (Config)#set igmp
```

2. Enable the IGMP snooping querier on the switch. If there are no other IGMP snooping queriers, this switch will become the IGMP snooping querier for the local network. If an external querier is discovered, this switch will not be a querier.

```
(DCSS Switching) (Config)#set igmp querier
(DCSS Switching) (Config)#exit
```

3. Create VLAN 100

```
(DCSS Switching)#vlan database
(DCSS Switching) (Vlan)#vlan 100
```

4. Enable IGMP snooping on VLAN 100.

```
(DCSS Switching) (Vlan)#set igmp 100
```

5. Enable the IGMP snooping querier on VLAN 100.

```
(DCSS Switching) (Vlan)#set igmp querier 100
```

6. Enable VLAN routing on VLAN 100.

```
(DCSS Switching) (Vlan)#vlan routing 100
(DCSS Switching) (Vlan)#exit
```

7. View the VLAN routing interface information.

```
(DCSS Switching) #show ip interface brief

Interface    State  IP Address      IP Mask         Method
----------   -----  --------------- --------------- -------
4/1          Down   0.0.0.0         0.0.0.0         None
```

8. Configure an IP address for VLAN 100. This address will be used as the IGMP snooping querier address if this switch becomes the querier.

```
(DCSS Switching)#configure
(DCSS Switching) (Config)#interface 4/1
(DCSS Switching) (Interface 0/1)#ip address 192.168.10.2 255.255.255.0
(DCSS Switching) (Interface 0/1)#exit
```

9. Specify the address to use as the source address for IGMP queries sent from any interface. The global querier address is the IP address of VLAN 100.

```
(DCSS Switching) (Config)#set igmp querier address 192.168.10.2
```

10. Enable IGMP snooping on ports 1–3.

```
(DCSS Switching) (Config#interface 0/1-0/3
(DCSS Switching) (Interface 0/1-0/3)#set igmp
```

11. Configure ports 1–3 as members of VLAN 100.

```
(DCSS Switching) (Interface 0/1-0/3)#vlan participation include 100
(DCSS Switching) (Interface 0/1-0/3)#exit
```

12. Enable IGMP on port 24, and configure the port as a trunk port that connects to the data center switch.

```
(DCSS Switching) (Config)#interface 0/24
(DCSS Switching) (Interface 0/24)#set igmp
(DCSS Switching) (Interface 0/24)#vlan participation include 100
(DCSS Switching) (Interface 0/24)#vlan tagging 100
(DCSS Switching) (Interface 0/24)#end
```

13. Verify the IGMP snooping configuration.

```
(DCSS Switching) #show igmpsnooping

Admin Mode..................................... Enable
Multicast Control Frame Count.................. 0
IGMP Router-Alert check........................ Disabled
IGMP header validation......................... Enabled
Interfaces Enabled for IGMP Snooping........... 0/1
                                                0/2
                                                0/3
                                                0/24
```

```
VLANs enabled for IGMP snooping.............. 100

(DCSS Switching) #show igmpsnooping querier vlan 100

VLAN 100 :   IGMP Snooping querier status
-----------------------------------------
IGMP Snooping Querier VLAN Mode.............. Enable
Querier Election Participate Mode.............. Disable
Querier VLAN Address.......................... 0.0.0.0
Operational State............................. Querier
Operational version........................... 2
Operational Max Resp Time..................... 10
```

After performing the configuration in this example, Host A sends a join message for multicast group 225.1.1.1. Host B sends a join message for group 225.1.1.2. Because IGMP snooping is enabled on the switch and on VLAN 100, the switch listens to the messages and dynamically adds Ports 1 and 2 to the multicast address table. Port 3 did not send a join message, so it does not appear in the table, as the following `show` command indicates.

```
(DCSS Switching) #show mac-address-table multicast

                                                                    Fwd
VLAN ID MAC Address          Source  Type    Description     Interface   Interface
------- ----------------     ------- ------- --------------- ---------   ---------
100     01:00:5E:01:01:01    IGMP    Dynamic  Network Assist   0/1        0/1
100     01:00:5E:01:01:02    IGMP    Dynamic  Network Assist   0/2        0/2
```

When the video server sends multicast data to group 225.1.1.1, Port 1 participates and receives multicast traffic, but Port 2 does not participate because it is a member of a different multicast group. Without IGMP snooping, all ports that are members of VLAN 100 would be flooded with traffic for all multicast groups, which would greatly increase the amount of traffic on the switch.

IGMPv3/SSM Snooping

IGMPv3 adds support for source filtering, which is the ability for a system to report interest in receiving packets **only** from specific source addresses, or from **all but** specific source addresses sent to a particular multicast address. This information is used by snooping switches to avoid delivering multicast packets from specific sources to networks where there are no interested receivers.

No additional configuration is required to enable IGMPv3/SSM snooping. It is enabled or disabled when snooping is enabled on a VLAN/interface. The forwarding database built using IGMPv3 reports is based on the Source IP address, the Multicast Group address, and VLAN. Consider the above configuration example. When Host A sends IGMPv3 IS_IN a report for Group 225.1.1.1 and Sources 192.168.10.1 and 192.168.20.1. As snooping is enabled globally on the switch and also on VLAN 100, two entries are added to MFDB so that multicast traffic with group IP = 225.1.1.1 and if Source IP = 192.168.10.1 or 192.168.20.1 is forwarded to port 1. All other multicast traffic destined to group 225.1.1.1 is dropped. The following command is used to display the SSM forwarding database.

```
(Routing) #show igmpsnooping ssm entries

VLAN                                                Source
ID    Group                 Source Ip               Filter Mode Interfaces
----  --------------------- ---------------------- ----------- --------------
100   225.1.1.1             192.168.10.1           include     0/1
100   225.1.1.1             192.168.20.1           include     0/1
```

LLDP and LLDP-MED

LLDP is a standardized discovery protocol defined by IEEE 802.1AB. It allows stations residing on an 802 LAN to advertise major capabilities physical descriptions, and management information to physically adjacent devices allowing a network management system (NMS) to access and display this information.

LLDP is a one-way protocol; there are no request/response sequences. Information is advertised by stations implementing the transmit function, and is received and processed by stations implementing the receive function. The transmit and receive functions can be enabled/disabled separately on each switch port.

LLDP-MED is an extension of the LLDP standard. LLDP-MED uses LLDP's organizationally-specific Type- Length-Value (TLV) extensions and defines new TLVs that make it easier for a VoIP deployment in a wired or wireless LAN/MAN environment. It also makes mandatory a few optional TLVs from LLDP and recommends not transmitting some TLVs.

The TLVs only communicate information; these TLVs do not automatically translate into configuration. An external application may query the MED MIB and take management actions in configuring functionality.

LLDP and LLDP-MED are used primarily in conjunction with network management tools to provide information about network topology and configuration, and to help troubleshoot problems that occur on the network. The discovery protocols can also facilitate inventory management within a company.

LLDP and the LLDP-MED extension are vendor-neutral discovery protocols that can discover devices made by numerous vendors. LLDP-MED is intended to be used on ports that connect to VoIP phones. Additional applications for LLDP-MED include device location (including for Emergency Call Service/E911) and Power over Ethernet management.

LLDP and Data Center Applications

DCBX uses TLV information elements over LLDP to exchange information, so LLDP must be enabled on the port to enable the information exchange.

Configuring LLDP

This example shows how to configure LLDP settings for the switch and to allow port 0/3 to transmit all LLDP information available.

To configure the switch:

1. Configure the transmission interval, hold multiplier, and reinitialization delay for LLDP PDUs sent from the switch.
```
(DCSS Switching) #configure
(DCSS Switching) (Config)#lldp timers interval 60 hold 5 reinit 3
```
2. Enable port 0/3 to transmit and receive LLDP PDUs.
```
(DCSS Switching) (Config)#interface 0/3
(DCSS Switching) (Interface 0/3)#lldp transmit
(DCSS Switching) (Interface 0/3)#lldp receive
```
3. Enable port 0/3 to transmit management address information in the LLDP PDUs and to send topology change notifications if a device is added or removed from the port.
```
(DCSS Switching) (Interface 0/3)#lldp transmit-mgmt
(DCSS Switching) (Interface 0/3)#lldp notification
```
4. Specify the TLV information to be included in the LLDP PDUs transmitted from port 0/3.
```
(DCSS Switching) (Interface 0/3)#lldp transmit-tlv sys-name sys-desc sys-cap port-desc
```

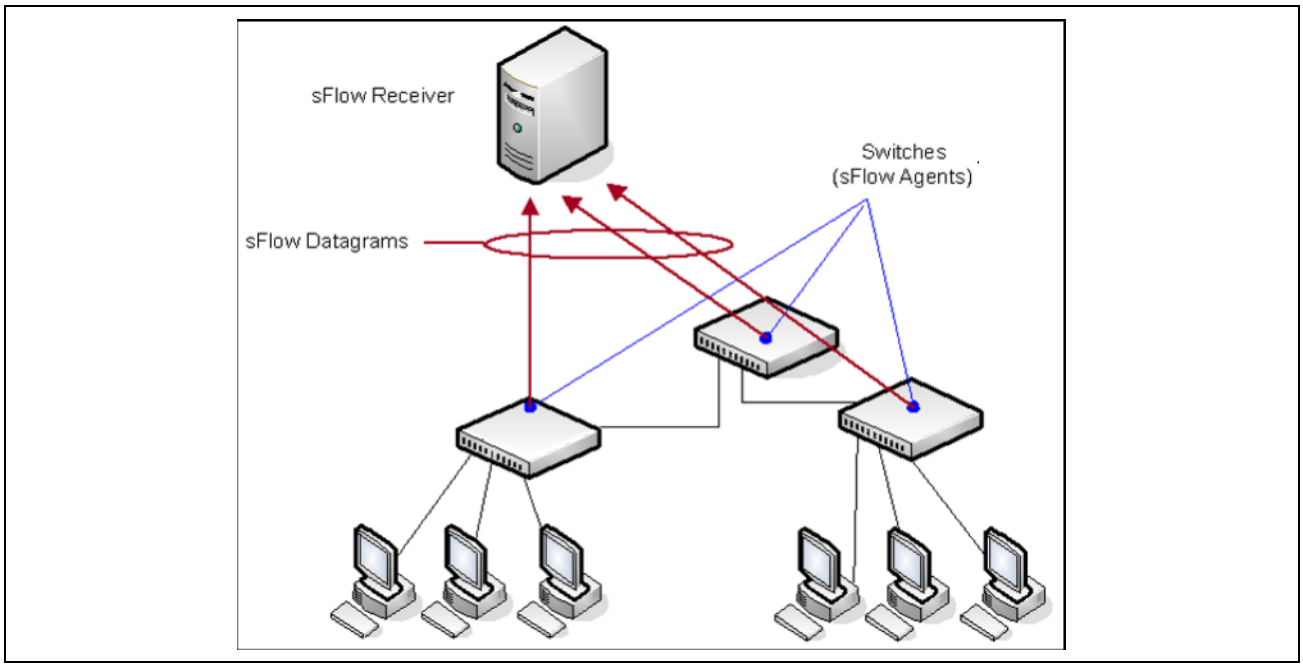
5. Set the port description to be transmitted in LLDP PDUs.
```
(DCSS Switching) (Interface 0/3)#description “Test Lab Port”
```
6. Exit to Privileged EXEC mode.
```
(DCSS Switching) (Interface 0/3)#end
```
7. View global LLDP settings on the switch.
```
(DCSS Switching)#show lldp

LLDP Global Configuration

Transmit Interval..................... 60 seconds
Transmit Hold Multiplier.............. 5
Reinit Delay.......................... 3 seconds
Notification Interval................. 5 seconds
```
8. View summary information about the LLDP configuration on port 0/3.
```
(DCSS Switching)#show lldp interface 0/3
LLDP Interface Configuration

Interface  Link    Transmit  Receive   Notify    TLVs     Mgmt
---------  ------  --------  --------  --------  -------  ----
0/3        Down    Enabled   Enabled   Enabled   0,1,2,3  Y

TLV Codes: 0- Port Description,   1- System Name
           2- System Description, 3- System Capabilities
```
9. View detailed information about the LLDP configuration on port 0/3.
```
(DCSS Switching)#show lldp local-device detail 0/3

LLDP Local Device Detail

Interface: 0/3

Chassis ID Subtype: MAC Address
Chassis ID: 00:10:18:82:15:7B
Port ID Subtype: MAC Address
Port ID: 00:10:18:82:15:7D
System Name:
System Description: Data Center Switch Software AS4600-54T, 48x10Gb, 2x40Gb, 2.1.0.0, Linux 2.6.34.6
Port Description: Test Lab Port
System Capabilities Supported: bridge, router
System Capabilities Enabled: bridge
Management Address:
    Type: IPv4
    Address: 10.27.22.149
```

sFlow

sFlow is an industry standard technology for monitoring high-speed switched and routed networks. DCSS software has a built-in sFlow agent that can monitor network traffic on each port and generate sFlow data to an sFlow receiver (also known as a collector). sFlow helps to provide visibility into network activity, which enables effective management and control of network resources. sFlow is an alternative to the NetFlow network protocol, which was developed by Cisco Systems. The switch supports sFlow version 5.

As illustrated in Figure 16 on page 101, the sFlow monitoring system consists of sFlow Agents (such as DCSS-based switches) and a central sFlow receiver. sFlow Agents use sampling technology to capture traffic statistics from monitored devices. sFlow datagrams forward sampled traffic statistics to the sFlow Collector for analysis. You can specify up to eight different sFlow receivers to which the switch sends sFlow datagrams.

Figure 16: sFlow Architecture



The advantages of using sFlow are:

- It is possible to monitor all ports of the switch continuously, with no impact on the distributed switching performance.
- Minimal memory is required. Samples are not aggregated into a flow-table on the switch; they are forwarded immediately over the network to the sFlow receiver.
- The sFlow system is tolerant to packet loss in the network because statistical modeling means the loss is equivalent to a slight change in the sampling rate.
- sFlow receiver can receive data from multiple switches, providing a real-time synchronized view of the whole network.
- The receiver can analyze traffic patterns based on protocols found in the headers (e.g., TCP/IP, IPX, Ethernet, AppleTalk…). This alleviates the need for a layer 2 switch to decode and understand all protocols.

sFlow Sampling

The sFlow Agent in the DCSS software uses two forms of sampling:

- Statistical packet-based sampling of switched or routed Packet Flows
- Time-based sampling of counters

Packet Flow Sampling and Counter Sampling are performed by sFlow Instances associated with individual Data Sources within an sFlow Agent. Both types of samples are combined in sFlow datagrams. Packet Flow Sampling creates a steady, but random, stream of sFlow datagrams that are sent to the sFlow Collector. Counter samples may be taken opportunistically to fill these datagrams.

To perform Packet Flow Sampling, an sFlow Sampler Instance is configured with a Sampling Rate. Packet Flow sampling results in the generation of Packet Flow Records. To perform Counter Sampling, an sFlow Poller Instance is configured with a Polling Interval. Counter Sampling results in the generation of Counter Records. sFlow Agents collect Counter Records and Packet Flow Records and send them as sFlow datagrams to sFlow Collectors.

Packet Flow Sampling

Packet Flow Sampling, carried out by each sFlow instance, ensures that any packet observed at a Data Source has an equal chance of being sampled, irrespective of the Packet Flow(s) to which it belongs.

Packet Flow Sampling is accomplished as follows:

- A packet arrives on an interface.
- The Network Device makes a filtering decision to determine whether the packet should be dropped.
- If the packet is not filtered (dropped), a decision is made on whether or not to sample the packet.
- A decision is made on whether or not to sample the packet.

  The mechanism involves a counter that is decremented with each packet. When the counter reaches zero a sample is taken.
- When a sample is taken, the counter indicating how many packets to skip before taking the next sample is reset. The value of the counter is set to a random integer where the sequence of random integers used over time is the Sampling Rate.

Counter Sampling

The primary objective of Counter Sampling is to efficiently, periodically export counters associated with Data Sources. A maximum Sampling Interval is assigned to each sFlow instance associated with a Data Source.

Counter Sampling is accomplished as follows:

- sFlow Agents keep a list of counter sources being sampled.
- When a Packet Flow Sample is generated the sFlow Agent examines the list and adds counters to the sample datagram, least recently sampled first. Counters are only added to the datagram if the sources are within a short period, 5 seconds say, of failing to meet the required Sampling Interval.
- Periodically, say every second, the sFlow Agent examines the list of counter sources and sends any counters that must be sent to meet the sampling interval requirement.

The set of counters is a fixed set.

Configuring sFlow

This example shows how to configure the switch so that ports 10-15 and port 23 send sFlow datagrams to an sFlow receiver at the IP address 192.168.20.34. The receiver owner is receiver1, and the timeout is 100000 seconds. A counter sample is generated on the ports every 60 seconds (polling interval), and 1 out of every 8192 packets is sampled.

To configure the switch:

1. Configure information about the sFlow receiver.

```
(DCSS Switching) #configure
(DCSS Switching) (Config)#sflow receiver 1 ip 192.168.20.34
(DCSS Switching) (Config)#sflow receiver 1 owner receiver1 timeout 100000
```

2. Configure the polling and sampling information for ports 10-20.

```
(DCSS Switching) (Config)#interface 0/10-0/15
(DCSS Switching) (Interface 0/10-0/15)#sflow poller 1
(DCSS Switching) (Interface 0/10-0/15)#sflow poller interval 60
(DCSS Switching) (Interface 0/10-0/15)#sflow sampler 1
(DCSS Switching) (Interface 0/10-0/15)#sflow sampler rate 8192
(DCSS Switching) (Interface 0/10-0/15)#exit
```

3. Configure the polling and sampling information for port 23.

```
(DCSS Switching) (Config)#interface 0/23
(DCSS Switching) (Interface 0/23)#sflow poller 1
(DCSS Switching) (Interface 0/23)#sflow poller interval 60
(DCSS Switching) (Interface 0/23)#sflow sampler 1
(DCSS Switching) (Interface 0/23)#sflow sampler rate 8192
(DCSS Switching) (Interface 0/23)#end
```

4. Verify the configured information.

```
(DCSS Switching) #show sflow receivers 1

Receiver Index.............................. 1
Owner String................................ receiver1
Time out.................................... 99400
IP Address:................................. 192.168.20.34
Address Type................................ 1
Port........................................ 6343
Datagram Version............................ 5
Maximum Datagram Size....................... 1400

(DCSS Switching) #show sflow pollers

 Poller        Receiver  Poller
 Data Source   Index     Interval
 -----------   -------   -------
 0/10          1         60
 0/11          1         60
 0/12          1         60
 0/13          1         60
 0/14          1         60
 0/15          1         60
 0/23          1         60
```

```
(DCSS Switching) #show sflow samplers

 Sampler      Receiver  Packet          Max Header
 Data Source  Index     Sampling Rate   Size
 -----------  -------   -------------   ----------
 0/10         1         8192            128
 0/11         1         8192            128
 0/12         1         8192            128
 0/13         1         8192            128
 0/14         1         8192            128
 0/15         1         8192            128
 0/23         1         8192            128
```

6 Configuring Data Center Features

Data Center Technology Overview

Note: The Data Center features and commands in this section are platform-dependent.

DCSS software supports Data Center Bridging (DCB) features to increase the reliability of Ethernet-based networks in the data center. The Ethernet enhancements that DCB provides are well suited for Fibre Channel over Ethernet (FCoE) environments ~~and iSCSI applications~~.

Table 9 provides a summary of the features this section describes.

Table 9: DCB Features

Feature	*Description*
PFC	Provides a way to distinguish which traffic on a physical link is paused when congestion occurs based on the priority of the traffic.
DCBX	Allows DCB devices to exchange configuration information, using type-length-value (TLV) information elements over LLDP, with directly connected peers.
ETS	Supports the ETS configuration and Application Priority TLVs, which are accepted from auto-upstream devices and propagated to auto-downstream devices.
QCN	Manages end-to-end congestion by enabling bridges to signal congestion information to end stations capable of transmission rate limiting to avoid frame loss. VLAN tag-encoded priority values are allocated to segregate frames subject to congestion control, allowing simultaneous support for both congestion control and other higher layer protocols.

Priority-Based Flow Control

Ordinarily, when flow control is enabled on a physical link, it applies to all traffic on the link. When congestion occurs, the hardware sends pause frames that temporarily suspend traffic flow to help prevent buffer overflow and dropped frames.

PFC provides a means of pausing individual priorities within a single physical link. By pausing the congested priority or priorities independently, protocols that are highly loss-sensitive can share the same link with traffic that has different loss tolerances.

This feature is used in networks where the traffic has differing loss tolerances. For example, Fibre Channel traffic is highly sensitive to traffic loss. If a link contains both loss-sensitive data and other less loss-sensitive data, the loss-sensitive data should use a no-drop priority that is enabled for flow control.

Priorities are differentiated by the priority field of the IEEE 802.1Q VLAN header, which identifies an IEEE 802.1p priority value. These priority values must be mapped to internal class-of-service (CoS) values.

The PFC feature allows you to specify the CoS values that should be paused (due to greater loss sensitivity) instead of dropped when congestion occurs on a link. Unless configured as no-drop, all CoS priorities are considered non-pausable ("drop") when priority-based flow control is enabled until no-drop is specifically turned on.

PFC Operation and Behavior

PFC uses a new control packet defined in IEEE 802.1Qbb and therefore is not compatible with IEEE 802.3 Annex 31B flow control. An interface that is configured for PFC will be automatically disabled for flow control. When PFC is disabled on an interface, the flow control configuration for the interface becomes active. Any flow control frames received on a PFC configured interface are ignored.

Each priority is configured as either *drop* or *no-drop*. If a priority that is designated as no-drop is congested, the priority is paused. Drop priorities do not participate in pause. You must configure the same no-drop priorities across the network in order to ensure end-to-end lossless behavior.

Operator configuration of PFC is used only when the port is configured in a manual role. When interoperating with other equipment in a manual role, the peer equipment must be configured with identical PFC priorities and VLAN assignments. Interfaces not enabled for PFC ignore received PFC frames. Ports configured in auto-upstream or auto-downstream roles receive their PFC configuration from the configuration source and ignore any manually-configured information.

Note: This feature is configurable on physical full duplex interfaces only. To enable PFC on a LAG interface, the member interfaces must have the same configuration.

When PFC is disabled, the interface defaults to the IEEE 802.3 flow control setting for the interface. PFC is disabled by default.

If you enable priority-based flow control for a particular priority value on an interface, make sure 802.1p priority values are mapped to CoS values (see "CoS" on page 130).

Data Center Bridging Exchange Protocol

The Data Center Bridging Exchange Protocol (DCBX) is used by DCB devices to exchange configuration information with directly connected peers. DCBX uses type-length-value (TLV) information elements over LLDP to exchange information, so LLDP must be enabled on the port to enable the information exchange. By default, LLDP is enabled on all ports. For more information, see "LLDP and LLDP-MED" on page 99.

The main objective of DCBX is to perform the following operations:

- **Discovery of DCB capability in a peer**: DCBX is used to learn about the capabilities of the peer device. It is a means to determine if the peer device supports a particular feature such as PFC.
- **DCB feature misconfiguration detection**: DCBX can be used to detect misconfiguration of a feature between the peers on a link. Misconfiguration detection is feature-specific because some features may allow asymmetric configuration.
- **Peer configuration of DCB features**: DCBX can be used by a device to perform configuration of DCB features in its peer device if the peer device is willing to accept configuration.

DCBX is expected to be deployed in Fibre Channel over Ethernet (FCoE) topologies in support of lossless operation for FCoE traffic. In these scenarios, all network elements are DCBX enabled. In other words, DCBX is enabled end-to-end.

The DCBX protocol supports the propagation of configuration information for the following features:

- Enhanced Transmission Selection (ETS)

- Priority-based Flow Control (PFC)
- Application Priorities

These features use DCBX to send and receive device configuration and capability information to the peer DCBX device.

The Application Priorities information is simply captured from the peer and potentially propagated to other peers by the DCBX component.

Interoperability with IEEE DCBX

To be interoperable with legacy industry implementations of DCBX protocol, DCSS software uses a hybrid model to support both the IEEE version of DCBX (IEEE 802.1Qaz) and legacy DCBX versions.

DCSS software automatically detects if a peer is operating with either of the two CEE DCBX versions or the IEEE standard DCBX version. This is the default mode. You can also configure DCBX to manually select one of the legacy versions or IEEE standard mode. In auto-detect mode, the switch starts operating in IEEE DCBX mode on a port, and if it detects a legacy DCBX device based on the OUI of the organization TLV, then the switch changes its DCBX mode on that port to support the version detected. There is no timeout mechanism to move back to IEEE mode. Once the DCBX peer times out, multiple peers are detected, the link is reset (link down/up) or as commanded by the operator, DCBX resets its operational mode to IEEE.

The interaction between the DCBX component and other components remains the same irrespective of the operational mode it is executing. For instance, the DCBX component interacts with PFC to get needed information to pack the TLVs to be sent out on the interface. Based on the operational control mode of the port, DCBX packs it in the proper frame format.

DCBX and Port Roles

Each port's behavior is dependent on the operational mode of that port and of other ports in the switch. The port mode is a DCBX configuration item that is passed to the DCBX clients to control the processing of their configuration information. There are four port roles:

- Manual
- Auto-Upstream
- Auto-Downstream
- Configuration Source

Ports operating in the manual role do not have their configuration affected by peer devices or by internal propagation of configuration. These ports have their operational mode, traffic classes, and bandwidth information specified explicitly by the operator. These ports advertise their configuration to their peer if DCBX is enabled on that port. Incompatible peer configurations are logged and counted with an error counter.

The default operating mode for each port is manual. A port that is set to manual mode sets the willing bit for DCBX client TLVs to false. Manually- configured ports never internally propagate or accept internal or external configuration from other ports, in other words, a manual configuration discards any automatic configuration. Manually-configured ports may notify the operator of incompatible configurations if client configuration exchange over DCBX is enabled. Manually- configured ports are always operationally enabled for DCBX clients, regardless of whether DCBX is enabled. Operationally enabled means that the port reports that it is able to operate using the current configuration.

A port operating in the auto-upstream role advertises a configuration, but it is also willing to accept a configuration from the link-partner and propagate it internally to the auto-downstream ports as well as receive configuration propagated internally by other auto-upstream ports. Specifically, the willing parameter is enabled on the port and the recommendation TLV is sent to the peer and processed if received locally. The first auto-upstream port to successfully accept a compatible configuration becomes the configuration source. The configuration source propagates its configuration to other auto-upstream and auto-downstream ports. Only the configuration source may propagate its configuration to other ports internally. Auto-upstream ports that receive internally propagated information ignore their local configuration and utilize the internally propagated information.

Peer configurations received on auto-upstream ports other than the configuration source result in one of two possibilities. If the configuration is compatible with the configuration source, then the DCBX client becomes operationally active on the upstream port. If the configuration is not compatible with the configuration source, then a message is logged indicating an incompatible configuration, an error counter is incremented, and the DCBX client is operationally disabled on the port. The expectation is that the network administrator configures the upstream devices appropriately so that all such devices advertise a compatible configuration.

A port operating in the auto-downstream role advertises a configuration but is not willing to accept one from the link partner. However, the port will accept a configuration propagated internally by the configuration source. Specifically, the willing parameter is disabled on auto-downstream. By default, auto-downstream ports have the recommendation TLV parameter enabled. Auto-downstream ports that receive internally propagated information ignore their local configuration and utilize the internally propagated information. Auto-downstream ports propagate PFC, ETS, and application priority information received from the configuration source.

In the Configuration Source role, the port has been manually selected to be the configuration source. Configuration received over this port is propagated to the other auto-configuration ports, however, no automatic election of a new configuration source port is allowed. Events that cause selection of a new configuration source are ignored. The configuration received over the configuration source port is maintained until cleared by the operator (setting the port to the manual role).

Configuration Source Port Selection Process

When an auto-upstream or auto-downstream port receives a configuration from a peer, the DCBX client first checks if there is an active configuration source. If there is a configuration source already selected, the received configuration is checked against the local port operational values as received from the configuration source, and if compatible, the client marks the port as operationally enabled. If the configuration received from the peer is determined to not be compatible, a message is logged, an error counter is incremented and the DCBX clients become operationally disabled on the port. Operationally disabled means that PFC will not operate over the port. The port continues to keep the link up and exchanges DCBX packets. If a compatible configuration is later received, the DCBX clients will become operationally enabled.

If there is no configuration source, a port may elect itself as the configuration source on a first-come, first-serve basis from the set of eligible ports. A port is eligible to become the configuration source if the following conditions are true:

- No other port is the configuration source.
- The port role is auto-upstream.
- The port is enabled with link up and DCBX enabled.
- The port has negotiated a DCBX relationship with the partner.
- The switch is capable of supporting the received configuration values, either directly or by translating the values into an equivalent configuration.

Whether or not the peer configuration is compatible with the configured values is NOT considered.

The newly elected configuration source propagates DCBX client information to the other ports and is internally marked as being the port over which configuration has been received. Configuration changes received from the peer over the configuration source port are propagated to the other auto-configuration ports. Ports receiving auto-configuration information from the configuration source ignore their current settings and utilize the configuration source information.

When a configuration source is selected, all auto-upstream ports other than the configuration source are marked as willing disabled.

To reduce flapping of configuration information, if the configuration source port is disabled, disconnected or loses LLDP connectivity, the system clears the selection of the configuration source port (if not manually selected) and enables the willing bit on all auto-upstream ports. The configuration on the auto-configuration ports is not cleared (configuration holdover). If the user wishes to clear the configuration on the system in this scenario, the user can put the configuration source port into manual mode.

When a new port is selected as the configuration source, it is marked as the configuration source, the DCBX configuration is refreshed on all auto-configuration ports and each port may begin configuration negotiation with their peer again (if any information has changed).

CoS Queuing

In a typical switch or router, each physical port consists of one or more queues for transmitting packets on the attached network. Multiple queues per port are often provided to give preference to certain packets over others based on user-defined criteria. When a packet is queued for transmission in a port, the rate at which it is serviced depends on how the queue is configured — and possibly the amount of traffic present in the other queues of the port. If a delay is necessary, packets are held in the queue until the scheduler authorizes the queue for transmission. As queues become full, packets have no place to be held for transmission and get dropped by the device.

The drop precedence of a packet is an indication of whether the packet is more or less likely to be dropped during times of queue congestion. Often referred to as packet coloring, a low drop precedence (green) allows the packet to be transmitted under most circumstances, a higher drop precedence (yellow) subjects the packet to dropping when bursts become excessive, while the highest drop precedence (red) discards the packet whenever the queue is congested. In some hardware implementations, the queue depth can be managed using tail dropping or a weighted random early discard, or WRED, technique. These methods often use customizable threshold parameters that are specified on a per-drop-precedence basis.

The DCSS QOS package contains Differentiated Services (DiffServ) support that allows traffic to be classified into streams and given certain QOS treatment in accordance with defined per-hop behaviors. However, the DiffServ feature does not offer direct configuration of the hardware COS queue resources.

The COS Queuing feature allows the switch administrator to directly configure certain aspects of device queuing to provide the desired QOS behavior for different types of network traffic when the complexities of DiffServ are not required. The priority of a packet arriving at an interface can be used to steer the packet to the appropriate outbound COS queue through a mapping table. With the CoS queuing feature, COS queue characteristics such as minimum guaranteed bandwidth, transmission rate shaping, etc. can be configured at the queue (or port) level.

For platforms that support the multistage scheduling architecture, the COS queue feature provides a method to configure Traffic Class Groups (TCGs) to extend the COS queue management. Multiple COS queues can be mapped to a single TCG. Each TCG can have a configured minimum guaranteed bandwidth allocation and a scheduling algorithm similar to the COS queue configuration. The TCG scheduling and bandwidth enforcement occurs after the COS queue scheduling and bandwidth enforcement is performed. Therefore all COS queues mapped to the same TCG share the scheduling and bandwidth properties of the TCG.

CoS Queuing Function and Behavior

Like CoS mapping, CoS queuing uses the concept of trusted and untrusted ports. CoS queuing includes user-configurable settings that affect hardware queue operation.

Trusted Port Queue Mappings

A trusted port is one that takes at face value a certain priority designation within arriving packets. Specifically, a port may be configured to trust one of the following packet fields:

- 802.1p User Priority
- IP Precedence
- IP DSCP

Packets arriving at the port ingress are inspected and their trusted field value is used to designate the COS queue that the packet is placed in when forwarded to the appropriate egress port. A mapping table associates the trusted field value with the desired COS queue.

Un-trusted Port Default Priority

Alternatively, a port may be configured as un-trusted, whereby it does not trust any incoming packet priority designation and uses the port default priority value instead. All packets arriving at the ingress of an un-trusted port are directed to a specific COS queue on the appropriate egress port(s) in accordance with the configured default priority of the ingress port. This process is also used for cases where a trusted port mapping is unable to be honored, such as when a non-IP packet arrives at a port configured to trust the IP precedence or IP DSCP value.

Queue Configuration

Queue configuration involves setting the following hardware port egress queue configuration parameters:

- Scheduler type: strict vs. weighted
- Minimum guaranteed bandwidth
- Maximum allowed bandwidth (i.e. shaping)
- Queue management type: tail-drop vs. WRED
- Tail drop parameters: threshold
- WRED parameters: minimum threshold, maximum threshold, drop probability

Defining these settings on a per-queue basis allows the user to create the desired service characteristics for different types of traffic. The tail drop and WRED parameters are specified individually for each supported drop precedence level.

In addition, the following settings can be specified on a per-interface basis:

- Queue management type: tail drop vs. WRED
- WRED decay exponent

Traffic Class Groups

In DCSS platforms that support multiple levels of egress scheduling, the Traffic Class Groups (TCGs) extend the egress queuing to make use of multiple levels of scheduling. A TCG defines a collection of egress COS Queues. The configuration parameters for the TCG specify the class of service characteristics applied to the aggregated traffic from the associated COS queues. This involves setting the following configuration parameters to each TCG.

- Map one or more COS queues to the TCG.
- Set the scheduling type for each TCG: Strict vs. WDRR

- Set the weight percentages for each TCG.
- Set the minimum guaranteed bandwidth for each TCG. The minimum bandwidth is specified in terms of the percentage of the total link bandwidth.
- Set the maximum allowed bandwidth for each TCG. The maximum bandwidth is specified in terms of the percentage of the total link bandwidth.

TCG configuration parameters are similar to that of COS queues. That is, the configuration of scheduling attributes such as minimum bandwidth, maximum bandwidth, and scheduling algorithm also apply to TCG. The behavior of a TCG with respect to scheduling algorithm and bandwidth allocation configuration is the same as that of COS Queues.

Each TCG is associated with a weight percentage which defines the priority of the TCG to be serviced when WDRR is configured as the scheduling type of the TCG. The weight of the TCG is used only after the minimum guaranteed bandwidth of each of the TCG is met and after all the strict priority TCGs are serviced. The weight of the TCG is then used to prioritize the TCGs among the TCGs that are configured for WDRR.

Enhanced Transmission Selection

Enhanced Transmission Selection (ETS) enables the sharing and redistribution of network bandwidth between various protocols. To support ETS, DCSS software accepts the ETS traffic class group and bandwidth information Application Priority TLV from auto-upstream devices and propagates it to auto-downstream devices. DCSS software supports the reception and propagation of ETS information in the automatic configuration port roles. On DCSS platforms that support hierarchical scheduling, bandwidth allocation and traffic class groups can be configured by ETS TLVs. Platforms that do not support hierarchical scheduling do not use the ETS information to configure traffic class groups or bandwidth allocations.

ETS Operation and Dependencies

Using priority-based processing and bandwidth allocations, different Traffic Class Groups (TCGs) within different types of traffic such as LAN, SAN and Management can be configured to provide bandwidth allocation or best effort transmit characteristics.

For ETS to be operational, the following dependency the following three configuration steps need to occur:

1. Configure COS queues to Traffic Class Group mapping for the egress ports.
2. Configure weight percentage (bandwidth allocation) for each TCG.
3. Enable appropriate scheduling algorithm for each TCG.

CoS information is exchanged with peer DCBX devices using ETS TLVs. As part of the ETS TLV, by default, DCBX advertises the following parameters, which are populated on per port basis.

- Mapping between ingress ports 802.1p priority to Traffic Class Group (TCG).
- Bandwidth percentage (weight percentage) of each Traffic Class Group.
- Scheduling algorithm for each Traffic Class Group.

The mapping between the ingress ports 802.1p priority and TCG is not direct. The mapping depends upon:

- The COS map defining the COS queue that a packet is egress forwarded to for the ingress 802.1p priority.
- Traffic Class Group map defining the COS queue to TCG mapping.

The indirect mapping between the 802.1p priorities and the associated TCG mapping is advertised by DCBX as part of the ETS TLVs. For this indirect mapping to be valid, the following two parameters must be configured (in addition to the configuration of the TCGs):

1. Configure 8021.p priority to COS mapping for the ingress ports.
2. Enable Trust mode on the ingress ports to trust the 802.1p priority present in the frames.

Quantized Congestion Notification (QCN)

QCN is a critical protocol for data center networks in which Ethernet is the common platform, to address the issues of congestion control. In data center networks, factors like flow control, lossless behavior, and latency are extremely important.

The QCN feature attempts to push the network congestion from the heart of core networks to the edges toward end stations. QCN avoids congestion spread by slowing down the end-hosts causing the congestion. QCN works across a single layer-2 domain. As soon as the traffic crosses a router (or an FCoE switch), it enters a different QCN domain.

The QCN congestion-point algorithm is implemented on queues where congestion is expected. Once enabled, it follows following three steps to rectify congestion:

- Congestion Detection — Monitoring the queue size and performing some calculations so that the algorithm can detect congestion as soon as possible.
- Culprit Flow Detection — Identifying the sender end station that is causing the congestion
- Congestion Notification — Issuing a Congestion Notification Message (CNM) to the culprit sender.

QCN operates between Congestion Points (CP), which detect and notify about congestion in the network, and Reaction Points (RP), which originate traffic into the congestion-managed network and receive/process the congestion notifications. The DCSS switch acts a Congestion Point in the network. More specifically, each DCSS switch consists of a set of Congestion points, one per port for each congestion-managed queue.

Configuring PFC

The network in this example handles standard data traffic and traffic that is time sensitive (such as voice and video). The time-sensitive traffic requires a higher priority than standard data traffic. All time-sensitive traffic is configured to use VLAN 100 and has an 802.1p priority of 5, which is mapped to hardware queue 4. The hosts that frequently send and receive time-sensitive traffic are connected to ports 3, 5, and 10, so PFC is enabled on these ports with 802.1p priority 5 traffic as no-drop. The configuration also enables VLAN tagging so that the 802.1p priority is identified. This example assumes that VLAN 100 has already been configured.

Caution! All ports may be briefly shutdown when modifying either flow control or PFC settings. PFC uses a control packet defined in 802.1Qbb and is not compatible with 802.3x FC.

1. Map 802.1p priority 5 to traffic class 4. For more information about traffic classes, see "CoS" on page 130.

```
(DCSS Switching) #configure
(DCSS Switching) (Config)#classofservice dot1p-mapping 5 4
```

2. Enter Interface Configuration mode for ports 3, 5, and 10.

```
(DCSS Switching) (Config)#interface 0/3,0/5,0/10
```

3. Enable PFC and configure traffic marked with 802.1p priority 5 to be paused rather than dropped when congestion occurs.

```
(DCSS Switching)(Interface 0/3,0/5,0/10)#datacenter-bridging
(DCSS Switching) (config-if-dcb)#priority-flow-control mode on
(DCSS Switching) (config-if-dcb)#priority-flow-control priority 5 no-drop
```

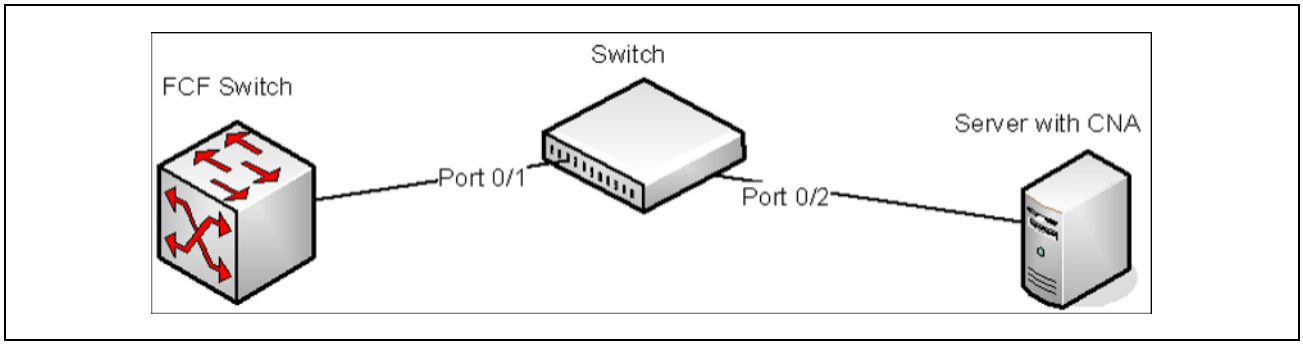
4. Enable VLAN tagging on the ports so the 802.1p priority is identified.

```
(DCSS Switching) (Interface 0/3,0/5,0/10)#vlan participation include 100
(DCSS Switching) (Interface 0/3,0/5,0/10)#vlan tagging 100
(DCSS Switching) (Interface 0/3,0/5,0/10)#exit
```

Configuring DCBX

In this example, port 0/1 on the DCSS switch connects to a FCoE-facing (FCF) switch. This port is designated as a default DCBX auto-upstream port. Port 0/2 on the DCSS switch is directly connected to a Converged Network Adapter (CNA) on a network server. The configuration advertised by the FCF is distributed from port 0/1 to port 0/2. In order to reduce configuration flapping, ports that obtain configuration information from a configuration source port will maintain that configuration for 2× the LLDP timeout, even if the configuration source port becomes operationally disabled.

Figure 17: DCBX Configuration



1. Map 802.1p priority 3 to traffic class 3. For more information about traffic classes, see "CoS" on page 130.

```
(DCSS Switching) #configure
(DCSS Switching) (Config)#classofservice dot1p-mapping 3 3
```

2. Enter Interface Configuration mode for port 1.

```
(DCSS Switching) (Config)#interface 0/1
```

3. Enable the LLDP transmit and receive capability on the port.

```
(DCSS Switching) (Interface 0/1)#lldp transmit
(DCSS Switching) (Interface 0/1)#lldp receive
```

4. Enable the port as the configuration source. This port is connected to a trusted FCF. Configuration received over this port is propagated to the other auto-configuration ports.

```
(DCSS Switching) (Interface 0/1)#lldp dcbx port-role configuration-source
(DCSS Switching) (Interface 0/1)#exit
```

5. Enter Interface Configuration mode for port 2.

```
(DCSS Switching) (Config)#interface 0/2
```

6. Enable the LLDP transmit and receive capability on the port.

```
(DCSS Switching) (Interface 0/2)#lldp transmit
(DCSS Switching) (Interface 0/2)#lldp receive
```

7. Configure the LLDP port role as *auto-down*, which means the port advertises a configuration but is not willing to accept one from the link partner. However, the port will accept a configuration propagated internally by the configuration source (port 0/1).

```
(DCSS Switching) (Interface 0/2)#lldp dcbx port-role auto-down
(DCSS Switching) (Interface 0/2)#exit
```

Configuring CoS Queuing and ETS

This example shows the manual configuration of the CoS queuing feature in a network where traffic needs to be prioritized based on the protocol's frame-loss tolerance. For example, FCoE traffic is highly sensitive to traffic loss. If a port has both loss-sensitive data and other less loss-sensitive data, then the loss-sensitive data is categorized into the same TCG to provide control over the bandwidth allocation and scheduling for the loss-sensitive traffic.

In this example, loss-sensitive traffic is sent with an 801.p priority value of 4, and less loss-sensitive traffic is sent with an 801.p priority value priority of 1. The following steps show how to configure the switch to prioritize the traffic.

1. Configure one to one mapping between 802.1p priority and COS Queue on the ingress port. Frames with 802.1p priority 1 are assigned to COS 1 queue and similarly frames with 802.1p priority 2 are assigned to COS2 and so on.

```
(DCSS Switching) (Config)#classofservice dot1p-mapping 0 0
(DCSS Switching) (Config)#classofservice dot1p-mapping 1 1
(DCSS Switching) (Config)#classofservice dot1p-mapping 2 2
(DCSS Switching) (Config)#classofservice dot1p-mapping 3 3
(DCSS Switching) (Config)#classofservice dot1p-mapping 4 4
(DCSS Switching) (Config)#classofservice dot1p-mapping 5 5
(DCSS Switching) (Config)#classofservice dot1p-mapping 6 6
(DCSS Switching) (Config)#classofservice dot1p-mapping 7 7
```

2. Enable 802.1p Trust mode on all the ports.

```
(DCSS Switching) (Config)#interface 0/1-0/16
(DCSS Switching) (Interface 0/1-0/16)#classofservice trust dot1p
(DCSS Switching) (Interface 0/1-0/16)#exit
```

3. Configure the mapping between COS queues and Traffic Classes Groups. Configure the Traffic Class Group so that 802.1p priority 4 is assigned to TCG1 and 802.1p priority 1 is assigned to TCG2 so that less loss sensitive traffic does not starve the loss sensitive traffic even during traffic bursts. Assign 802.1p priority 7 traffic to TCG0.

```
(DCSS Switching) (Config)#classofservice traffic-class-group 4 1
(DCSS Switching) (Config)#classofservice traffic-class-group 1 2
(DCSS Switching) (Config)#classofservice traffic-class-group 7 0
```

4. Enable VLAN tagging on the ports so the 802.1p priority is identified. The interfaces in this example are members of VLAN 100, which has been previously configured.

```
(DCSS Switching) (Config)#interface 0/1-0/16
(DCSS Switching) (Interface 0/1-0/16)#vlan participation include 100
(DCSS Switching) (Interface 0/1-0/16)#vlan tagging 100
(DCSS Switching) (Interface 0/1-0/16)#exit
```

5. Configure the weight percentage of TCG0 to 10%, and the weights of TCG1 and TCG2 to 45% each.

```
(DCSS Switching) (Config)#traffic-class-group weight 10 45 45
```

6. Associate weighted round robin scheduling with TCG1 and TCG2.

```
(DCSS Switching) (Config)#no traffic-class-group strict 1 2
```

7. Configure TCG0 for strict priority scheduling.

```
(DCSS Switching) (Config)#traffic-class-group strict 0
```

8. Associate TCG0 with CoS queue 7 so that it serves the high priority internal control traffic with CoS 7.

```
(DCSS Switching) (Config)#classofservice traffic-class-group 7 0
```

9. Configure the minimum bandwidth percentage for all the TCGs to be zero.

```
(DCSS Switching) (Config)#traffic-class-group min-bandwidth 0 0 0
```

After performing Step 1–Step 9, the data traffic with an 802.1p priority is sent through TCG1, and 45% of the bandwidth (excluding TCG0 bandwidth) is reserved for TCG1. This protects the TCG1 traffic from traffic that is transmitted on TCG2. Any burst in traffic being transmitted in TCG2 does not affect traffic in TCG1. If TCG2 is not being utilized to the full potential then TCG1 can still use that bandwidth for transmitting TCG1 traffic.

With the configuration in this example, TCG0 with strict priority gets highest priority and can consume the full bandwidth of the pipeline. TCG1 and TCG2 share the remaining bandwidth after TCG0 consumes its share of the pipeline.

Based on this configuration, when the switch sends the configuration ETS TLVs to the peer, the values that are given to DCBX are as follows:

- **Willing Bit** — This bit is set to TRUE for auto-upstream interfaces if there is no configuration source or FALSE if there is a configuration source, and FALSE for auto-downstream and manual ports.
- ~~**Credit-based Shaper support and**~~ **Max TC** — The maximum number of supported traffic classes is 8.

- **Priority Assignment Table** — Table 10 contains the default values advertised by DCBX to the peer DCBX device. If available, the mapping translated from the configuration source is used. This table defines the mapping between the egress Traffic Class Group and ingress 802.1p priority.

Table 10: 802.1p-to-TCG Mapping

802.1p Priority	*Traffic Class*
0	0
1	0
2	0
3	0
4	0
5	0
6	0
7	0

- **TC Bandwidth and TSA Assignment Table** — Table 11 contains the default values advertised by DCBX to the peer DCBX device. If available, the assignments translated from the configuration source are used. This table defines the bandwidth allocated to each Traffic Class Group and the respective scheduling algorithm for each TCG; the scheduling algorithm is enumerated in the IEEE 802.1Q specification.

Table 11: TCG Bandwidth and Scheduling

Traffic Class	*Bandwidth percentage*	*Scheduling Algorithm*
0	10	strict priority (tail-drop) (0)
1	45	strict priority (tail-drop) (0)
2	45	strict priority (tail-drop) (0)

OpenFlow

The purpose of the OpenFlow feature is to demonstrate hardware and software capabilities and to provide a platform on which to build custom networking features, such as the Data Center Tenant Networking feature.

OpenFlow 1.0 mode enables the switch to interoperate with standard OpenFlow controllers such as NOX. The NOX controller has several built-in OpenFlow controller applications that can be used with DCSS switches.

The only difference between the DCSS Tenant Networking mode and OpenFlow 1.0 mode is the switch management paradigm. In Tenant Networking mode, the DCSS switch communicates with the OpenFlow Manager to obtain the configuration for OpenFlow Controllers, CAPWAP tunnels, and Rate Limiters. In OpenFlow 1.0 mode these configuration parameters are defined through the switch user interface.

The underlying DCSS OpenFlow implementation inherently supports multiple hardware tables. Even when operating in OpenFlow 1.0 mode, the OpenFlow Controller can access the different hardware tables by slightly modifying the OFPT_FLOW_MOD message. A pure OpenFlow 1.0 standards-compliant Controller can only access one hardware table. The administrator can also configure the default hardware table accessed by the OpenFlow 1.0 protocol.

Enabling and Disabling OpenFlow

The OpenFlow feature can be enabled or disabled by the network administrator. Although this feature is administratively enabled, it is not operational until the switch has an IP address. A separate operational state indicates whether the OpenFlow feature is operational. If the feature is not operational, then another state indicates the reason for the feature to be disabled.

After administratively disabling the feature, the network administrator must wait until the OpenFlow Feature is operationally disabled before re-enabling the feature. The OpenFlow feature can be administratively disabled at any time.

The administrator can allow the switch to automatically assign an IP address to the OpenFlow feature or they can manually select the address. When the address is assigned automatically and the interface with the assigned address goes offline, the switch will select another active interface if one is available. The OpenFlow feature becomes operationally disabled and re-enabled when a new IP address is selected. If the address is assigned statically, the OpenFlow feature becomes operational only when a switch interface with the matching IP address becomes active.

The automatic IP addresses selection is done in the following order of preference.

1. The loopback interfaces.
2. The routing interfaces.
3. The network interface.

The service port cannot be used as an IP interface for the OpenFlow feature. If routing is enabled, the Network interface cannot be used as the OpenFlow interface.

Once the IP address is selected, it is used until the interface goes offline, the feature is disabled, or a more preferred interface becomes available. The selected IP address is used as the end-point of SSL connections and the end-point of CAPWAP tunnels. When the OpenFlow feature is operationally disabled the switch drops connections with the OpenFlow Managers and Controllers, the switch also purges all flows and configuration settings programmed by the managers and controllers.

If the administrator changes the OpenFlow variant while the OpenFlow feature is enabled, the switch automatically disables and re-enables the OpenFlow feature causing all flows to be deleted and connections to the controllers to be dropped.

If the administrator changes the default hardware table for OpenFlow 1.0 and if the switch is currently operating in OpenFlow 1.0 variant, the OpenFlow feature is automatically disabled and re-enabled.

Interacting with the OpenFlow Manager

The OpenFlow Manager is a device that uses the Open vSwitch management protocol to send commands and retrieve status from the switch.

The DCSS OpenFlow feature supports the OpenFlow Manager only when the DCTENANT_NET component is selected in CCHelper. If the DCTENANT_NET component is not selected, the code for interacting with the OpenFlow manager is excluded from the file system whenever practical, and conditionally compiled out from common files. If the DCTENANT_NET component is selected, but the OpenFlow variant is not configured to be "Tenant Networking" then the communications with the OpenFlow Manager is not supported.

In order to interact with the OpenFlow Manager, the OpenFlow feature must be administratively enabled. The administrator must also configure IP addresses of the OpenFlow Managers using the switch UI. The OpenFlow Manager interaction is handled by the Open vSwitch module called OVSDB.

Deploying OpenFlow

The OpenFlow Manager uses the Management protocol to tell the switch how to communicate with the OpenFlow Controllers and the IP addresses of switches in which CAPWAP tunnels must be set up.

If the administrator selects the OpenFlow 1.0 variant of the OpenFlow protocol, the Controller IP addresses are manually assigned through the switch user interface and the CAPWAP tunnel destination IP addresses are also manually assigned.

OpenFlow Scenarios

The OpenFlow feature is mainly used in a data center network where devices are located in different parts of the network and require layer-2 connectivity.

The tunneling feature enables the devices to communicate over a layer-3 infrastructure. The flow management feature enables customers to avoid scaling problems and loops associated with the layer-2 network.

The OpenFlow feature can also be used in a research environment, but there are two limitations that make the “research” use case less attractive. First, only one OpenFlow controller can manage the switch at a time, meaning that concurrent experiments are not supported unless concurrency is handled at the controller level. Second, the OpenFlow controller has complete access to all ports and VLANs, meaning that using the switch for mixed production and experimental traffic is not advisable.

OpenFlow Interaction with Other Functions

The OpenFlow component interacts with multiple DCSS components by either communicating with these components or sharing common resources with the components.

OpenFlow Variants

OpenFlow 1.0

In **OpenFlow 1.0** mode, the switch is a hybrid OpenFlow switch and supports the OpenFlow 1.0 standard. Hybrid OpenFlow switch means OpenFlow acts as a protocol in conjunction with existing switch functionality. OpenFlow 1.0 mode enables the switch to inter-operate with the standard OpenFlow controllers such as NOX, Beacon, and Big Switch. If COTS versions of these controllers are not available, testing is limited to verification via the OVS_VXCTL tool.

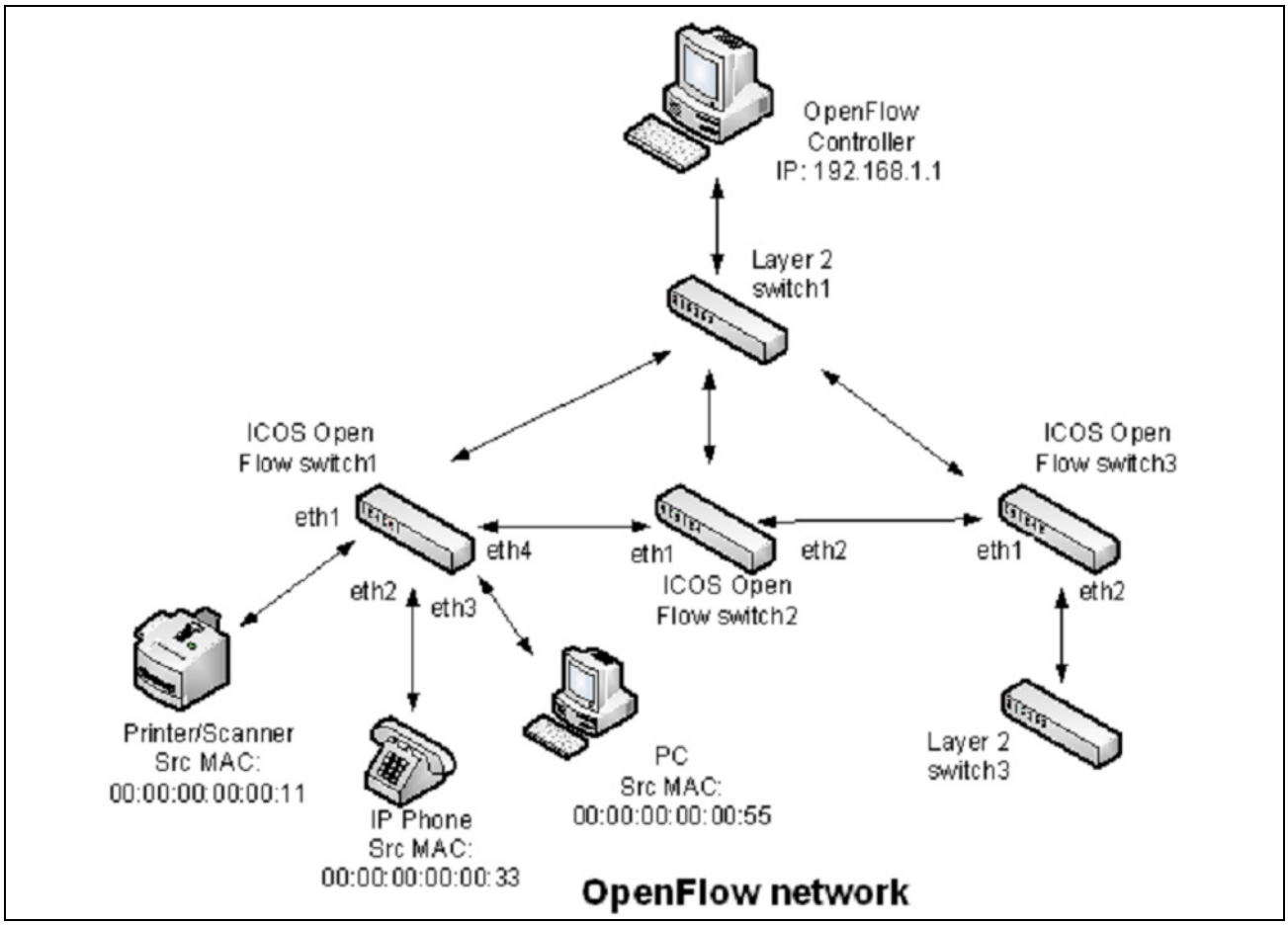
Data Center Tenant Networking

In **Tenant Networking** mode, the DCSS switch communicates with the OpenFlow Manager to obtain the configuration for OpenFlow Controllers, CAPWAP tunnels, and Rate Limiters. In OpenFlow 1.0 mode, these configuration parameters are defined through the switch user interface.

Configuring OpenFlow

The following example uses the network interface's IP address. All DCSS switches shown in Figure 18 have the same OpenFlow configuration.

Figure 18: OpenFlow Network Example



Use the following commands to configure an OpenFlow network:

1. Configure the network protocol as DHCP with the following command:

```
(DCSS Switching) #network protocol dhcp
```

2. Since the controller IP address in this example is configured from the Switch CLI, set the OpenFlow variant mode to **openflow1.0** with the following command:

```
(DCSS Switching) (Config)# openflow variant openflow10
```

3. Set the controller IP address with the following command:

```
(DCSS Switching) (Config)#openflow controller 192.168.1.1 6633 tcp
```

4. To insert the flow into the **OpenFlow 1.0** match table which can match on all OpenFlow 1.0 fields, set the OpenFlow default flow table to **Full-Match** with the following command:

```
(DCSS Switching) (Config)#openflow default-table full-match
```

5. Enable OpenFlow on the switch with the following command:

```
(DCSS Switching) (Config)#openflow enable
```

6. Verify the OpenFlow configuration with the following command:

```
(DCSS Switching) #show openflow

Supported Versions.............................. 1.0
Administrative Mode............................. Enable
Operational Status.............................. Enabled
Disable Reason.................................. None
IP Address...................................... 10.27.65.64
Static IP Mode.................................. Disable
Static IP Address............................... 10.1.1.1
Network MTU..................................... 1518
OpenFlow Variant................................ OpenFlow 1.0
Default Table................................... full-match

  OpenFlow Manager IP:port Addresses
--------------------------------------


(DCSS Switching) #show openflow configured controller

     IP Address          IP Port                  Connection Mode
---------------  ---------------  --------------------------------
    192.168.1.1             6633                              tcp
```

7. The controller installs rules in the DCSS switches. In this example, the following rules have been installed:

   DCSS Switch 1

   - Forward any traffic with ingress port 0/1 with Source MAC 00:00:00:00:00:11 to egress port 0/4
   - Forward any traffic with ingress port 0/1 with Source MAC 00:00:00:00:00:33 to egress port 0/4
   - Forward any traffic with ingress port 0/1 with Source MAC 00:00:00:00:00:55 to egress port 0/4

   DCSS Switch 2

   - Forward any traffic with ingress port 0/1 with Source MAC 00:00:00:00:00:11 to egress port 0/2
   - Forward any traffic with ingress port 0/1 with Source MAC 00:00:00:00:00:33 to egress port 0/2
   - Forward any traffic with ingress port 0/1 with Source MAC 00:00:00:00:00:55 to egress port 0/2

   DCSS Switch 3

   - Forward any traffic with ingress port 0/1 with Source MAC 00:00:00:00:00:11 to egress port 0/2
   - Forward any traffic with ingress port 0/1 with Source MAC 00:00:00:00:00:33 to egress port 0/2
   - Forward any traffic with ingress port 0/1 with Source MAC 00:00:00:00:00:55 to egress port 0/2

8. To verify the installed flows for DCSS Switch 1, use the following command:

```
(DCSS Switching) #show openflow installed flows

Flow 0C9E0D00 type "1DOT0"

Match criteria:
Flow table              24 : Priority            32768
```

```
Ingress port           0/1 : Src MAC  00:00:00:00:00:11
Actions:
Egress port            0/4
Status:
Duration                 7 : Idle                    5 : installed in hardware    1

Flow F6880900 type "1DOT0"

Match criteria:
Flow table              24 : Priority             32768
Ingress port           0/2 : Src MAC  00:00:00:00:00:33
Actions:
Egress port            0/4
Status:
Duration                11 : Idle                    9 : installed in hardware    1

Flow 36370100 type "1DOT0"

Match criteria:
Flow table              24 : Priority             32768
Ingress port           0/3 : Src MAC  00:00:00:00:00:55
Actions:
Egress port            0/4
Status:
Duration              1121 : Idle                 1119 : installed in hardware    1
```

9. To verify the installed flows for DCSS Switch 2, use the following command:

```
(DCSS Switching) #show openflow installed flows

Flow 0C9E0D00 type "1DOT0"

Match criteria:
Flow table              24 : Priority             32768
Ingress port           0/1 : Src MAC  00:00:00:00:00:11
Actions:
Egress port            0/2
Status:
Duration                 7 : Idle                    5 : installed in hardware    1

Flow F6880900 type "1DOT0"

Match criteria:
Flow table              24 : Priority             32768
Ingress port           0/1 : Src MAC  00:00:00:00:00:33
Actions:
Egress port            0/2
Status:
Duration                11 : Idle                    9 : installed in hardware    1

Flow 36370100 type "1DOT0"

Match criteria:
Flow table              24 : Priority             32768
Ingress port           0/1 : Src MAC  00:00:00:00:00:55
Actions:
Egress port            0/2
Status:
Duration              1121 : Idle                 1119 : installed in hardware    1
```

10. To verify the installed flows for DCSS Switch 3, use the following command:

```
(DCSS Switching) #show openflow installed flows

Flow 0C9E0D00 type "1DOT0"

Match criteria:
Flow table              24 : Priority             32768
Ingress port           0/1 : Src MAC  00:00:00:00:00:11
Actions:
Egress port            0/2
Status:
Duration                 7 : Idle                    5 : installed in hardware    1

Flow F6880900 type "1DOT0"

Match criteria:
Flow table              24 : Priority             32768
Ingress port           0/1 : Src MAC  00:00:00:00:00:33
Actions:
Egress port            0/2
Status:
Duration                11 : Idle                    9 : installed in hardware    1

Flow 36370100 type "1DOT0"

Match criteria:
Flow table              24 : Priority             32768
Ingress port           0/1 : Src MAC  00:00:00:00:00:55
Actions:
Egress port            0/2
Status:
Duration              1121 : Idle                 1119 : installed in hardware    1
```

7 Configuring Quality of Service

ACLs

Access Control Lists (ACLs) are a collection of permit and deny conditions, called rules, that provide security by blocking unauthorized users and allowing authorized users to access specific resources.

ACLs can also provide traffic flow control, restrict contents of routing updates, and decide which types of traffic are forwarded or blocked. ACLs can reside in a firewall router, a router connecting two internal networks, or a Layer 3 switch.

DCSS software supports ACL configuration in both the ingress and egress direction. Egress ACLs provide the capability to implement security rules on the egress flows (traffic leaving a port) rather than the ingress flows (traffic entering a port). Ingress and egress ACLs can be applied to any physical port, port-channel (LAG), or VLAN routing port.

Depending on whether an ingress or egress ACL is applied to a port, when the traffic enters (ingress) or leaves (egress) a port, the ACL compares the criteria configured in its rules, in order, to the fields in a packet or frame to check for matching conditions. The ACL forwards or blocks the traffic based on the rules.

Note: Every ACL is terminated by an implicit **deny all** rule, which covers any packet not matching a preceding explicit rule

You can set up ACLs to control traffic at Layer 2, Layer 3, or Layer 4. MAC ACLs operate on Layer 2. IP ACLs operate on Layers 3 and 4. DCSS supports both IPv4 and IPv6 ACLs.

MAC ACLs

MAC ACLs are Layer 2 ACLs. You can configure the rules to inspect the following fields of a packet:

- Source MAC address
- Source MAC mask
- Destination MAC address
- Destination MAC mask
- VLAN ID
- Class of Service (CoS) (802.1p)
- EtherType

L2 ACLs can apply to one or more interfaces. Multiple access lists can be applied to a single interface; sequence number determines the order of execution. You can assign packets to queues using the assign queue option.

IP ACLs

IP ACLs classify for Layers 3 and 4 on IPv4 or IPv6 traffic.

Each ACL is a set of up to ten rules applied to inbound traffic. Each rule specifies whether the contents of a given field should be used to permit or deny access to the network, and may apply to one or more of the following fields within a packet:

- Destination IP with wildcard mask
- Destination L4 Port
- Every Packet
- IP DSCP
- IP Precedence
- IP TOS
- Protocol
- Source IP with wildcard mask
- Source L4 port
- Destination Layer 4 port

ACL Redirect Function

The redirect function allows traffic that matches a permit rule to be redirected to a specific physical port or LAG instead of processed on the original port. The redirect function and mirror function are mutually exclusive. In other words, you cannot configure a given ACL rule with mirror and redirect attributes.

ACL Mirror Function

ACL mirroring provides the ability to mirror traffic that matches a permit rule to a specific physical port or LAG. Mirroring is similar to the redirect function, except that in flow-based mirroring a copy of the permitted traffic is delivered to the mirror interface while the packet itself is forwarded normally through the device. You cannot configure a given ACL rule with both mirror and redirect attributes.

Using ACLs to mirror traffic is considered to be flow-based mirroring since the traffic flow is defined by the ACL classification rules. This is in contrast to port mirroring, where all traffic encountered on a specific interface is replicated on another interface.

ACL Logging

ACL Logging provides a means for counting the number of matches against an ACL rule. When you configure ACL Logging, you augment the ACL deny rule specification with a *log* parameter that enables hardware hit count collection and reporting. The switch uses a fixed five minute logging interval, at which time trap log entries are written for each ACL logging rule that accumulated a non-zero hit count during that interval. You cannot configure the logging interval.

Time-Based ACLs

The time-based ACL feature allows the switch to dynamically apply an explicit ACL rule within an ACL for a predefined time interval by specifying a time range on a per-rule basis within an ACL, so that the time restrictions are imposed on the ACL rule.

With a time-based ACL, you can define when and for how long an individual rule of an ACL is in effect. To apply a time to an ACL, first define a specific time interval and then apply it to an individual ACL rule so that it is operational only during the specified time range, for example, during a specified time period or on specified days of the week.

A time range can be absolute (specific time) or periodic (recurring). If an absolute and periodic time range entry are defined within the same time range, the periodic timer is active only when the absolute timer is active.

Note: Adding a conflicting periodic time range to an absolute time range will cause the time range to become inactive. For example, consider an absolute time range from 8:00 AM Tuesday March 1st 2011 to 10 PM Tuesday March 1st 2011. Adding a periodic entry using the 'weekend' keyword will cause the time-range to become inactive because Tuesdays are not on the weekend.

A named time range can contain up to 10 configured time ranges. Only one absolute time range can be configured per time range. During the ACL configuration, you can associate a configured time range with the ACL to provide additional control over permitting or denying a user access to network resources.

Benefits of using time-based ACLs include:

- Providing more control over permitting or denying a user access to resources, such as an application (identified by an IP address/mask pair and a port number).
- Providing control of logging messages. Individual ACL rules defined within an ACL can be set to log traffic only at certain times of the day so you can simply deny access without needing to analyze many logs generated during peak hours.

ACL Limitations

The following limitations apply to ingress and egress ACLs.

- Maximum of 100 ACLs.
- Maximum number configurable rules per list is 1023.
- Maximum ACL rules (system-wide) is 16384.
- You can configure mirror or redirect attributes for a given ACL rule, but not both.
- The switch hardware supports a limited number of counter resources, so it may not be possible to log every ACL rule. You can define an ACL with any number of logging rules, but the number of rules that are actually logged cannot be determined until the ACL is applied to an interface. Furthermore, hardware counters that become available after an ACL is applied are not retroactively assigned to rules that were unable to be logged (the ACL must be un-applied then re-applied). Rules that are unable to be logged are still active in the ACL for purposes of permitting or denying a matching packet. If console logging is enabled and the severity is set to Info (6) or a lower severity, a log entry may appear on the screen.
- The order of the rules is important: when a packet matches multiple rules, the first rule takes precedence. Also, once you define an ACL for a given port, all traffic not specifically permitted by the ACL is denied access.

ACL Configuration Process

To configure ACLs, follow these steps:

1. Create a MAC ACL by specifying a name.
2. Create an IP ACL by specifying a number.
3. Add new rules to the ACL.
4. Configure the match criteria for the rules.
5. Apply the ACL to one or more interfaces.

Preventing False ACL Matches

Be sure to specify ACL access-list, permit, and deny rule criteria as fully as possible to avoid false matches. This is especially important in networks with protocols such as FCoE that have newly-introduced EtherType values. For example, rules that specify a TCP or UDP port value should also specify the TCP or UDP protocol and the IPv4 or IPv6 EtherType. Rules that specify an IP protocol should also specify the EtherType value for the frame.

In general, any rule that specifies matching on an upper-layer protocol field should also include matching constraints for each of the lower-layer protocols. For example, a rule to match packets directed to the well-known UDP port number 22 (SSH) should also include matching constraints on the IP protocol field (protocol=0x11 or UDP) and the EtherType field (EtherType=0x0800 or IPv4). Table 12 lists commonly-used EtherTypes numbers:

Table 12: Common EtherType Numbers

EtherType	*Protocol*
0x0800	Internet Protocol version 4 (IPv4)
0x0806	Address Resolution Protocol (ARP)
0x0842	Wake-on LAN Packet
0x8035	Reverse Address Resolution Protocol (RARP)
0x8100	VLAN tagged frame (IEEE 802.1Q)
0x86DD	Internet Protocol version 6 (IPv6)
0x8808	MAC Control
0x8809	Slow Protocols (IEEE 802.3)
0x8870	Jumbo frames
0x888E	EAP over LAN (EAPOL – 802.1X)
0x88CC	Link Layer Discovery Protocol
0x8906	Fibre Channel over Ethernet
0x8914	FCoE Initialization Protocol
0x9100	Q in Q

Table 13 lists commonly-used IP protocol numbers:

Table 13: Common IP Protocol Numbers

IP Protocol Number	*Protocol*
0x00	IPv6 Hop-by-hop option
0x01	ICMP
0x02	IGMP
0x06	TCP
0x08	EGP
0x09	IGP
0x11	UDP

ACL Configuration Examples

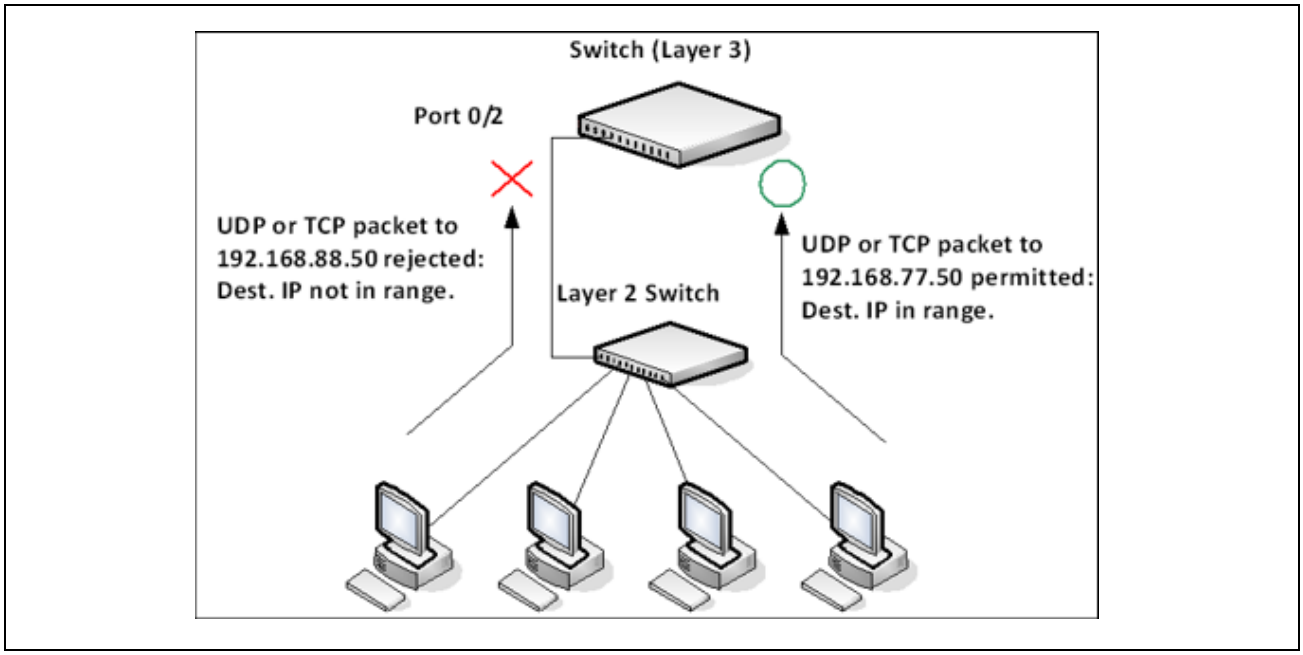
This section contains the following examples:

- Configuring an IP ACL
- Configuring a MAC ACL
- Configuring a Time-Based ACL

Configuring an IP ACL

The commands in this example set up an IP ACL that permits hosts in the 192.168.77.0/24 subnet to send TCP and UDP traffic only to the host with an IP address of 192.168.77.50. The ACL is applied to port 2 on the switch.

Figure 19: IP ACL Example Network Diagram



To configure the switch:

1. Create an extended ACL and configures a rule for the ACL that permits packets carrying TCP traffic that matches the specified Source IP address (192.168.77.0/24), and sends these packets to the specified Destination IP address (192.168.77.50).

```
(DCSS Switching) #config
(DCSS Switching) (Config)#access-list 100 permit tcp 192.168.77.0 0.0.0.255 192.168.77.50 0.0.0.0
```

2. Define the rule to set similar conditions for UDP traffic as for TCP traffic.

```
(DCSS Switching) (Config)#access-list 100 permit udp 192.168.77.0 0.0.0.255 192.168.77.3 0.0.0.255
```

3. Apply the rule to inbound (ingress) traffic on port 2. Only traffic matching the criteria will be accepted on this port.

```
(DCSS Switching) (Config)#interface 0/2
(DCSS Switching) (Interface 0/2)#ip access-group 100 in
(DCSS Switching) (Interface 0/2)#end
```

4. Verify the configuration.

```
(DCSS Switching) #show ip access-lists 100

ACL ID: 100
Inbound Interface(s): 0/2


Rule Number: 1
Action..................................... permit
Match All.................................. FALSE
Protocol................................... 6(tcp)
Source IP Address.......................... 192.168.77.0
Source IP Wildcard Mask.................... 0.0.0.255
Destination IP Address..................... 192.168.77.50
Destination IP Wildcard Mask............... 0.0.0.0

Rule Number: 2
Action..................................... permit
Match All.................................. FALSE
Protocol................................... 17(udp)
Source IP Address.......................... 192.168.77.0
Source IP Wildcard Mask.................... 0.0.0.255
Destination IP Address..................... 192.168.77.3
Destination IP Wildcard Mask............... 0.0.0.255
```

Configuring a MAC ACL

The following example creates a MAC ACL named mac1 that denies all IPX traffic on all ports. All other type of traffic is permitted.

To configure the switch:

1. Create a MAC Access List named mac1

```
(DCSS Switching) #config
(DCSS Switching) (Config)#mac access-list extended mac1
```

2. Configure a rule to deny all IPX traffic, regardless of the source or destination MAC address.

```
(DCSS Switching) (Config-mac-access-list)#deny any any ipx
```

3. Configure a rule to permit all other types of traffic, regardless of the source or destination MAC address.

```
(DCSS Switching) (config-mac-access-list)#permit any any
(DCSS Switching) (config-mac-access-list)#exit
```

4. Bind the ACL to all ports.

```
(DCSS Switching) (Config)#mac access-group mac1 in
(DCSS Switching) (Config)#exit
```

5. View information about the configured ACL.

```
(DCSS Switching) #show mac access-lists

Current number of all ACLs: 2  Maximum number of all ACLs: 100
```

```
MAC ACL Name                    Rules  Direction  Interface(s)     VLAN(s)
------------------------------  -----  ---------  --------------   ----------
mac1                            2      inbound    0/1, 0/2,
                                                  0/3, 0/4,
                                                  0/5, 0/6,
                                                  0/7, 0/8,
                                                  0/9, 0/10,
--More-- or (q)uit

(DCSS Switching) #show mac access-lists mac1

ACL Name: mac1
Inbound Interface(s): 0/1, 0/2, 0/3, 0/4, 0/5, 0/6, 0/7, 0/8, 0/9, 0/10, 0/11, 0/12, 0/13, 0/14, 0/
15, 0/16, 0/17, 0/18, 0/19, 0/20, 0/21, 0/22, 0/23, 0/24, 0/25, 0/26, 0/27, 0/28, 0/29, 0/30, 0/31,
0/32, 0/33, 0/34, 0/35, 0/36, 0/37, 0/38, 0/39, 0/40, 0/41, 0/42, 0/43, 0/44, 0/45, 0/46, 0/47, 0/
48, 0/49, 0/50, 0/51, 0/52, 3/1, 3/2, 3/3, 3/4, 3/5, 3/6, 3/7, 3/8, 3/9, 3/10, 3/11, 3/12, 3/13, 3/
14, 3/15, 3/16, 3/17, 3/18, 3/19, 3/20, 3/21, 3/22, 3/23, 3/24, 3/25, 3/26, 3/27, 3/28, 3/29, 3/30,
3/31, 3/32, 3/33, 3/34, 3/35, 3/36, 3/37, 3/38, 3/39, 3/40, 3/41, 3/42, 3/43, 3/44, 3/45, 3/46, 3/
47, 3/48, 3/49, 3/50, 3/51, 3/52, 3/53, 3/54, 3/55, 3/56, 3/57, 3/58, 3/59, 3/60, 3/61, 3/62, 3/63,
3/64


Rule Number: 1
Action..................................... deny
Ethertype.................................. ipx

Rule Number: 2
Action..................................... permit
Match All.................................. TRUE
```

Configuring a Time-Based ACL

The following example configures an ACL that denies HTTP traffic from 8:00 pm to 12:00 pm and 1:00 pm to 6:00 PM on weekdays and from 8:30 am to 12:30 PM on weekends. The ACL affects all hosts connected to ports that are members of VLAN 100. The ACL permits VLAN 100 members to browse the Internet only during lunch and after hours.

To configure the switch:

1. Create a time range called *work-hours*.

```
(DCSS Switching) #config
(DCSS Switching) (Config)#time-range work-hours
```

2. Configure an entry for the time range that applies to the morning shift Monday through Friday.

```
(DCSS Switching) (config-time-range)#periodic weekdays 8:00 to 12:00
```

3. Configure an entry for the time range that applies to the afternoon shift Monday through Friday.

```
(DCSS Switching) (config-time-range)#periodic weekdays 13:00 to 18:00
```

4. Configure an entry for the time range that applies to Saturday and Sunday.

```
(DCSS Switching) (config-time-range)#periodic weekend 8:30 to 12:30
(DCSS Switching) (config-time-range)#exit
```

5. Create an extended ACL that denies HTTP traffic during the *work-hours* time range.

```
(DCSS Switching) (Config)#access-list 101 deny tcp any any eq http time-range work-hours
```

6. Apply the ACL to ingress traffic in VLAN 100.

```
(DCSS Switching) (Config)#ip access-group 101 vlan 100 in
(DCSS Switching) (Config)#exit
```

7. Verify the configuration.

```
(DCSS Switching) #show ip access-lists 101

ACL ID: 101
Inbound VLAN ID(s): 100

Rule Number: 1
Action..................................... deny
Match All.................................. FALSE
Protocol................................... 6(tcp)
Destination L4 Port Keyword................ 80(www/http)
Time Range Name............................ work-hours
Rule Status................................ inactive
```

CoS

The CoS feature lets you give preferential treatment to certain types of traffic over others. To set up this preferential treatment, you can configure the ingress ports, the egress ports, and individual queues on the egress ports to provide customization that suits your environment.

The level of service is determined by the egress port queue to which the traffic is assigned. When traffic is queued for transmission, the rate at which it is serviced depends on how the queue is configured and possibly the amount of traffic present in other queues for that port. Some traffic is classified for service (i.e., packet marking) before it arrives at the switch. If you decide to use these classifications, you can map this traffic to egress queues by setting up a CoS Mapping table.

Each ingress port on the switch has a default priority value (set by configuring VLAN port priority) that determines the egress queue its traffic gets forwarded to. Packets that arrive without a priority designation, or packets from ports you've identified as "untrusted," get forwarded according to this default.

Trusted and Untrusted Port Modes

Ports can be configured in *trusted* mode or *untrusted* mode with respect to ingress traffic.

Ports in Trusted Mode: When a port is configured in trusted mode, the system accepts at face value a priority designation encoded within packets arriving on the port. You can configure ports to trust priority designations based on one of the following fields in the packet header:

- 802.1 Priority: values 0–7
- IP DSCP: values 0–63

A mapping table associates the designated field values in the incoming packet headers with a traffic class priority (actually a CoS traffic queue).

Ports in Untrusted Mode: If you configure an ingress port in untrusted mode, the system ignores any priority designations encoded in incoming packets, and instead sends the packets to a traffic queue based on the ingress port's default priority.

Traffic Shaping on Egress Traffic

For slot/port interfaces, you can specify a traffic shaping rate for the port (in Kbps) for egress traffic. The traffic shaping rate specifies an upper limit on the transmission bandwidth used.

Defining Traffic Queues

For each queue, you can specify:

- Minimum bandwidth guarantee: A percentage of the port's maximum negotiated bandwidth reserved for the queue.
- Scheduler type – strict/weighted:
  - Strict priority scheduling gives an absolute priority, with traffic in the highest priority queues always sent first, and traffic in the lowest priority queues always sent last.
  - Weighted scheduling requires a specification of priority for each queue relative to the other queues, based on their minimum bandwidth values.

Supported Queue Management Methods

The switch supports the following methods, configurable per-interface-queue, for determining which packets are dropped when the queue is full:

- Taildrop: Any packet forwarded to a full queue is dropped regardless of its importance.
- Weighted Random Early Detection (WRED) drops packets selectively based their drop precedence level. For each of four drop precedence levels on each WRED-enabled interface queue, you can configure the following parameters:
  - Minimum Threshold: A percentage of the total queue size below which no packets of the selected drop precedence level are dropped.
  - Maximum Threshold: A percentage of the total queue size above which all packets of the selected drop precedence level are dropped.
  - Drop Probability: When the queue depth is between the minimum and maximum thresholds, this value provides a scaling factor for increasing the number of packets of the selected drop precedence level that are dropped as the queue depth increases.

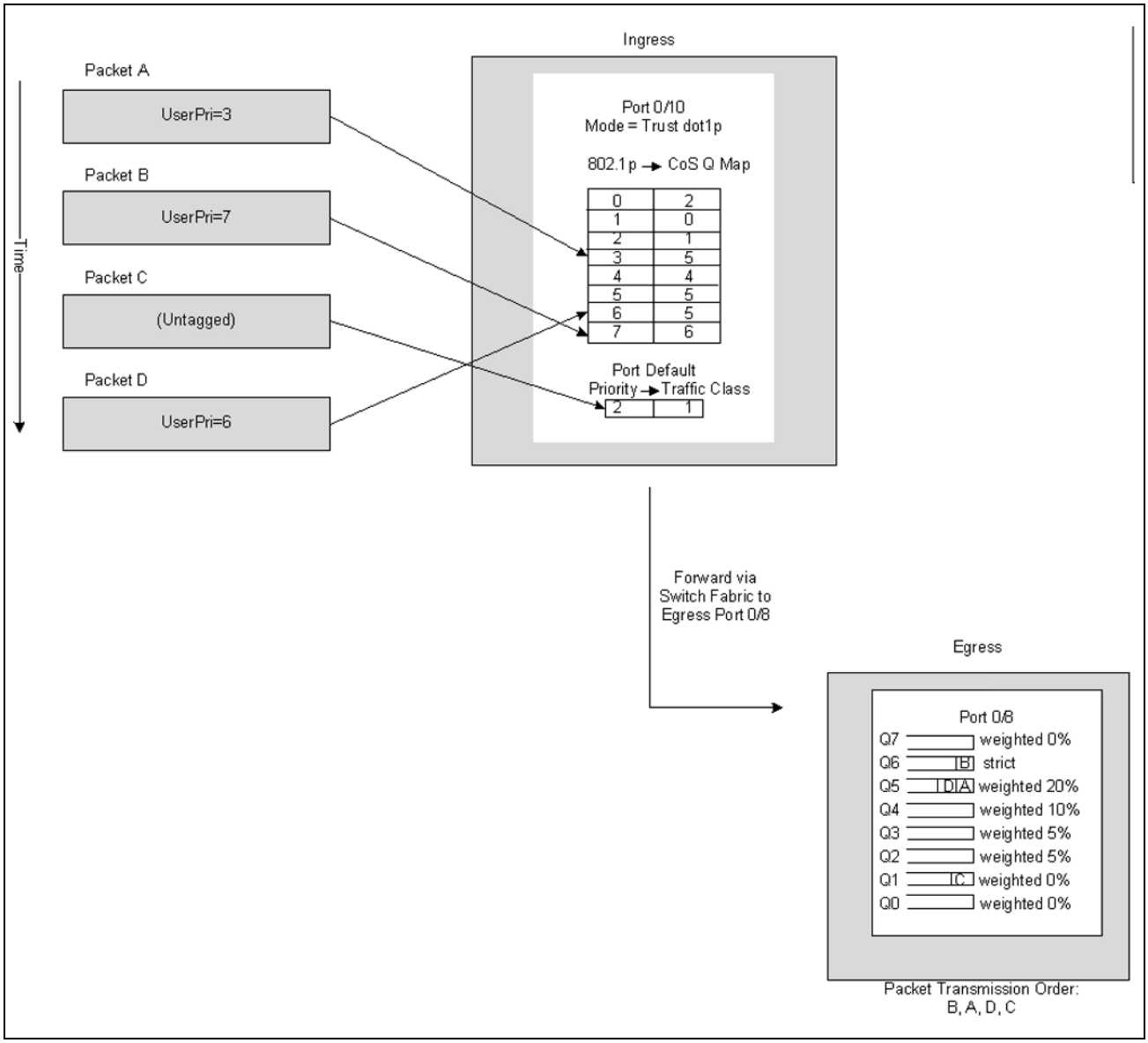
CoS Configuration Example

Figure 20 illustrates the network operation as it relates to CoS mapping and queue configuration.

Four packets arrive at the ingress port 0/10 in the order A, B, C, and D. Port 0/10 is configured to trust the 802.1p field of the packet, which serves to direct packets A, B, and D to their respective queues on the egress port. These three packets utilize the 802.1p to CoS Mapping Table for port 0/10.

In this example, the 802.1p user priority 3 is configured to send the packet to queue 5 instead of the default queue 3. Since packet C does not contain a VLAN tag, the 802.1p user priority does not exist, so Port 0/10 relies on its default port priority (2) to direct packet C to egress queue 1.

Figure 20: CoS Mapping and Queue Configuration



Continuing this example, the egress port 0/8 is configured for strict priority on queue 6, and a weighted scheduling scheme is configured for queues 5-0. Assuming queue 5 has a higher weighting than queue 1 (relative weight values shown as a percentage, with 0% indicating the bandwidth is not guaranteed), the queue service order is 6 followed by 5 followed by 1. Assuming each queue unloads all packets shown in the diagram, the packet transmission order as seen on the network leading out of Port 0/8 is B, A, D, C. Thus, packet B, with a higher user precedence than the others, is able to work its way through the device with minimal delay and is transmitted ahead of the other packets at the egress port.

The following commands configure port 10 (ingress interface) and Port 8 (egress interface).

1. Configure the Trust mode for port 10.

```
(DCSS Switching) #config
(DCSS Switching) (Config)#interface 0/10
(DCSS Switching) (Interface 0/10)#classofservice trust dot1p
```

2. For port 10, configure the 802.1p user priority 3 to send the packet to queue 5 instead of the default queue (queue 3).

```
(DCSS Switching) (Interface 0/10)#classofservice dot1p-mapping 3 5
```

3. For port 10, specify that untagged VLAN packets should have a default priority of 2.

```
(DCSS Switching) (Interface 0/10)#vlan priority 2
(DCSS Switching) (Interface 0/10)#exit
```

4. For Port 8, the egress port, configure a weighted scheduling scheme for queues 5–0.

```
(DCSS Switching) (Config)#interface 0/8
(DCSS Switching) (Interface 0/8)#cos-queue min-bandwidth 0 0 5 5 10 20 40 0
```

5. Configure Port 8 to have strict priority on queue 6.

```
(DCSS Switching) (Interface 0/8)#cos-queue strict 6
(DCSS Switching) (Interface 0/8)#end
```

6. View the configuration.

```
(DCSS Switching) #show interfaces cos-queue 0/8

Interface..................................... 0/8
Interface Shaping Rate........................ 0
WRED Decay Exponent........................... 9

Queue Id   Min. Bandwidth   Scheduler Type   Queue Management Type
--------   --------------   --------------   ---------------------
0          0                Weighted         Tail Drop
1          0                Weighted         Tail Drop
2          5                Weighted         Tail Drop
3          5                Weighted         Tail Drop
4          10               Weighted         Tail Drop
5          20               Weighted         Tail Drop
6          40               Strict           Tail Drop
7          0                Weighted         Tail Drop
```

DiffServ

Standard IP-based networks are designed to provide *best effort* data delivery service. Best effort service implies that the network delivers the data in a timely fashion, although there is no guarantee that it will. During times of congestion, packets may be delayed, sent sporadically, or dropped. For typical Internet applications, such as email and file transfer, a slight degradation in service is acceptable and in many cases unnoticeable. Conversely, any degradation of service has undesirable effects on applications with strict timing requirements, such as voice or multimedia.

DiffServ Functionality and Switch Roles

How you configure DiffServ support in DCSS software varies depending on the role of the switch in your network:

- **Edge device**: An edge device handles ingress traffic, flowing towards the core of the network, and egress traffic, flowing away from the core. An edge device segregates inbound traffic into a small set of traffic classes, and is responsible for determining a packet's classification. Classification is primarily based on the contents of the Layer 3 and Layer 4 headers, and is recorded in the Differentiated Services Code Point (DSCP) added to a packet's IP header.
- **Interior node**: A switch in the core of the network is responsible for forwarding packets, rather than for classifying them. It decodes the DSCP in an incoming packet, and provides buffering and forwarding services using the appropriate queue management algorithms.

Before configuring DiffServ on the switch, you must determine the QoS requirements for the network as a whole. The requirements are expressed in terms of rules, which are used to classify inbound or outbound traffic on a particular interface.

Elements of DiffServ Configuration

During configuration, you define DiffServ rules in terms of classes, policies, and services:

- **Class**: A class consists of a set of rules that identify which packets belong to the class. Inbound traffic is separated into traffic classes based on Layer 2, Layer 3, and Layer 4 header data. The class type **All** is supported; this specifies that every match criterion defined for the class must be true for a match to occur.
- **Policy**: A policy defines the QoS attributes for one or more traffic classes. An attribute identifies the action taken when a packet matches a class rule. An example of an attribute is to mark a packet. The switch supports the ability to assign traffic classes to output CoS queues, and to mirror incoming packets in a traffic stream to a specific egress interface (physical port or LAG).

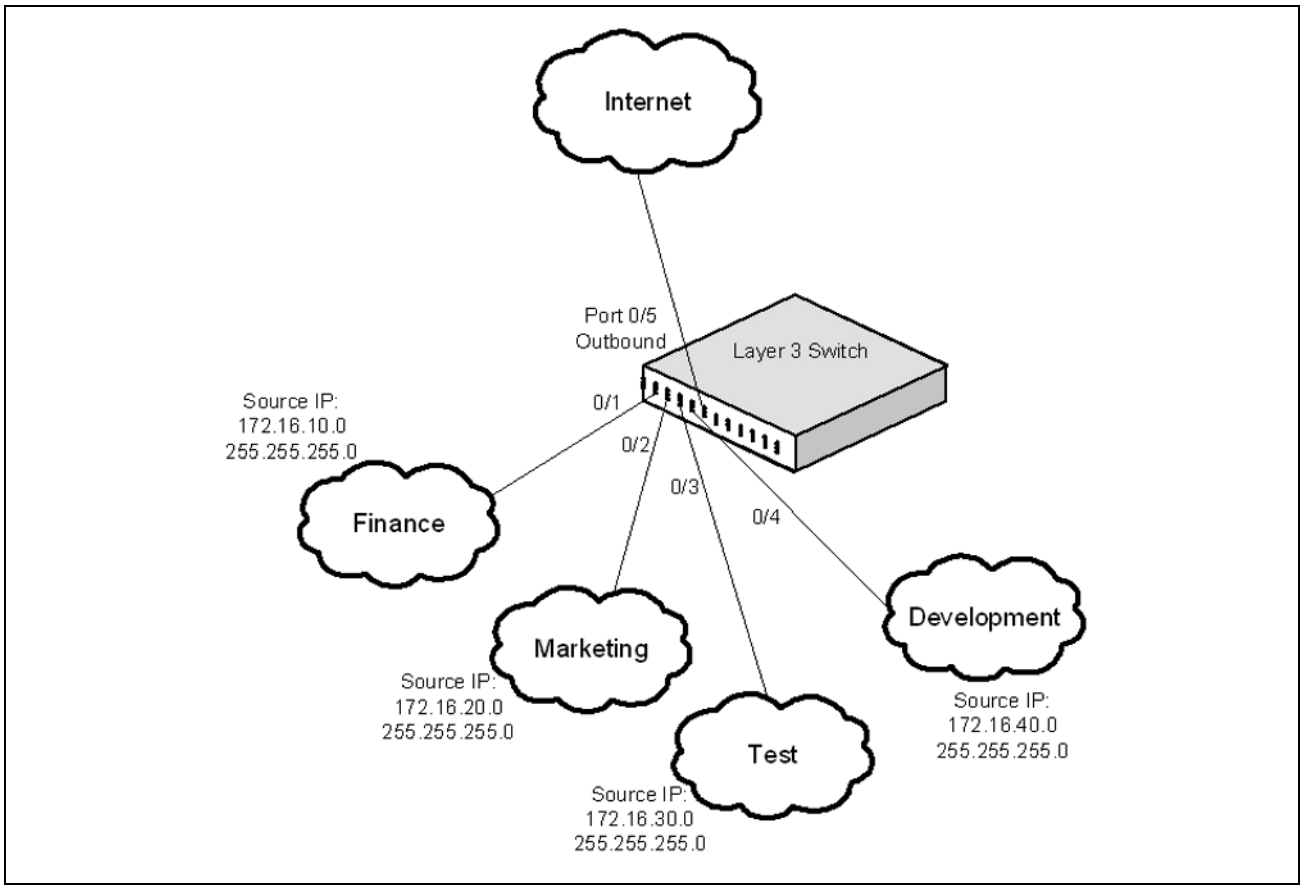
  DCSS software supports the **Traffic Conditioning Policy** type which is associated with an inbound traffic class and specifies the actions to be performed on packets meeting the class rules:

  – Marking the packet with a given DSCP, IP precedence, or CoS value. Traffic to be processed by the DiffServ feature requires an IP header if the system uses IP Precedence or IP DSCP marking.
  – Policing packets by dropping or re-marking those that exceed the class's assigned data rate.
  – Counting the traffic within the class.
- **Service:** Assigns a policy to an interface for inbound traffic.

Configuring DiffServ to Provide Subnets Equal Access to External Network

This example shows how a network administrator can provide equal access to the Internet (or other external network) to different departments within a company. Each of four departments has its own Class B subnet that is allocated 25% of the available bandwidth on the port accessing the Internet.

Figure 21: DiffServ Internet Access Example Network Diagram



The following commands show how to configure the DiffServ example depicted in Figure 21.

1. Enable DiffServ operation for the switch.

```
(DCSS Switching) #config
(DCSS Switching) (Config)#diffserv
```

2. Create a DiffServ class of type *all* for each of the departments, and name them. Also, define the match criteria—Source IP address—for the new classes.

```
(DCSS Switching) (Config)#class-map match-all finance_dept
(DCSS Switching) (Config-classmap)#match srcip 172.16.10.0 255.255.255.0
(DCSS Switching) (Config-classmap)#exit

(DCSS Switching) (Config)#class-map match-all marketing_dept
(DCSS Switching) (Config-classmap)#match srcip 172.16.20.0 255.255.255.0
(DCSS Switching) (Config-classmap)#exit

(DCSS Switching) (Config)#class-map match-all test_dept
(DCSS Switching) (Config-classmap)#match srcip 172.16.30.0 255.255.255.0
(DCSS Switching) (Config-classmap)#exit
```

```
(DCSS Switching) (Config)#class-map match-all development_dept
(DCSS Switching) (Config-classmap)#match srcip 172.16.40.0 255.255.255.0
(DCSS Switching) (Config-classmap)#exit
```

3. Create a DiffServ policy for inbound traffic named *internet_access,* adding the previously created department classes as instances within this policy. This policy uses the assign-queue attribute to put each department's traffic on a different egress queue. This is how the DiffServ inbound policy connects to the CoS queue settings established below.

```
(DCSS Switching) (Config)#policy-map internet_access in
(DCSS Switching) (Config-policy-map)#class finance_dept
(DCSS Switching) (Config-policy-classmap)#assign-queue 1
(DCSS Switching) (Config-policy-classmap)#exit

(DCSS Switching) (Config-policy-map)#class marketing_dept
(DCSS Switching) (Config-policy-classmap)#assign-queue 2
(DCSS Switching) (Config-policy-classmap)#exit

(DCSS Switching) (Config-policy-map)#class test_dept
(DCSS Switching) (Config-policy-classmap)#assign-queue 3
(DCSS Switching) (Config-policy-classmap)#exit

(DCSS Switching) (Config-policy-map)#class development_dept
(DCSS Switching) (Config-policy-classmap)#assign-queue 4
(DCSS Switching) (Config-policy-classmap)#exit
(DCSS Switching) (Config-policy-map)#exit
```

4. Attach the defined policy to interfaces 0/1 through 0/4 in the inbound direction

```
(DCSS Switching) (Config)#interface 0/1-0/4
(DCSS Switching) (Interface 0/1-0/4)#service-policy in internet_access
(DCSS Switching) (Interface 0/1-0/4)#exit
```

5. Set the CoS queue configuration for the (presumed) egress interface 0/5 such that each of queues 1, 2, 3 and 4 get a minimum guaranteed bandwidth of 25%. All queues for this interface use weighted round robin scheduling by default. The DiffServ inbound policy designates that these queues are to be used for the departmental traffic through the assign-queue attribute. It is presumed that the switch will forward this traffic to interface 0/1 based on a normal destination address lookup for internet traffic.

```
(DCSS Switching) (Config)#interface 0/5
(DCSS Switching) (Interface 0/5)#cos-queue min-bandwidth 0 25 25 25 25 0 0 0
(DCSS Switching) (Interface 0/5)#end
```

Index