White Paper

Cisco **StackWise** Technology

This white paper provides an overview of the Cisco StackWise™ technology and the specific mechanisms that it uses to create a unified, logical switching architecture through the linkage of multiple, fixed configuration switches. This paper focuses on the following critical aspects of the Cisco StackWise technology: stacking interconnect behavior, stack creation and modification; Layer 2 and Layer 3 forwarding; and quality-of-service (QoS) mechanisms. The goal of the paper is to help the reader understand how the Cisco StackWise technology delivers advanced performance for voice, video, and gigabit Ethernet applications.

Technology Overview

The Cisco StackWise technology provides an innovative new method for collectively utilizing the capabilities of a stack of switches. Individual switches intelligently join to create a single switching unit with a 32-Gbps switching backplane. Configuration and routing information is shared by every switch in the stack, creating a single switching unit. Switches can be added and deleted from a working stack without affecting performance.

The switches are united into a single logical unit via special stack interconnect cables that create a bidirectional closed-loop path. This bidirectional path acts as a backplane for all the connected switches. Network topology and routing information is updated continuously through the backplane. All stack members have full access to the backplane bandwidth. The stack is managed as a single unit by a master switch, which is elected from one of the stack member switches.

Each switch in the stack has the capability to behave as a master or subordinate in the hierarchy. The master switch is elected and serves as the control center for the stack. The subordinates act as forwarding processors. Each switch is assigned a number. Up to nine separate switches can be joined together. The stack can have switches added and removed without affecting stack performance.

Each stack of Cisco Catalyst 3750 Series switches has a single IP address and is managed as a single object. This single IP management applies to activities such as fault detection, virtual LAN (VLAN) creation and modification, security, and QoS controls. Each stack has only one configuration file, which is distributed to

Figure 1
Stack of Cisco Catalyst 3750 Series Switches with StackWise Technology

Cisco Systems, Inc.
All contents are Copyright © 1992–2003 Cisco Systems, Inc. All rights reserved. Important Notices and Privacy Statement.

each member in the stack. This allows each switch in the stack to share the same network topology, Media Access Control (MAC) address, and routing information.

The Stack Interconnect Functionality

Cisco StackWise technology unites up to nine individual Cisco Catalyst 3750 switches into a single logical unit, using special stack interconnect cables and stacking software. The stack behaves as a single switching unit that is managed by a master switch elected from one of the member switches. The master switch automatically creates and updates all the switching and optional routing tables. A working stack can accept new members or delete old ones without service interruption.

Bidirectional Flow

To efficiently load balance the traffic, packets are allocated between two logical counter-rotating paths. Each counter-rotating path supports 16-Gbps of traffic for a total of 32 Gbps. The egress queues calculate path usage to ensure that the traffic load is equally partitioned.

Whenever a frame is ready for transmission onto the path, a calculation is made to see which path has the most available bandwidth. The entire frame is then copied onto this half of the path. Traffic is serviced depending upon its class of service (CoS) or differentiated services code point (DSCP) designation. Low latency traffic is given priority.

When a break is detected in a cable, the traffic immediately loops back across the single remaining 16-Gbps path to continue forwarding.

Online Stack Adds and Removals

Switches can be added and deleted to a working stack without affecting the backplane performance. When a new switch is added, the master switch automatically configures the unit with the currently running Cisco IOS® software image and configuration of the stack. The stack will gather information such as switching table information and update the MAC tables as new addresses are learned. The network manager does not have to do anything to bring up the switch before it is ready to operate. Similarly, switches can be removed from a working stack without any operational effect on the remaining switches. When the stack discovers that a series of ports is no longer present, it will update this information without affecting forwarding or routing.

Physical Daisy Chain Linkage

The switches are physically connected via a daisy chain as noted in Figure 2. A break in any one of the cables will result in the ring bandwidth being reduced to half of its full capacity. Subsecond timing mechanisms detect traffic problems and immediately institute failover. This mechanism restores dual path flow when the timing mechanisms detect renewed activity on the cable.

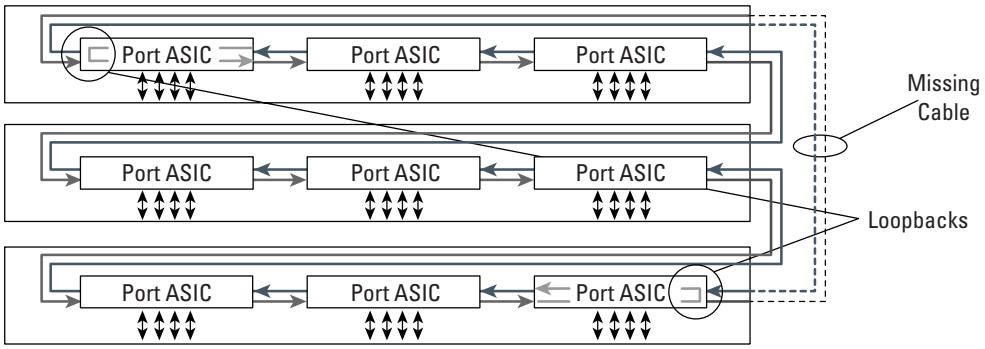
All contents are Copyright © 1992–2003 Cisco Systems, Inc. All rights reserved. Important Notices and Privacy Statement.

Figure 2
Cisco StackWise Technology Resilient Cabling

Subsecond Failover

Within microseconds of a breakage of one part of the ring, all data is switched to the active half of the bidirectional path (Figure 3).

Figure 3
Loopback After Cable Break



The switches continually monitor the stack ports for activity and correct data transmission. If error conditions cross a certain threshold, or there is insufficient electromagnetic contact of the cable with its port, the switch detecting this then sends a message to its nearest neighbor opposite from the breakage. Both switches then divert all their traffic onto the remaining good path.

Single Management IP Address

The stack receives a single IP address as a part of the initial configuration. After the stack IP address is created, the physical switches linked to it become part of the master switch group. When connected to a group, each switch will use the stack IP address. When a new master is elected, it uses this IP address to continue interacting with the network.

Cisco Systems, Inc.
All contents are Copyright © 1992–2003 Cisco Systems, Inc. All rights reserved. Important Notices and Privacy Statement.

Stack Creation and Modification

Stacks are created when individual switches are joined together with stacking cables. When the stack ports detect electromechanical activity, each port starts to transmit information about its switch. When the complete set of switches is known, the stack elects one of the members to be the master switch, which will be responsible for maintaining and updating configuration files, routing information, and other stack information. The entire stack will have a single IP address that will be used by all the switches.

1:N Master Redundancy

1:N master redundancy allows each stack member to serve as a master, providing the highest reliability for forwarding. Each switch in the stack can serve as a master, creating an N+1 availability scheme for network control. In the unlikely event of a single unit failure, all other units continue to forward traffic and maintain operation.

Master Switch Election

The stack behaves as a single switching unit that is managed by a master switch elected from one of the member switches. The master switch automatically creates and updates all the switching and optional routing tables. Any member of the stack can become the master switch. Upon installation, or reboot of the entire stack, an election process occurs among the switches in the stack. There is a hierarchy of selection criteria for the election.

1. *User priority*—The network manager can select a switch to be master.
2. *Hardware and software priority*—This will default to the unit with the most extensive feature set. Catalyst 3750 switches with Enhanced Multilayer Software Image (EMI) versions take precedence over Standard Multilayer Software Image (SMI) versions.
3. *Default configuration*—If a switch has pre-existing configuration information, it will take precedence over switches that have not been configured.
4. *Uptime*—The switch that has been running the longest is selected.
5. *MAC address*—Each switch reports its MAC address to all its neighbors for comparison. The switch with the lowest MAC address is selected.

Master Switch Activities

The master switch acts as the primary point of contact for IP functions such as Telnet sessions, pings, command-line interface (CLI), and routing information exchange. The master is responsible for downloading forwarding tables to each of the subordinate switches. Multicast and unicast routing tasks are implemented from the master. QoS and access control list (ACL) configuration information is distributed from the master to the subordinates. When a new subordinate switch is added, or an existing switch removed, the master will issue a notification of this event and all the subordinate switches will update their tables accordingly.

Shared Network Topology Information

The master switch is responsible for collecting and maintaining correct routing and configuration information. It keeps this information current by periodically sending copies or updates to all the subordinate switches in the stack. When a new master is elected, it re-applies the running configuration from the previous master to ensure user and network continuity.

All contents are Copyright © 1992–2003 Cisco Systems, Inc. All rights reserved. Important Notices and Privacy Statement.

Subordinate Switch Activities

Each switch has tables for storing its own local MAC addresses as well as tables for the other MAC addresses in the stack. The master switch keeps tables of all the MAC addresses reported to the stack. The master also creates a map of all the MAC addresses in the entire stack and distributes it to all the subordinates. Each switch becomes aware of every port in the stack. This eliminates repetitive learning processes and creates a much faster and more efficient switching infrastructure for the system.

Subordinate switches keep their own spanning trees for each VLAN that they support. The master switch keeps a copy of all spanning tree tables for each VLAN in the stack. When a new VLAN is added, or removed, all the existing switches will receive a notification of this event and update their tables accordingly.

Subordinate switches wait to receive copies of the running configurations from the master, and begin start transmitting data upon receipt of the most current information. This ensures that all the switches will use only the most current information and that there is only one network topology used for forwarding decisions.

Multiple Mechanisms for High Availability

The Cisco StackWise technology supports a variety of mechanisms for creating high resiliency in a stack.

- *CrossStack EtherChannel® technology*—Multiple switches in a stack can create an EtherChannel connection. Loss of an individual switch will not affect connectivity for the other switches.
- *Equal cost routes*—Switches can support dual homing to different routers for redundancy.
- 1:N master redundancy—Every switch in the stack can act as the master. If the current master fails, another master is elected from the stack.
- *Stacking cable resiliency*—When a break in the bidirectional loop occurs, the switches automatically begin sending information over the half of the loop that is still intact. If the entire 32 Gbps of bandwidth is being used, QoS mechanisms will control traffic flow to keep jitter and latency-sensitive traffic flowing while throttling lower priority traffic.
- *Online insertion and removal*—Switches can be added and deleted without affecting performance of the stack.
- Distributed Layer 2 forwarding—In the event of a master switch failure, individual switches will continue to forward information based on the tables they last received from the master.
- *RPR+ for Layer 3 resiliency*—Each switch is initialized for routing capability and is ready to be elected as master if the current master fails. Subordinate switches are not reset so that Layer 2 forwarding can continue uninterrupted.

Layer 2 and Layer 3 Forwarding

Cisco StackWise technology offers an innovative method for the management of Layer 2 and Layer 3 forwarding. Layer 2 forwarding is done in distributed method. Layer 3 is done in a centralized manner. This delivers the greatest possible resiliency and efficiency for routing and switching activities across the stack.

Forwarding Resiliency During Master Change

When one master switch becomes inactive and while a new master is elected, the stack continues to function. Layer 2 connectivity continues unaffected. The new master uses its hot standby unicast table to continue processing unicast traffic. Multicast tables and routing tables are flushed and reloaded to avoid loops.

All contents are Copyright © 1992–2003 Cisco Systems, Inc. All rights reserved. Important Notices and Privacy Statement.

High Availability Architecture for Routing Resiliency Using RPR+

The mechanism used for high availability in routing during the change in masters is called Routing Processor Redundancy + (RPR+). It is used in the Cisco 12000 and 7500 series routers and the Catalyst® 6500 Series Switch products for high availability. Each subordinate switch with routing capability is initialized and ready to take over routing functions if the master fails. Each subordinate switch is fully initialized and connected to the master. The subordinates have identical interface addresses, encapsulation types, and interface protocols and services. The subordinate switches continually receive and integrate synchronized configuration information sent by the current master and monitor their readiness to operate through the continuous execution of self-tests. Re-establishment of routes and links happens more quickly than in normal Layer 3 devices because of the lack of time needed to initialize the routing interfaces.

Adding New Members

When the switching stack has established a master, any new switch added afterward automatically becomes a subordinate. All the current routing and addressing information is downloaded into the subordinate so that it can immediately begin transmitting traffic. Its ports become identified with the IP address of the master switch. Global information, such as QoS configuration settings, is downloaded into the new subordinate member.

Cisco IOS Images Must Be Identical

The Cisco StackWise technology requires that all units in the stack run the same release of Cisco IOS Software. When the stack is first built, though, the units can have SMI and EMI version differences. This will automatically cause the EMI versions to become the master switches so the stack can use routing functions. However, the first time you upgrade the Cisco IOS Software release, all units will be required to run the same SMI or EMI version as the master.

Automatic Cisco IOS Software Upgrade/Downgrade from the Master Switch

When a new switch is added to an existing stack, the master switch communicates with the switch to determine if the Cisco IOS image is the same as the one on the stack. If it is the same, the master switch sends the stack configuration to the device and the ports are brought online. If the Cisco IOS image is not the same, one of three things will occur:

1. If the hardware of the new switch is supported by the Cisco IOS image running on the stack, the master will by default download the Cisco IOS image in the master's Flash memory to the new switch, send down the stack configuration, and bring the switch online.
2. If the hardware of the new switch is supported by the Cisco IOS image running on the stack and the user has configured a Trivial File Transfer Protocol (TFTP) server for Cisco IOS image downloads, then the master will automatically download the Cisco IOS image from the TFTP server to the new switch, configure it, then bring it online.
3. If the hardware of the new switch is not supported by the Cisco IOS image running on the stack, the master will put the new switch into a suspended state, notify the user of a version incompatibility, and wait until the user upgrades the master to a Cisco IOS image that supports both types of hardware. The master will then upgrade the rest of the stack to this version, including the new switch, and bring the stack online.

All contents are Copyright © 1992–2003 Cisco Systems, Inc. All rights reserved. Important Notices and Privacy Statement.

Upgrades Apply to All Devices in the Stack

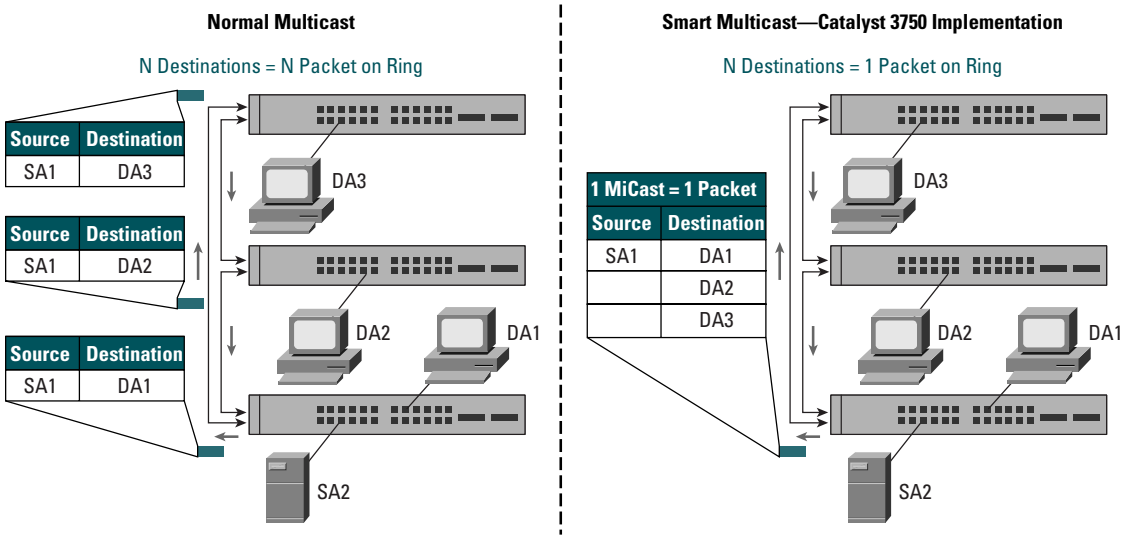
Because the switch stack behaves like a single unit, upgrades apply universally to all members of the stack at once. This means that if an original stack contains both EMI and SMI functions on the various switches, the first time a Cisco IOS Software upgrade is applied, all units in the stack will take on the characteristic of the image applied. While this makes it much more efficient to add functionality to the stack, it is important to make sure all applicable upgrade licenses have been purchased before allowing units to be upgraded from SMI to EMI functions. Otherwise, those units will be in violation of Cisco IOS Software policy.

Smart Unicast and Multicast—One Packet, Many Destinations

The Cisco StackWise technology uses an extremely efficient mechanism for transmitting unicast and multicast traffic. Each data packet is put on the ring only once. This includes multicast packets. Each data packet has a 24-byte header with an activity list for the packet as well as a QoS designator. The activity list specifies the port destination or destinations and what should be done with the packet. In the case of multicast, the master switch identifies which of the ports should receive a copy of the packets and adds a destination index for each port. One copy of the packet is put on the ring. Each switch port that owns one of the destination index addresses then copies this packet. This creates a much more efficient mechanism for the stack to receive and manage multicast information (Figure 4).

Figure 4

Comparison of Normal Multicast in Stackable Switches and Smart Multicast in Cisco Catalyst 3750 Series Switches Using Cisco StackWise Technology



QoS Mechanisms

QoS provides granular control where the user meets the network. This is particularly important for networks migrating to converged applications where differential treatment of information is essential. QoS is also necessary for the migration to Gigabit Ethernet speeds, where congestion must be avoided.

All contents are Copyright © 1992–2003 Cisco Systems, Inc. All rights reserved. Important Notices and Privacy Statement.

QoS Applied at the Edge

The Cisco StackWise technology supports four shaped or shared round-robin ingress queues. Any queue can be shaped or shared. The aggregate egress output can be shaped as well. The network manager can control each of the four queues to set priority for each queue and set a ratio of bandwidth that can be used by each of the queues. The default state is having all queues shared.

When the queues are configured as shared, they are defined as a percentage. For example: queue one is 50 percent, queue two is 30 percent, queue three is 20 percent, and queue four is 10 percent.

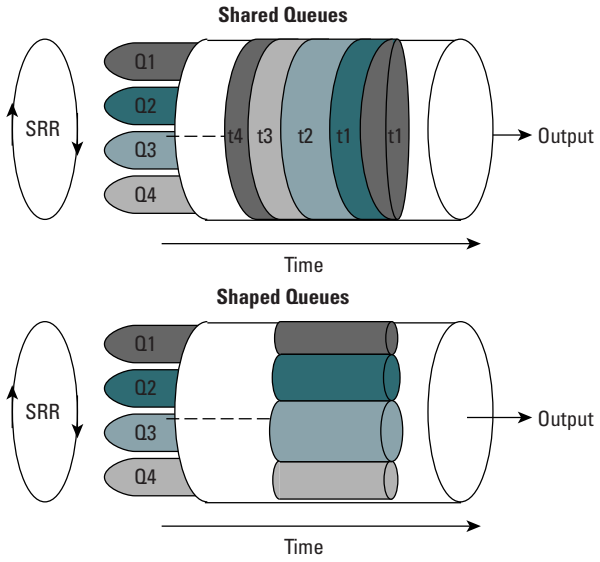
When queues are configured as shaped, they are defined in terms of bandwidth. For example: queue one is 10 Mb, queue two is 40 Mb, queue three is 30 Mb, and queue four is 20 Mb. If one of the queues is not using all its allocated bandwidth, the others can use that queue resource, if they are exceeding their limits. The sum of the bandwidth totals cannot exceed the maximum line rate of the port.

Control traffic for the switch uses one of the four available queues. The network manager can configure which of the queues is used to manage this traffic.

Individual queues can be configured to support either frame or byte count limits.

Figure 5 shows shaped and shared queues.

Figure 5
Comparison of Shaped and Shared Queues



All contents are Copyright © 1992–2003 Cisco Systems, Inc. All rights reserved. Important Notices and Privacy Statement.

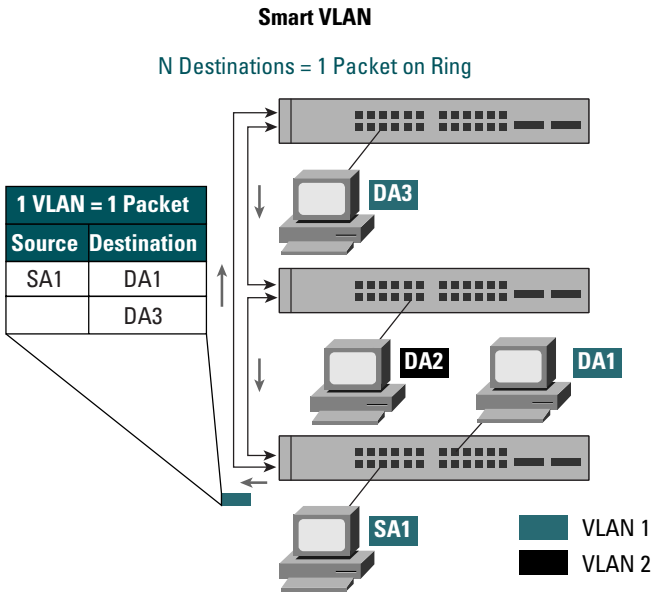
Jumbo Frame Support

The Cisco StackWise technology supports granular jumbo frames up to 9 KB on the 10/100/1000 copper ports.

Smart VLANs

VLAN operation is the same as multicast operation. If the master detects information that is destined for multiple VLANs, it creates one copy of the packet with many destination addresses. This enables the most effective use of the switching backplane (Figure 6).

Figure 6
Smart VLAN Operations



Cross-Stack EtherChannel Connections

Because all the ports in a stack behave as one logical unit, EtherChannel technology can operate across multiple, physical devices in the stack. Cisco IOS Software can aggregate up to eight separate physical ports from any switches in the stack into one logical channel uplink. Up to 12 EtherChannel groups are supported on a stack.

Management

Products using the Cisco StackWise™ technology can be managed by the CLI or by network management packages. Cisco Cluster Management Suite (CMS) Software has been developed specifically for management of Cisco stackable switches. Special wizards for stack units in Cisco CMS Software allow the network manager to configure all the ports in a stack with the same profile. Predefined wizards for data, voice, video, multicast, security, and inter-VLAN routing functions allow the network manager to set all the port configurations at once.

The Cisco StackWise™ technology is also manageable by CiscoWorks.

Summary

Cisco StackWise™ technology allows you to increase the resiliency and the versatility of your network edge to accommodate evolution for speed and converged applications.

All contents are Copyright © 1992–2003 Cisco Systems, Inc. All rights reserved. Important Notices and Privacy Statement.

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

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

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

Asia Pacific Headquarters
Cisco Systems, Inc.
Capital Tower
168 Robinson Road
#22-01 to #29-01
Singapore 068912
www.cisco.com
Tel: +65 6317 7777
Fax: +65 6317 7799

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

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

All contents are Copyright © 1992–2003 Cisco Systems, Inc. All rights reserved. Catalyst, Cisco, Cisco IOS, Cisco Systems, the Cisco Systems logo, and EtherChannel are registered trademarks of Cisco Systems, Inc. and/or its affiliates in the U.S. and certain other countries.

All other trademarks mentioned in this document or Web site are the property of their respective owners. The use of the word partner does not imply a partnership relationship between Cisco and any other company. (0303R)

WH/LW4341 0303