Understanding Token Ring Switching

This appendix discusses several aspects of Token Ring switching and how they relate to the Catalyst 3920. This appendix provides information on the following:

- Switches versus Bridges and Routers on page A-1
- Bridging Modes on page A-2
- Forwarding Modes on page A-3
- Dedicated Token Ring on page A-4
- Token Ring VLANs on page A-4
- Spanning-Tree Protocol on page A-7

Switches versus Bridges and Routers

Because the number of stations that can be connected to any single ring is limited, large Token Ring LANs are divided into smaller rings. Furthermore, because stations must contend for the token with other stations on the same ring, attaching fewer stations to a ring gives each one a greater number of opportunities to transmit and receive information. This microsegmentation of the network results in a larger number of rings or segments.

The traditional method of connecting multiple Token Ring segments is to use a source-routing bridge. For example, bridges are often used to link workgroup rings to the backbone ring. However, the introduction of the bridge can significantly reduce performance at the user's workstation. Further problems may be introduced by aggregate traffic loading on the backbone ring.

To maintain performance and avoid overloading the backbone ring, you can locate servers on the same ring as the workgroup that needs to access the server. However, dispersing the servers throughout the network makes them more difficult to back up, administer, and secure than if they are located on the backbone ring and limits the number of servers that particular stations can access.

Collapsed backbone routers offer greater throughput than bridges, and can interconnect a larger number of rings without becoming overloaded. Routers provide both bridging and routing function between ring and have sophisticated broadcast control mechanisms. These mechanisms become increasingly important as the number of devices on the network increase.

The main drawback of using routers as the campus backbone is the relatively high price-per-port and the fact that the throughput typically does not increase as ports are added. A Token Ring switch is designed to provide wire speed throughput regardless of the number of ports in the switch. In addition, the switch can be configured to provide very low latency between Token Ring ports by using cut-through switching.

As a local collapsed backbone device, a Token Ring switch offers a lower per-port cost and can incur lower interstation latency than a router. In addition, the switch can be used to directly attach large numbers of clients or servers, thereby replacing concentrators. Typically, a Token Ring switch is used in conjunction with a router, providing a high-capacity interconnection between Token Ring segments while retaining the broadcast control and wide-area connectivity provided by the router.

Bridging Modes

The Catalyst 3920 supports the following bridging modes:

- Source-Route Bridging
- Source-Route Transparent Bridging
- Source-Route Switching

Source-Route Bridging

Source-route bridging (SRB) is the original method of bridging used to connect Token Ring segments. A source-route bridge makes all forwarding decisions based upon data in the routing information field (RIF). It does not learn or look up MAC addresses. Therefore, SRB frames without a RIF are not forwarded.

Clients or servers that support source routing typically send an explorer frame to determine the path to a given destination. There are two types of explorer frames: all-routes explorer and spanning-tree explorer. All SRB bridges copy all-routes explorer frames and add their own routing information. For frames that are received from or sent to ports that are in the spanning-tree forwarding state, bridges copy spanning-tree explorer frames and add their own routing information. Because all-routes explorer frames will traverse all paths between two devices, they are used in path determination. Spanning-tree explorer frames are used to send datagrams because the spanning tree will ensure that only one copy of an spanning-tree explorer frame is sent to each ring.

Note The spanning tree used with source-routing is different from the IEEE spanning tree used in transparent bridges. The Catalyst 3920 supports both types of spanning-tree algorithms.

Source-Route Transparent Bridging

Source-route transparent (SRT) bridging is an IEEE standard that combines source-route bridging and transparent bridging. An SRT bridge forwards frames that do not contain a RIF based on the destination MAC address. Frames that contain a RIF are forwarded based upon source-routing.

The SRT bridge only runs the IEEE STP. It does not support the IBM STP.

Source-Route Switching

Similar to a transparent bridge, the Catalyst 3920 can forward broadcast, multicast, and unicast frames based on MAC address. If, however, you have source-route bridges in your network, the Catalyst 3920 can forward frames based on the RIF. This dual frame-forwarding technology is called source-route switching.

In source-route switching, the switch learns and forwards frames based on source route descriptors for stations that are one or more source-route bridge hops away. A route descriptor is a portion of a RIF that indicates a single hop. It is defined as a ring number and a bridge number. When a source-routed frame enters the switch, the switch learns the route descriptor for the hop closest to the switch. Frames received from other ports with the same next-hop route descriptor as their destination will be forwarded to that port.

The key difference between SRB and source-route switching is that while a source-route switch looks at the RIF, it never updates the RIF. Therefore, all ports in a source-route switch group have the same ring number.

Source-route switching provides the following benefits:

- The switch does not need to learn the MAC addresses of the devices on the other side of a source-route bridge. Therefore, the number of MAC addresses that the switch must learn and maintain is significantly reduced.
- The switch can support parallel source-routing paths.
- An existing ring can be partitioned into several segments without requiring a change in the existing ring numbers or the source-route bridges.
- The switch can support duplicate MAC addresses if the stations reside on LAN segments with different LAN IDs (ring numbers).

Forwarding Modes

The Catalyst 3920 supports the following forwarding modes:

- Store-and-Forward
- Cut-Through
- Adaptive Cut-Through

Store-and-Forward

Store-and-forward is the traditional mode of operation for a bridge and is one of the modes supported by the Catalyst 3920. In store-and-forward, the port adapter reads the entire frame into memory and then determines whether the frame should be forwarded. At this point, the frame is also examined for any errors (frames with errors are not forwarded). If the frame contains no errors, it is sent to the destination port for forwarding.

While store-and-forward reduces the amount of error traffic on the LAN, it also causes a delay in frame forwarding that is dependent upon the length of the frame.

Cut-Through

In cut-through mode, the Catalyst 3920 transfers nonbroadcast packets between ports without buffering the entire frame into memory. Instead, when a port on the Catalyst 3920 that is operating in cut-through mode receives the first few bytes of a frame, it analyzes the packet header to determine the destination of the frame, establishes a connection between the input and output ports, and, when the token becomes available, it transmits the frame onto the destination ring.

In accordance with specification ISO/IEC 10038, the Catalyst 3920 uses Access Priority 4 to gain priority access to the token on the output ring if the outgoing port is operating in half-duplex mode. This increases the proportion of packets that can be cut through and makes it possible for the Catalyst 3920 to reduce the average interstation latency.

In certain circumstances, however, the cut-through technique cannot be applied and the Catalyst 3920 must buffer frames into memory.

For example, buffering must be performed in the following circumstances:

- The Catalyst 3920 has two packets to transmit to the same ring.
- A packet is switched between 4- and 16-Mbps rings.
- The destination ring is beaconing.

Adaptive Cut-Through

With adaptive cut-through mode, the user can configure the switch to automatically use the best forwarding mode based on user-defined thresholds. In adaptive cut-through mode, the ports operate in cut-through mode unless the number of forwarded frames that contain errors exceeds a specified percentage. When this percentage is exceeded, the switch automatically changes the mode of the port to store-and-forward. Then, once the number of frames containing errors falls below a specified percentage, the operation mode of the ports is once again set to cut through.

Dedicated Token Ring

Classic 4- and 16-Mbps Token Ring adapters must be connected to a port on a concentrator. These adapters are also limited to operating in half-duplex mode. In half-duplex mode, the adapter can only be sending or receiving a frame; it cannot do both simultaneously.

Dedicated Token Ring, developed by the IEEE, defines a method in which the switch port can emulate a concentrator port, thereby eliminating the need for an intermediate concentrator. In addition, dedicated Token Ring defines a new full-duplex data passing mode called Transmit Immediate, which eliminates the need for a token and allows the adapter to transmit and receive simultaneously.

Dedicated Token Ring is particularly useful for providing improved access to servers. A server can be attached directly to a switch. This allows the server to take advantage of the full 16 Mbps available for sending and receiving and results in an aggregate bandwidth of 32 Mbps.

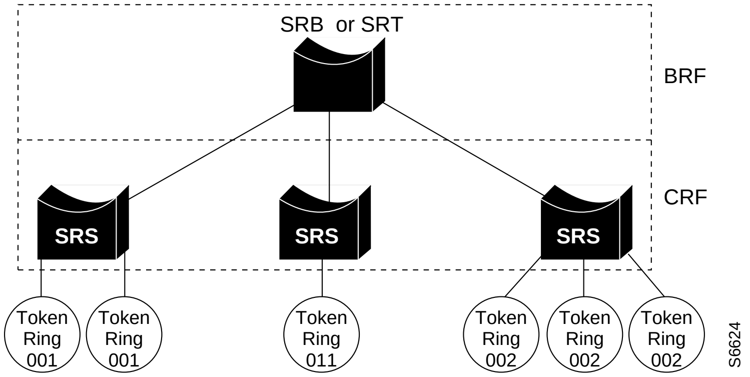
Token Ring VLANs

Within a Token Ring VLAN, distributed rings can be formed by defining groups of ports that have the same ring number. The IEEE calls such a port group a Token Ring Concentrator Relay Function (TrCRF). A TrCRF is limited to the ports in a single Catalyst 3920 or those within a stack of Catalyst 3920s.

The ring number of the TrCRF can be defined or learned from external bridges. Within the TrCRF, source-route switching is used for forwarding based on either MAC addresses or route descriptors. If desired, the entire VLAN can operate as a single ring.

Frames can be switched between ports within a single TrCRF.

Figure A-1 Token Ring VLANs



As shown in Figure A-1, multiple TrCRFs can be interconnected using a single Token Ring Bridge Relay Function (TrBRF). For source routing, the switch appears as a single bridge between the distributed rings. The TrBRF can function as an SRB or SRT bridge running either the IBM or IEEE STP. If SRB is used, duplicate MAC addresses can be defined on different distributed rings.

To accommodate SNA traffic, you can use a combination of SRT and SRB modes. In a mixed mode the TrBRF considers some ports (internal ports connected to TrCRFs) to be operating in SRB mode while others are operating in SRT mode.

The TrBRF can be extended across a network of switches via high-speed uplinks between the switches. These links must have the ability to multiplex multiple VLANs and provide the necessary information to support distributed rings.

VLAN Trunking Protocol

You use the Cisco VLAN Trunking Protocol (VTP) to set up and manage VLANs across an entire management domain. When new VLANs are added to a device (Cisco router or LAN switch) in a management domain, VTP can be used to automatically distribute the information to other trunks of all of the devices in the management domain. This distribution ensures VLAN naming consistency and connectivity between all devices in the domain by allowing each device in the domain to learn of any new VLANs added to other devices in the domain. The VTP is transmitted on all trunk connections, including Inter-Switch Link (ISL) and ATM LAN Emulation (LANE). Although the Catalyst 3920 does not contain any trunk ports, it may be part of a stack that contains a switch with an ISL module.

On boot up, the Catalyst switch with a trunk port sends out periodic requests for VTP configuration on all of its trunks until it receives a summary advertisement from a neighbor. It uses that summary advertisement to determine whether its currently stored configuration is obsolete. If the stored configuration is obsolete, the Catalyst switch requests all VTP information from the neighbor.

The Catalyst switch transmits VTP frames on its trunk ports, advertising its management domain name, configuration revision number, and VLAN information that it has learned. Other Catalyst switches in the domain use these advertisements to learn about any new VLANs that are configured in the transmitting switch. This process of advertising and learning allows a new VLAN to be created and configured on only one switch in the management domain. This information is then learned automatically by all of the other devices in the domain.

The Catalyst switch can operate in three different VTP modes: server, client, or transparent.

- In server mode, the switch permits changes to the administrative domain's global VLAN configuration from the local device. Redundancy in a network domain can be created by using multiple VTP servers.
- In client mode, the switch accepts configuration changes from other devices in the administrative domain, but will not permit local changes to the database.
- In transparent mode, the switch forwards any VTP packets received on the default VLANs of any trunk onto the default VLANs of all other trunks.

  Use VTP transparent mode to have a Catalyst switch not participate in VTP and yet not have it cut off VTP configuration from propagating beyond it. In transparent mode, VTP packets received on one trunk are automatically propagated unchanged to all other trunks on the device but are ignored on the device itself.

Note To enable ring number learning for TrCRFs, the VTP mode must be set to transparent (which is the default) and the ring number on the VTP VLAN Parameter Configuration for the TrCRF must be set to auto (which is the default). If you have set the VTP to client or server, you cannot set the ring number to auto.

VTP Start Up

When a Catalyst switch is booted for the first time (and when it is rebooted after a nonvolatile random-access memory [NVRAM] reset), it comes up in no-domain mode. The no-domain mode means there is no domain name configured in the switch. While in no-domain mode, a switch will not attempt to advertise its own current configuration. If and when it receives an advertisement from any neighbor on any trunk, it will immediately accept the management domain name from the neighbor's advertisement as its own. After receiving all of the neighbor's configuration data, it will begin advertising this data regularly (after a reboot) on all of its trunks.

Security

A checksum is calculated using an arbitrary security value that is appended to the front end and the back end of the data in a VTP configuration. When a VTP device has received all of the parts of the VTP configuration, it recalculates the checksum using its own security value derived from the password that has been configured locally. The device will not accept the new configuration if the checksums do not match.

On all Cisco VTP devices, the default initial configuration of the security value is all zeroes. Therefore, VTP devices will always accept one another's VLAN configurations as long as none of the security values on any of the devices have been modified. To make use of the security feature, a password needs to be set. The password must be the same for the management domain on all devices in the domain. Neither the password nor the security value itself is ever advertised over the network.

Caution If passwords are set, a management domain does not function properly if the same management domain password is not assigned to each Catalyst switch in the domain.

Spanning-Tree Protocol

The STP is a broadcast algorithm used by network bridge connections to dynamically discover a loop-free subset of the network topology while maintaining a path between every pair of LANs or VLANs in the network.

To accomplish this, the STP blocks ports that, if active, would create bridging loops. If the primary link fails, it activates one of the blocked bridge ports to provide a new path through the network.

In a traditional bridged network, there is one STP for each bridge connection. Each bridge maintains its own database of configuration information and transmits and receives only on those ports belonging to the bridge. The type of STP that runs on a bridge depends on the transmission mode of the bridge connection (whether the connection is transparent, SRB, source-route switching, or SRT).

In a switched network, you can configure virtual networks. A switch can have ports that belong to different VLANs, some of which may span several switches. To prevent loops in the bridged connections between the VLANs, you should configure the STP. As discussed in the “Token Ring VLANs” section, in a Token Ring switch, there are two levels of VLANs. The grouping of ports (TrCRFs) is connected by logical bridges (TrBRFs).

Therefore, in a Token Ring switched network, to ensure loops are removed from the topology you must configure a separate STP for each logical bridge (TrBRF) and for each of the port groupings (TrCRF) configured for a VLAN.

How the STP Algorithm Works

The following is a general summary of how the STP eliminates loops in the network:

1. Each bridge is assigned an 8-byte unique bridge identifier.

   The first 2 bytes are a priority field, and the last 6 bytes contain one of the bridge’s MAC addresses. The bridge with the lowest bridge identifier among all bridges on all LAN segments is the root bridge. The network administrator can assign a lower bridge priority to a selected bridge to control which bridge becomes the root, or the administrator can use default bridge priorities and allow the STP to determine the root.

2. Each bridge port is associated with a path cost.

   The path cost represents the cost of transmitting a frame to a bridged segment through that port. A network administrator typically configures a cost for each port based on the speed of link (for example, the cost of a port connected to a 16-Mbps LAN could be assigned a lower path cost than a port connected to a 4-Mbps LAN).

3. Each bridge determines its root port and root path cost.

   The root port is the port that represents the shortest path from itself to the root bridge. The root path cost is the total cost to the root. All ports on the root bridge have a zero cost.

4. All participating bridges elect a designated bridge from among the bridges on that LAN segment.

   A designated bridge is the bridge on each LAN segment that provides the minimum root path cost. Only the designated bridge is allowed to forward frames to and from that LAN segment toward the root.

5 All participating bridges select ports for inclusion in the spanning tree.

The selected ports will be the root port plus the designated ports for the designated bridge. Designated ports are those where the designated bridge has the best path to reach the root. In cases where two or more bridges have the same root path cost, the bridge with the lowest bridge identifier becomes the designated bridge.

6 Using the preceding steps, all but one of the bridges directly connected to each LAN segment are eliminated, thereby removing all multiple LAN loops.

How Spanning-Tree Information is Shared

The STP calculation requires that bridges communicate with other bridges in the network that are running the STP. Each bridge is responsible for sending and receiving configuration messages called bridge protocol data units (BPDUs).

BPDUs are exchanged between neighboring bridges at regular intervals (typically 1 to 4 seconds) and contain configuration information that identifies the:

- Bridge that is presumed to be the main bridge or root (root identifier)
- Distance from the sending bridge to the root bridge (called the root path cost)
- Bridge and port identifier of the sending bridge
- Age of the information contained in the configuration message

If a bridge fails and stops sending BPDUs, the bridges detect the lack of configuration messages and initiate a spanning-tree recalculation.

BPDU Field Formats

Figure A-2 shows the format of the fields inside a BPDU.

Note All fields in the BPDU are common to all STPs except for the Port ID field. If the BPDU is an IEEE or Cisco STP BPDU message, the Port ID field specifies the transmitting port number of the originating bridge. If the BPDU is an IBM STP BPDU message, then the Port ID field specifies the ring and bridge number through which the message was sent.

Figure A-2 BPDU Field Formats

2	**1**	**1**	**1**	**8**	**4**	**8**	**2**	**2**	**2**	**2**	**2**
Protocol Identifier	Version	Message Type	Flags	Root ID	Root Path Cost	Bridge ID	Port ID	Message Age	Maximum Age	Hello Time	Forward Delay

BPDU Configuration Message Fields

Protocol Identifier—Identifies the protocol. This field contains the value zero.

Version—Identifies the version. This field contains the value zero.

Message Type—Identifies the message type. This field contains the value zero.

Flags—1-byte field, of which only the first two bits are used. The topology change (TC) bit signals a topology change. The topology change acknowledgment (TCA) bit is set to acknowledge receipt of a configuration message with the TC bit set.

Root ID—Identifies the root bridge by listing its 2-byte priority followed by its 6-byte ID.

Root Path Cost—Cost of the path from the bridge sending the configuration message to the root bridge.

Bridge ID—Priority and ID of the bridge sending the message.

Port ID—Port number (IEEE or Cisco STP BPDU) or the ring and bridge number (IBM STP BPDU) from which the configuration message was sent. This field allows loops created by multiple attached bridges to be detected and corrected.

Message Age—Indicates the amount of time that has elapsed since the root sent the configuration message on which the current configuration message is based.

Maximum Age—Indicates when the current configuration message should be deleted.

Hello Time—Indicates the time between root bridge configuration messages.

Forward Delay—Indicates the length of time that bridges should wait before transitioning to a new state after a topology change. If a bridge transitions too soon, it is possible that not all network links will be ready to change their state, and loops can result.

Catalyst 3920 Spanning-Tree Support

The Catalyst 3920 supports the following STPs:

- IEEE 802.1d
- IBM
- Cisco

The following sections briefly describe the type of transmission mode supported by each STP.

IEEE 802.1d STP

The IEEE STP can be used at the TrCRF or the TrBRF level. This type of spanning tree supports bridge domains and allows the bridge to construct a loop-free topology across an extended LAN. Specifically, the IEEE 802.1d STP supports the following bridge modes:

- Transparent Bridging
- Source-Route Switching
- Source-Route Transparent Bridging

The IEEE 802.1d STP BPDU format is:

Destination Address	Source Address	SAP	BPDU

Transparent Bridging

When a bridge connection is transparent mode:

- The bridge connection learns the source MAC addresses.
- Frames are forwarded based upon the destination address.

Source-Route Switching

When a bridge connection is source-route switching:

- The bridge connection learns route descriptors for frames that contain a RIF and learns the source MAC addresses for frames that do not contain a RIF.
- Source-route frames are forwarded based on the route descriptor.
- Non-source-route frames are forwarded based on the destination address.

Source-Route Transparent Bridging

When a bridge connection is source-route transparent:

- Transparent bridging and source-route bridging modes are combined.
- The bridge connection learns route descriptors for frames that contain a RIF and learns the source MAC addresses for frames that do not contain a RIF.
- Non-source-route frames are forwarded based on the destination address.
- Source-route frames are forwarded based on the route descriptor.
- All-routes explorer and spanning-tree explorer frames are issued and forwarded.
- The IEEE STP is used to eliminate loops for non-source-route and spanning-tree explorer frames.

IBM STP

The IBM STP can be used at the TrBRF level. This type of spanning tree was developed to manage the limited broadcast path through source-route bridges.

Source-Route Bridging

When a bridge connection is source-route:

- The bridge connection learns the source MAC address for frames that originate from the local ring and the route descriptor for frames that originate on the other side of a source-route bridge.
- Non-source-route frames are not forwarded.
- Source-route frames are forwarded based on the route descriptor.
- All-routes explorer and spanning-tree explorer frames are issued and forwarded.
- The IBM STP is used to eliminate loops only for spanning-tree explorer frames.

The IBM STP BPDU format is:

Destination Address	Source Address	SAP	BPDU

Cisco STP

The Cisco STP can be used at the TrCRF level. This type of spanning tree was developed to address a looping problem that can be introduced when you use VLANs in a Token Ring environment.

One of the rules in processing source-route traffic is that a source-route frame should never be forwarded to a ring that it has previously traversed. If the RIF of a source-route frame already contains the ring number for the next hop, the bridge assumes that the frame has already been on that ring and drops the frame.

With Token Ring VLANs, however, this rule can cause a problem. With the existing STP, a frame that originated on one physical ring of a Token Ring VLAN and is processed by an external SRT bridge would not be forwarded to another physical ring of the same Token Ring VLAN. Therefore, the IEEE 802.1d STP was used as a basis to create the Cisco STP. The Cisco STP ensures that traffic from one physical ring of a VLAN is not blocked from the other physical rings that comprise the VLAN.

Table A-1 summarizes the activities occurring in the TrCRF and TrBRF when the Cisco STP is run.

Table A-1 Cisco STP Summary

TrCRF Bridging Mode	TrCRF	TrBRF
SRB	• Runs the IEEE STP. • Processes IBM STP BPDUs from external bridges.	• Performs as a source-route bridge. • Runs the IBM STP to external bridges. • Drops transparent IEEE STP BPDUs of the TrCRF.
SRT	• Runs the Cisco STP. • Replaces bridge group address of destination address field with a Cisco-specific group address to prevent external bridges from analyzing TrCRF BPDUs. • Generates BPDUs with the Routing Information Identifier bit in the source address field set in the outbound frame and a 2-byte RIF added. This frame format ensures that the TrCRF remains local to the logical ring and is not transparently bridged or source routed to other LANs. Only TrCRFs connected via physical loops receive the BPDUs. • Processes IEEE STP BPDUs from external bridges.	• Performs as a source-route transparent bridge. • Runs the IEEE STP to external bridges. • Forwards transparent and source-route traffic. • Forwards source-route traffic to all other TrCRFs in the TrBRF whether they are in SRT or SRB mode.

The Cisco STP BPDU format is:

Destination Address	Source Address	RIF	SAP	BPDU

Spanning-Tree BPDU Formats Summary

For each BPDU format:

- The destination address is specified in the Bridge Group Address table.
- The source address is the base MAC address used by the switch.
- The SAP field should be set to 0x424203.

For the Cisco STP BPDU format, the source address must have the “msp masked” on to indicate the presence of a RIF in the header. The information carried in the RIF for the Cisco STP BPDU is a 2-byte field and must be set to 0x0200.