Understanding Token Ring Switching

This appendix discusses several aspects of Token Ring switching and how they relate to the Catalyst 3900. This appendix provides information on the following:

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 3900 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 3900 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 3900 can forward broadcast, multicast, and unicast frames based on MAC address. If, however, you have source-route bridges in your network, the Catalyst 3900 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 3900 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 3900. 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 3900 transfers nonbroadcast packets between ports without buffering the entire frame into memory. Instead, when a port on the Catalyst 3900 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 3900 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 3900 to reduce the average interstation latency.

In certain circumstances, however, the cut-through technique cannot be applied and the Catalyst 3900 must buffer frames into memory.

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

- The Catalyst 3900 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.

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 ISL trunk connections.

The Catalyst 3900 supports VTP Version 2, which includes provisions for the propagation of Token Ring-specific parameters associated with VLANs, such as hop count and bridge numbers.

On boot up, the Catalyst 3900 switch 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 3900 requests all VTP information from the neighbor.

The Catalyst 3900 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.

VTP Modes

The Catalyst 3900 switch can operate in three different VTP modes: server, client, or transparent.

- In server mode, the Catalyst 3900 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.

  When you make a change to the VLAN configuration on a VTP server, the change is propagated to all switches in the VTP domain. VTP advertisements are transmitted out all trunk connections, including ISL and LANE. You cannot create or modify VLANs on a VTP server until the management domain name is specified or learned.

- In client mode, the Catalyst 3900 accepts configuration changes from other devices in the administrative domain, but will not permit local changes to the database.

- In transparent mode, the Catalyst 3900 forwards on 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.

  If you create or modify VLANs on a switch that is in transparent mode, the changes affect only the individual switch.

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 Pruning

VTP pruning enhances network bandwidth use by reducing unnecessary flooded traffic, such as broadcast, multicast, unknown, and flooded unicast packets. VTP pruning increases available bandwidth by restricting flooded traffic to those ISL trunk links that the traffic must use to access the appropriate network devices. By default, VTP pruning is disabled.

Note Although VTP pruning can be enabled on a switch that is in VTP Server or Transparent mode, only switches that are in VTP Server or Client mode can participate in VTP pruning. VTP Clients, while they can participate in VTP pruning, cannot alter the pruning mode for the administrative domain.

Note Make sure that all devices in the administrative domain support VTP pruning before you enable it.

Figure A-1 shows a switched network without VTP pruning enabled. Port 1 on Switch 1 and port 2 on Switch 4 are assigned to the VLAN 200. A broadcast is sent from the host connected to Switch 1. Switch 1 floods the broadcast and every switch in the network receives it, even though Switches 3, 5, and 6 have no ports in the VLAN 200.

Figure A-1 Flooding Traffic without VTP Pruning

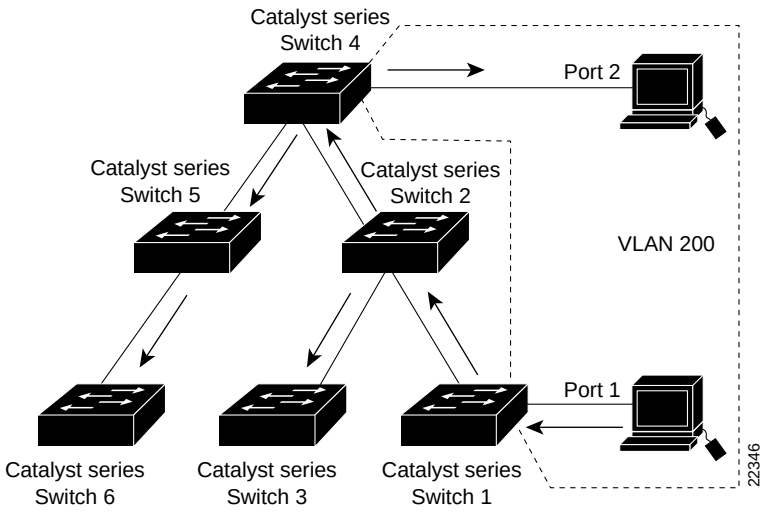
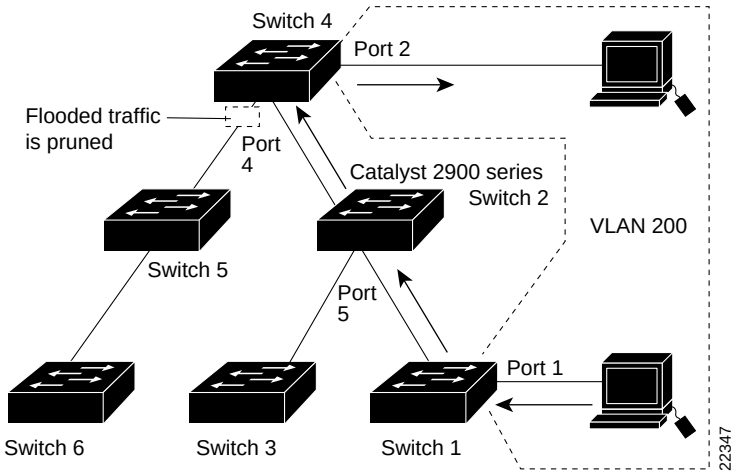


Figure A-2 shows the same switched network with VTP pruning enabled. The broadcast traffic from Switch 1 is not forwarded to Switches 3, 5, and 6 because traffic for the VLAN 200 has been pruned on the links indicated (port 5 on Switch 2 and port 4 on Switch 4).

Figure A-2 Flooding Traffic with VTP Pruning



Enabling VTP pruning on a VTP server enables pruning for the entire administrative domain. VTP pruning takes effect several seconds after you enable it. By default, VLANs 2 through 1000 are pruning-eligible. VTP pruning does not prune traffic from VLANs that are pruning-ineligible. VLAN 1, the default TrCRF (1003), the default TrBRF (1005), and TrCRFs are always pruning-ineligible, therefore traffic from these VLANs cannot be pruned.

VTP Start Up

When a Catalyst 3900 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.

Token Ring VLANs

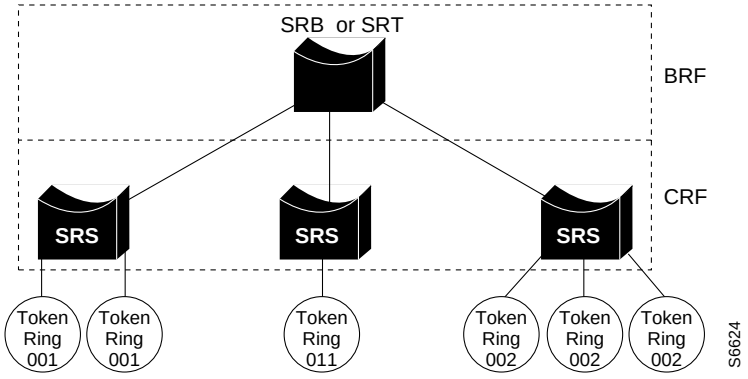
Within a Token Ring VLAN, distributed rings can be formed by defining groups of ports that have the same ring number. The IEEE calls such a port group a Token Ring Concentrator Relay Function (TrCRF). A TrCRF is limited to the ports in a single Catalyst 3900 or those within a stack of Catalyst 3900s.

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.

As shown in Figure A-3, 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.

Figure A-3 Token Ring VLANs



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.

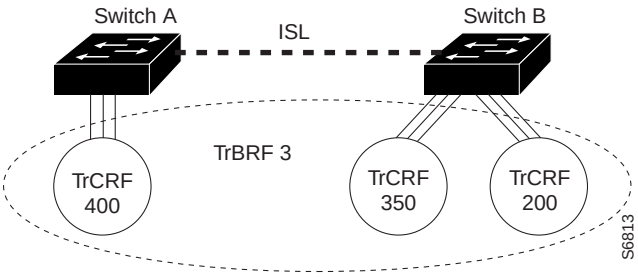
Token Ring VLANs and ISL

The TrBRF can be extended across a network of switches via high-speed uplinks between the switches. If you have an ISL module installed in your Catalyst 3900 Token Ring switch, the following types of TrCRFs can exist in your network: undistributed, backup, and default.

Undistributed TrCRFs

The *undistributed* TrCRF is the standard type of TrCRF in the Catalyst 3900 switch. The undistributed TrCRF is located on one switch and has a logical ring number associated with it. Multiple undistributed TrCRFs located on the same or separate switches can be associated with a single parent TrBRF. The parent TrBRF acts as a multiport bridge, forwarding traffic between the undistributed TrCRFs. Figure A-4 illustrates the undistributed TrCRF.

Figure A-4 Undistributed TrCRFs



Backup TrCRFs

The *backup* TrCRF enables you to configure an alternate route for traffic between undistributed TrCRFs located on separate switches that are connected by a TrBRF, in case the ISL connection between the switches becomes inactive.

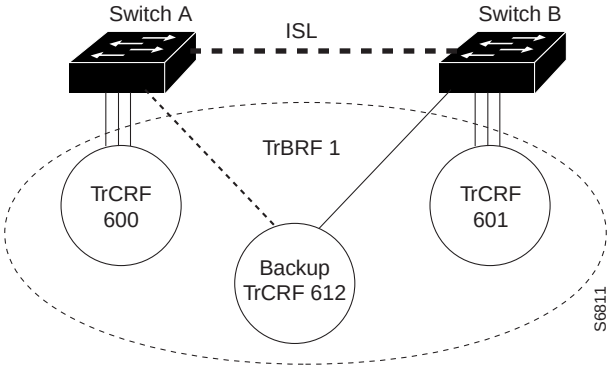
While a TrBRF can contain multiple TrCRFs, it can contain only *one* TrCRF that is configured as a backup TrCRF. That backup TrCRF can contain only *one* port from each related switch. If, however, you have more than one TrBRF defined on a switch, you can have more than one backup TrCRF defined on a switch; one defined for each TrBRF.

To create a backup TrCRF, create the TrCRF, assign it to the TrBRF that traverses the switches, mark it as a backup TrCRF, and then assign one port on each switch to the backup TrCRF.

Caution If the backup TrCRF port is attached to a Token Ring MAU, it will not provide a backup path unless the ring speed and port mode are set by another device. Therefore, it is recommended that you manually configure the ring speed and port mode for the port assigned to the backup TrCRF.

Under normal circumstances, only one port in the backup TrCRF is active. The active port is the port with the lowest MAC address. If the ISL connection between the switches become inactive, the port that is a part of the backup TrCRF on each affected switch will automatically become active, rerouting traffic between the undistributed TrCRFs through the backup TrCRF. When the ISL connection is reestablished, all but one port in the backup TrCRF will be disabled. Figure A-5 illustrates the backup TrCRF.

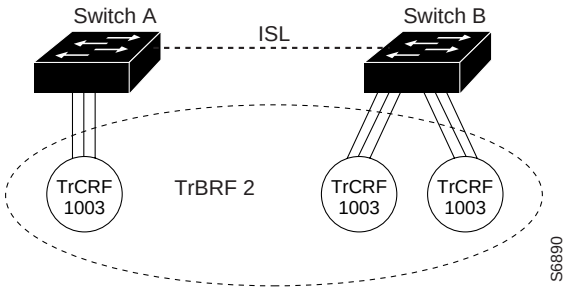
Figure A-5 Backup TrCRF



Default TrCRF

As a rule, TrCRFs cannot span different switches. There is one exception; the default TrCRF (1003). The default TrCRF can contains ports that are located on multiple switches. It is associated with the default TrBRF (1005), which can span switches via ISL. As the default TrCRF is the only TrCRF that can be associated with the default TrBRF, the default TrBRF does not perform any bridging functions, but simply uses source-route switching to forward traffic between the ports of the TrCRF. Figure A-6 illustrates the default TrCRF.

Figure A-6 Default TrCRF



Understanding ATM

ATM is a hardware and software architecture that switches small units of data called *cells*. The latency in a cell switch is very small because of the short cell size. Short cells have a tiny store-and-forward delay. In the absence of port contention and buffering, cells are switched quickly in the hardware. In addition to the low latency, ATM is beneficial to large networks because it:

- Is a multiplexing and switching technology that is designed for flexibility and performance.
- Supports Quality of Service (QOS) options for flexibility and high bandwidth options (up to Gigabits per second) for performance.
- Offers switched virtual circuits (SVCs), which are automatically set up and torn down when data needs to be transferred.

- Supports environments where applications with different performance requirements need to be executed on the same computer, multiplexer, router, switch and network. The flexibility of ATM means that voice, video, data, and future payloads can be transported.
- Has worldwide support. The ATM Forum, an industry forum made up of many companies (including Cisco), works with formal standards bodies to specify ATM.

ATM Cell

The components of the 53-byte ATM cell are the following:

- Generic flow control. Intended to be used for controlling user access and flow control. This field is used when passing ATM traffic through a user-network (UNI) interface to alleviate short-term overload conditions. A network-to-network (NNI) interface does not use this field for GFC purposes; rather, an NNI uses this field to define a larger VPI value for trunking purposes.
- Virtual path identifier (VPI). Identifies the route (path) to be taken by the ATM cell. In an idle or null cell, the VPI field is set to all zeros. (A cell containing no information in the payload field is either “idle” or “null”). A virtual path connection (VPC) is a group of virtual connections between two points in the network. Each virtual connection may involve several ATM links. VPIs provide a way to bundle ATM traffic being sent to the same destination.
- Virtual channel identifier (VCI). Identifies the circuit or connection number on that path. In an idle or null cell (one containing no payload information), the VCI field is set to all zeros. Other non-zero values in this field are reserved for special purposes. For example, the values VPI=0 and VCI=5 are used exclusively for ATM signaling purposes when requesting an ATM connection. A VCC is a connection between two communicating ATM entities; the connection may consist of a concatenation of many ATM links.
- Payload type identifier (PTI). Indicates the type of data being carried in the payload. The first bit is a 0 if the payload contains user information and is a 1 if it carries connection management information. The second bit indicates if the cell experienced congestion over a path. If the payload is user information, the third bit indicates if the information is from customer premises equipment (CPE).
- Cell loss priority (CLP). 1-bit descriptor in the ATM cell header is set by the ATM adaptation layer (AAL) to indicate the relative importance of a cell. This bit is set to 1 to indicate that a cell can be discarded, if necessary, such as when an ATM switch is experiencing traffic congestion. If a cell should not be discarded, such as when supporting a specified or guaranteed QOS, this bit is set to 0. This bit may also be set by the ATM layer if an ATM connection exceeds the QOS parameters established during connection setup.
- Header error control (HEC). 8-bit cyclic redundancy check (CRC) computed on all fields in an ATM UNI/NNI cell header. The HEC is capable of detecting all single-bit errors and certain multiple-bit errors. This field provides protection against incorrect message delivery caused by addressing errors. However, it provides no error protection for the ATM cell payload proper. The physical layer uses this field for cell delineation functions during data transport.
- Payload. Maximum of 48 bytes. There is no error control for the payload.

PVC

With a PVC, everything is statically configured and no signaling is involved. The PVC is mapped to a network in a subinterface point-to-point configuration. The logical data link layer can use Subnetwork Access Protocol (SNAP) encapsulation (as defined in RFC 1483). This encapsulation allows multiple protocols to be multiplexed over one PVC. Alternately, the logical data link layer can use LAN emulation (LANE) Version 1 over PVC.

The PVC is statically mapped at each ATM node. The path of the PVC is identified at each switch by an incoming VCI and VPI and an outgoing VCI and VPI.

Note The Catalyst 3900 does not support PVC configuration.

SVC

Establishing an ATM SVC involves an agreement between the end nodes and all the switches in between. Each end node has a special signaling channel to the connected switch called the UNI. Switches have a signaling channel between them called the NNI. Cells that arrive on the signaling channel are reassembled into frames in the reliable Service-Specific Connection Oriented Protocol (SSCOP). The signaling information follows the Q.2931 standard.

Establishing an SVC potentially involves signaling between the following:

- Router and a private ATM switch (private UNI)
- Router and a public ATM switch (public UNI)
- Private ATM switch and a public ATM switch (public UNI)

The UNI is defined by the ATM Forum UNI specification.

Interfaces to public ATM networks are identified by an E.164 address. Interfaces to private ATM networks are identified by a network service access point (NSAP) address. These addresses are contained in different fields of the same 20-octet address.

Once an SVC is established, it functions like a PVC. SVCs can be used in point-to-point subinterface configuration or point-to-multipoint nonbroadcast multiaccess (NBMA) configuration.

ATM Adaptation Layers

The purpose of the AAL is to receive the data from the various sources or applications and convert or adapt it to 48-byte segments that will fit into the payload of an ATM cell. Because ATM benefits from its ability to accommodate data from various sources with differing characteristics, the adaptation layer must be flexible.

There are four classes of traffic supported by the adaptation layer:

- Class A is supported by AAL1 and is typically used for servers such as DS1 or DS3 circuit emulation.
- Class B is supported by AAL2 and is typically used for compressed voice and video.
- Class C is supported by AAL3/4 and is used for transmitting VBR traffic and SDMS.
- Class D is supported by AAL5 and also supports VBR traffic but with minimal overhead.

Because ATM is inherently a connection-oriented transport mechanism and because the current applications of ATM are heavily oriented toward LAN traffic, many of the current ATM products, including the Catalyst 3900, support the Class D adaptation layer with AAL5.

Components of an ATM Network

The building blocks of an ATM internetwork may consist of the following:

- Routers with ATM interfaces
- Computers with a native ATM Network Interface Card (NIC)
- LightStream 1010 or other ATM switches
- ATM physical layer, supporting SONET OC-3 with single or multimode fiber, TAXI with multimode fiber, or DS3/E3 with coaxial cable
- LAN switches with ATM interfaces

ATM and VLANs

The ATM expansion module supports up to 63 VLANs (or ELANs). Each VLAN corresponds to an ELAN. Each association between the ATM expansion module and a VLAN creates a virtual ATM port. A virtual ATM port is the equivalent of an LEC.

Understanding LAN Emulation

LANE makes the ATM network transparent to LAN traffic by mapping connectionless LAN traffic over the connection-oriented ATM network. It uses point-to-multipoint connections to service the connectionless broadcast service that is required by LAN protocols.

Cisco's Token Ring implementation of LANE makes an ATM interface look like one or more Token Ring interfaces; the ELAN looks like a ring. Setting up LECs allows the Catalyst 3900 to operate in a LAN environment containing ATM devices, such as Cisco 7000 or 4500 series routers with an ATM interface or ATM port adapter connected to a LightStream 1010 ATM switch.

The Catalyst 3900 supports the LANE standard as defined by the ATM Forum specification *LAN Emulation over ATM Version 1.0*, ATM_FORUM 94-0035. This service emulates the following LAN-specific characteristics:

- Connectionless services
- Multicast services
- LAN Media Access Control (MAC) driver services

LANE service provides connectivity between ATM-attached devices and LAN-attached devices. This includes connectivity between ATM-attached stations and LAN-attached stations as well as connectivity between LAN-attached stations across an ATM network.

Because LANE connectivity is defined at the MAC layer, upper protocol layer functions of LAN applications can continue unchanged when the devices join ELANs. This feature protects corporate investments in legacy LAN applications.

An ATM network can support multiple independent ELANs. Membership of an end system in any of the ELANs is independent of the physical location of the end system.

Components of LANE

Up to 256 ELANs can be set up in an ATM switch cloud. A Catalyst 3900 ATM module can participate in up to 63 of these ELANs.

LANE is defined on a client-server LAN model, as follows:

- LANE client (LEC)

  An LEC emulates a LAN interface to higher layer protocols and applications. It forwards data to other LANE components and performs LANE address resolution functions.

  Each LEC is a member of only one ELAN. However, a router or a Catalyst 3900 ATM module can include LECs for multiple ELANs: one LEC for each ELAN of which it is a member.

  If a router has clients for multiple ELANs, the router can route traffic between the ELANs.

  Note If the Catalyst 3900 has multiple ATM modules and each has a client that active for the same ELAN, the Catalyst 3900 will not bridge between the ELANs on the different modules. The Catalyst 3900 acts as an edge device on an ATM cloud.

- LANE server (LES)

  The LANE server for an ELAN is the control center. It provides joining, address resolution, and address registration services to the LANE clients in that ELAN. Clients can register destination unicast and multicast MAC addresses with the LANE server. The LANE server also handles LANE ARP (LE-ARP) requests and responses.

  The current implementation has a limit of one LANE server per ELAN.

- LANE broadcast and unknown server (BUS)

  The LANE BUS sequences and distributes multicast and broadcast packets and handles unicast flooding.

  One combined LES and BUS is required per ELAN.

- LANE configuration server (LECS)

  The LANE configuration server contains the database that determines which ELAN a device belongs to (each configuration server can have a different named database). Each LEC contacts the LECS once, when it joins an ELAN, to determine which ELAN it should join. The LECS returns the ATM address of the LES for that ELAN.

  One LECS is required per ATM LANE switch cloud.

  The LECS database can have the following four types of entries:

  — ELAN name, ATM address of LANE server pairs

  — LANE client MAC address, ELAN name pairs

  — LANE client ATM template, ELAN name pairs

  — Default ELAN name

  Note ELAN names must be unique on an interface. If two interfaces participate in LANE, the second interface may be in a different switch cloud.

- Simple Server Redundancy Protocol (SSRP)

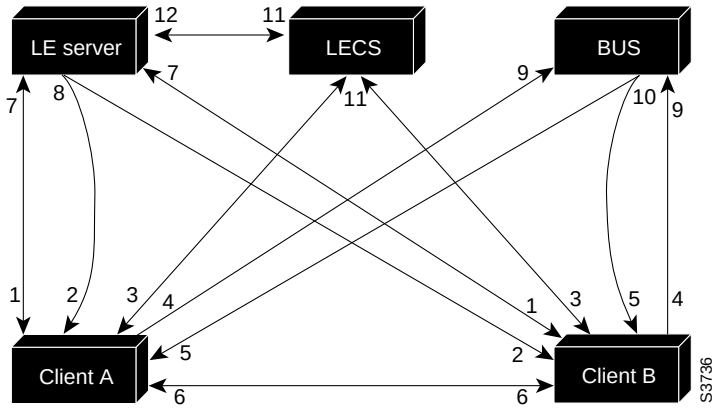
  The LANE simple server redundancy feature creates fault tolerance using standard LANE protocols and mechanisms. If a failure occurs on the LANE configuration server or on the LANE server/broadcast-and-unknown server, the ELAN can continue to operate using the services of a backup LANE server.

The Catalyst 3900 ATM module currently supports only the LEC function. A Cisco 7000, Cisco 7200, Cisco 7500, RSP 7000, Cisco 4500, Cisco 4700, or Catalyst 5000 with an ATM Interface Processor (AIP) can supply all LANE functions.

LANE Operation and Communication

Communication among LANE components is typically handled by several types of SVCs. Some SVCs are unidirectional; others are bidirectional. Some are point-to-point and others are point-to-multipoint. Figure A-7 illustrates the various types of SVCs.

Figure A-7 LANE VCC Types



1–7	Control direct	4–9	Multicast send
2–8	Control distribute	5–10	Multicast forward
3–11	Configure direct (client)	6–6	Data direct
		11–12	Configure direct (server)

The following section describes various processes that occur, starting with a client requesting to join an ELAN.

Join Process

The following process (illustrated in Figure A-7) normally occurs after an LEC has been enabled on the ATM module in a Catalyst 3900:

Step 1 The client requests to join an ELAN. The client sets up a connection to the LECS to find the ATM address of the LANE server for its ELAN. See the bidirectional, point-to-point link (link 1–7 in Figure A-7).

An LEC finds the LECS using the following methods in the listed order:

- Locally configured ATM address
- Interim Local Management Interface (ILMI)
- Well-known address defined by the ATM Forum

Step 2 The LECS identifies the LES. Using the same VCC, the LECS returns the ATM address and the name of the LES for the client's ELAN.

Step 3 The client tears down the configure direct VCC.

Step 4 The client contacts the server for its LAN. The client sets up a connection to the LES for its ELAN (bidirectional, point-to-point control direct VCC [link 1–7 in Figure A-7]) to exchange control traffic. Once a Control Direct VCC is established between an LEC and LES, it remains up.

Step 5 The LES verifies that the client is allowed to join the ELAN. The server for the ELAN sets up a connection to the LECS to verify that the client is allowed to join the ELAN (bidirectional, point-to-point server configure VCC [link 11–12 in Figure A-7]).

The server's configuration request contains the client's MAC address, its ATM address, and the name of the ELAN. The LECS checks its database to determine whether the client can join that LAN; then it uses the same VCC to inform the server whether or not the client is allowed to join.

Step 6 The LES allows or disallows the client to join the ELAN. If allowed, the LES adds the LEC to the unidirectional, point-to-multipoint control distribute VCC (link 2–8 in Figure A-7) and confirms the join over the bidirectional, point-to-point control direct VCC (link 1–7 in Figure A-7). If disallowed, the LES rejects the join over the bidirectional, point-to-point control direct VCC (link 1–7 in Figure A-7).

Step 7 The LEC sends LAN emulation address resolution protocol (LE-ARP) packets for the broadcast address, which is all 1s. Sending LE-ARP packets for the broadcast address returns the ATM address of the BUS. Then the client sets up the multicast send VCC (link 4–9 in Figure A-7) and the BUS adds the client to the multicast forward VCC (link 5–10 in Figure A-7) to and from the broadcast-and-unknown server.

Addressing

On a LAN, packets are addressed by the MAC-layer address of the destination and the source stations. To provide similar functionality for LANE, MAC-layer addressing must be supported. Every LEC must have a MAC address. In addition, every LANE component (server, client, BUS, and configuration server) must have a unique ATM address.

All LECs on the same interface have a different, automatically assigned MAC address. That MAC address is also used as the end-system identifier part of the ATM address, as explained in the following section.

LANE ATM Addresses

A LANE ATM address has the same syntax as an NSAP rather than E-164. It consists of the following:

- A 13-byte prefix that includes the following fields defined by the ATM Forum: authority and format identifier (AFI) field (1 byte); Data Country Code (DCC) or International Code Designator (ICD) field (2 bytes); Domain Specific Part Format Identifier (DFI) field (1 byte); Administrative Authority (3 bytes); Reserved (2 bytes); Routing Domain (2 bytes); and Area (2 bytes).
- A 6-byte end-system identifier (MAC).
- A 1-byte selector field.

ILMI Address Registration

The Catalyst 3900 ATM module uses ILMI registration to build its ATM address and to register this address with the ATM switch. To build its ATM address, the Catalyst 3900 obtains its ATM address prefix from the ATM switch. Then it combines the ATM address prefix with its own MAC address and the selector value. As an alternative, you can hard-code the address. Once the Catalyst ATM module has determined its ATM address, it uses ILMI to register this address with the ATM switch.

Address Resolution

As communication occurs on the ELAN, each client dynamically builds a local LE-ARP table. The LE-ARP table maps ELAN MAC addresses (Layer 2) to ATM addresses (also Layer 2). A client's LE-ARP table can also have static, preconfigured entries. The LE-ARP table maps MAC addresses to ATM addresses.

When a client first joins an ELAN, its LE-ARP table has no dynamic entries and the client has no information about destinations on or beyond its ELAN. To learn about a destination when a packet is to be sent, the client begins the following process to find the ATM address corresponding to the known MAC address:

Step 1 The client sends an LE-ARP request to the LANE server for this ELAN (point-to-point control direct VCC [link 1–7 in Figure A-7]).

Step 2 If the MAC address is registered with the server, it returns the corresponding ATM address. If not, the LES forwards the LE-ARP request to all clients on the ELAN (point-to-multipoint control distribute VCC [link 2–8 in Figure A-7]).

Step 3 Any client that recognizes the MAC address responds with its ATM address (point-to-point control direct VCC [link 1–7 in Figure A-7]).

Step 4 The LES forwards the response (point-to-multipoint control distribute VCC [link 2–8 in Figure A-7]).

Step 5 The client adds the MAC address-ATM address pair to its LE-ARP cache.

Step 6 Now the client can establish a VCC to the desired destination and proceed to transmit packets to that ATM address (bidirectional, point-to-point data direct VCC [link 6–6 in Figure A-7]).

For unknown destinations, the client sends a packet to the BUS, which forwards the packet to all clients in the ELAN. The BUS floods the packet because the destination might be behind a bridge that has not yet learned this particular address.

Traffic Handling

The Catalyst 3900 allows you to define up to 64 traffic profiles. These profiles can be used to define the maximum rates for each traffic type. The traffic profiles are defined on the Traffic Profile Table Setup panel. Refer to “Configuring Traffic Profiles Tables” section on page 6-44 of the “Configuring the Catalyst 3900”chapter for more information.

For each VLAN (or ELAN), a traffic profile must be mapped to each DD-VCCs. The process of mapping depends on whether the traffic is incoming or outgoing:

- For incoming calls, the LEC tries to find a traffic profile that best matches the traffic parameters in the call. You can define the maximum discrepancy between the specified parameters and actual values on a per ELAN basis on the ELAN Configuration panel. Refer to “Configuring LEC Parameters” section on page 6-40 of the “Configuring the Catalyst 3900” chapter for more information.
- For outgoing calls, you can define up to 10 profiles to use. The destination ATM address is ANDed with the address mask. The resulting ATM address is compared with the ATM address in the mapping table. If there is a match, each defined profile (0 through 9) is used in sequence until a call SETUP is accepted by the destination node. You define the address, address mask, and profiles to be used on the Traffic Profile Mapping panel. Refer to “Configuring Traffic Profile Mapping” section on page 6-42 of the “Configuring the Catalyst 3900”chapter for more information.

Multicast Traffic

When an LEC has broadcast or multicast traffic, or unicast traffic with an unknown address to send, the following process occurs:

Step 1 The client sends the packet to the BUS (unidirectional, point-to-point multicast send VCC [link 4–9 in Figure A-7]).

Step 2 The BUS forwards (floods) the packet to all clients (unidirectional, point-to-multipoint multicast forward VCC [link 5–10 in Figure A-7]).

This VCC branches at each ATM switch. The switch forwards these packets to multiple outputs. (The switch does not examine the MAC addresses; it simply forwards all packets it receives.)

Understanding ISL

The Catalyst family of switches provides a means of multiplexing VLANs between switches and routers using ISL on Fast Ethernet or LAN emulation on ATM. You can use any combination of these trunk technologies to form enterprise-wide VLANs.

A *trunk* is a physical link that carries the traffic of multiple VLANs between two switches or between a switch and a router, thereby allowing the VLANs to be extended across switches. Trunks use high-speed interfaces such as Fast Ethernet, Fiber Data Distributed Interface (FDDI), or ATM.

ISL was originally developed for Ethernet switches. It uses a Fast Ethernet interface to provide connectivity between switches and extends the VLAN capabilities of the switch by tagging the standard Fast Ethernet frame with the necessary VLAN information. Like ATM, ISL can provide a high-speed link between switches. Unlike ATM, however, ISL forwards the data across the high-speed link without breaking the frames into cells. The frame is sent intact across the ISL connection.

The Token Ring implementation of ISL encapsulates Token Ring frames in Fast Ethernet frames.

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 “VLAN Trunking Protocol” section on page A-4, 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.

Note If you have redundant ISL links in your switch configuration, you must enable STP at the TrCRF level for the default TrCRF (1003) and enable STP at the TrBRF level for all other TrCRFs.

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-8 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-8 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 3900 Spanning-Tree Support

The Catalyst 3900 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.

Duplicate Ring Protocol

The Cisco Duplicate Ring Protocol (DRiP) runs on Cisco devices that support switched VLAN networking and is used to identify active VLANs and help prevent the configuration of duplicate rings (TrCRFs) across switches.

Through packet advertisements, DRiP maintains the status of TrCRFs. It then uses this information to determine whether there are multiple TrCRFs active in a TrBRF.

DRiP information is used for the following:

- Enables the switch to filter out excessive All-Routes Explorer (ARE) frames. The DRiP information is used in conjunction with the local configuration to determine which of the TrCRFs configured within a TrBRF have active ports. This information is used by Token Ring and ISL ports to discard AREs that have already been on an attached ring.
- Detects the configuration of duplicate TrCRFs across switches, which would cause a TrCRF to be distributed across ISL trunks. The DRiP information is used in conjunction with the local configuration information to determine which TrCRFs are already active on the switch. If DRiP determines that a TrCRF is configured on more than one switch, it will disable the ports associated with the TrCRF. In this case, the status of the port displayed on Port Configuration panel will be “Disabled by DRiP”.
- Detects the failure of an ISL path and enables a backup path.