C H A P T E R **4**

Implementation of the Cell/Area Zone

This chapter outlines recommendations, best practices, and configurations for implementing a cell/area zone architecture in an EttF environment. The cell/area zone is where actual end nodes connect into the network, so careful planning must be done to achieve the optimal design from both the network and device perspective.

As mentioned in earlier chapters, EtherNet/IP is the enabling standard at this layer. Ethernet networks have been successfully used on the factory floor for the past 15 years, mainly in non-time-critical applications. Ethernet technology (more accurately IEEE standard 802.1 and IEEE standard 802.3 technologies) has evolved from a 10 Mbps, half-duplex, bus/tree topology into a 100 Mbps and 1 Gbps, full-duplex, switch/router-based hierarchical star topology. This evolution has created an opportunity for using Ethernet in industrial networks that support time-critical applications. EtherNet/IP is a communication system suitable for use in industrial environments and time-critical applications. EtherNet/IP uses standard Ethernet and TCP/IP technologies and an open Application Layer protocol called Control and Information Protocol (CIP). CIP is also used in ControlNet and DeviceNet networks. In EtherNet/IP networks, exchange of time-critical data is based on the producer/consumer model where a transmitting device (host or end node) produces data on the network, and many receiving devices can consume this data simultaneously. Implementation of the producer/consumer data exchange is based on the IP multicast service mapped over the Ethernet multicast service. EtherNet/IP-supported functions include the following:

- Time-critical data exchange
- Human-machine interface (HMI)
- Device configuration and programming
- Remote access to web pages embedded in EtherNet/IP devices
- Device and network diagnostics

The configuration details outlined below (for example, VLAN numbers, hostnames, port numbers, and so on) are merely examples and should be adjusted accordingly to a particular factory environment.

Cell/Area Zone Network Device Provisioning

Networking devices in the cell/area zone include Cisco Catalyst 2955s and the downlinks on the Catalyst 3750. The recommended Cisco IOS Software version at the time of this writing is C2955-12.1.22-EA9 (Crypto Image) and C3750-12.2.25-SEB4 (Advanced Crypto Image). The images can be downloaded from the following URL: http://www.cisco.com/kobayashi/sw-center/index.shtml.

Beginning with the Catalyst 2955s, the startup process is as follows:

Step 1   Load the Cisco IOS image on all devices.

Step 2   Configure all uplink 1 GE ports (gi0/1–2) as trunk ports carrying only one VLAN:

```
interface GigabitEthernet0/1
switchport trunk native vlan 20
switchport trunk allowed vlan 20
switchport mode trunk
end
```

Step 3   Configure all FastEthernet (fa0/1–fa0/12) interfaces as switchport access ports for the particular VLAN carried on the uplink trunk:

```
!
interface FastEthernet0/1
 switchport access vlan 20
 switchport mode access
 spanning-tree portfast
 spanning-tree bpdufilter enable
 spanning-tree bpduguard enable
 ip dhcp snooping limit rate 10
end
```

Step 4   On FastEthernet ports connected to an end device, manually configure speed and duplex settings to those supported by the end device:

```
cell-c2955-12#conf t
Enter configuration commands, one per line.  End with CNTL/Z.
cell-c2955-12(config)#int f0/1
cell-c2955-12(config-if)#speed 100
cell-c2955-12(config-if)#duplex full
```

The end device must also be configured to match the settings from above.

Step 5   Enable broadcast suppression filters on all Gigabit Ethernet uplinks to help prevent broadcast floods in case of a misconfiguration or a rogue device:

```
Switch# configure terminal
Switch(config)# interface gigabitethernet0/1
Switch(config-if)# storm-control broadcast level 20
```

Step 6   Shut down all interfaces that are not in use.

For the Catalyst 3750, the startup process is as follows:

Step 1   Load the correct Cisco IOS image on both devices in the stack,

Step 2   Configure downlink GE ports (gi1/0/14 and gi2/0/14) as trunk ports carrying the one VLAN configured for the Catalyst 2955s. Note that each of these is on a different switch in the stack.

```
!
interface GigabitEthernet1/0/14
 switchport trunk encapsulation dot1q
 switchport trunk native vlan 999
 switchport trunk allowed vlan 20
 switchport mode trunk

interface GigabitEthernet2/0/14
 switchport trunk encapsulation dot1q
 switchport trunk native vlan 999
 switchport trunk allowed vlan 20
 switchport mode trunk
```

Step 3   Configure the L3 SVIs where the VLANs will terminate. This will also be the IGMP querier interface needed for IGMP snooping.

```
!
interface Vlan20
 ip address 10.17.20.1 255.255.255.0
 ip pim sparse-dense-mode
end
!

interface Vlan30
 ip address 10.17.30.1 255.255.255.0
 ip pim sparse-dense-mode
end
```

Step 4   Enable broadcast suppression filters on all Gigabit Ethernet uplinks to help prevent broadcast floods in case of a misconfiguration or rogue device:

```
Switch# configure terminal
Switch(config)# interface range gigabitethernet1/0/14 , gigabitethernet2/0/14
Switch(config-if)# storm-control broadcast level 20
```

**Step 5**   Shut down all interfaces that are not in use.

Virtual LAN Segmentation

VLAN Overview

A virtual LAN (VLAN) is a switched network segmented on a functional, application, or organizational basis as opposed to a physical or geographical basis. Switches filter destination MAC addresses and forward VLAN frames only to ports that serve the VLAN to which the traffic belongs. A VLAN consists of several end systems, either hosts or network equipment (such as switches and routers), all of which are members of a single logical broadcast domain. A VLAN no longer has physical proximity constraints

for the broadcast domain. This VLAN is supported on various pieces of network equipment (for example, LAN switches) that support VLAN trunking protocols between them. Each VLAN supports a separate spanning tree (IEEE 802.1d).

A VLAN can span multiple switches such that in the topology shown in Figure 4-1, PAC 1 is controlling I/Os 1, 2, and 3 on Switch 1; and PAC 2 is controlling I/Os 4 and 5 on Switch 2 on the same VLAN. In this case, all devices on this VLAN are on the same broadcast domain.
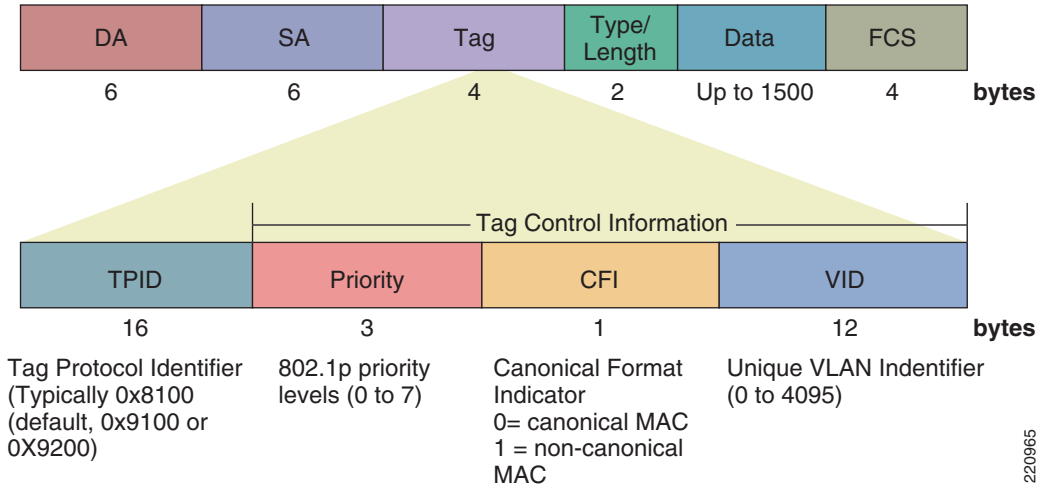
Figure 4-1   ***VLAN Spanning Multiple Switches***
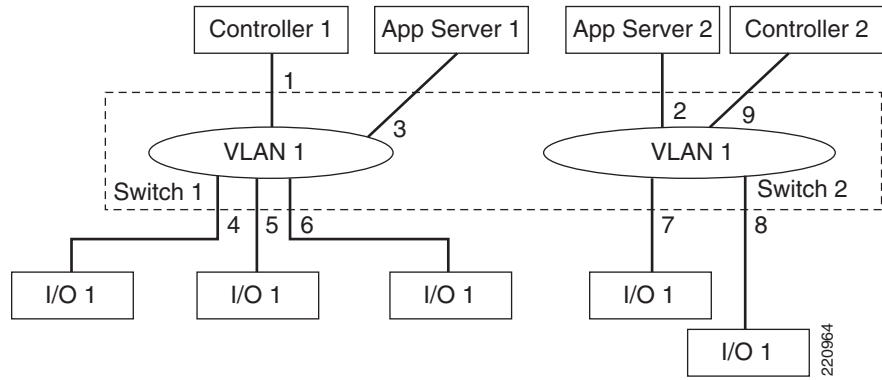
VLAN Details

A VLAN is created by inserting a four-byte VLAN header into the basic Ethernet frame between the source address and length/type fields, as shown in Figure 4-2.

Figure 4-2   ***VLAN Spanning Multiple Switches***

VLANs In the Cell/Area Zone

Cell/area zone devices on the factory floor should include only traffic (application, consumer/producer) that is relevant to running that particular cell. For this reason, EttF 1.1 recommends logically segmenting traffic with the use of VLANs. As shown in Figure 4-3, only one VLAN is recommended for all data traffic relevant to that particular cell/area zone. Because 80–90 percent of traffic is local to one cell, this is the optimal design. Furthermore, producer/consumer traffic is technically confined to a single VLAN because of a feature in Rockwell Automation device firmware. The Catalyst 3750 aggregates all VLANs in the cell/area zone and terminates them with L3 switched virtual interfaces (SVIs).

Figure 4-3   *VLANs in the Cell/Area Zone*
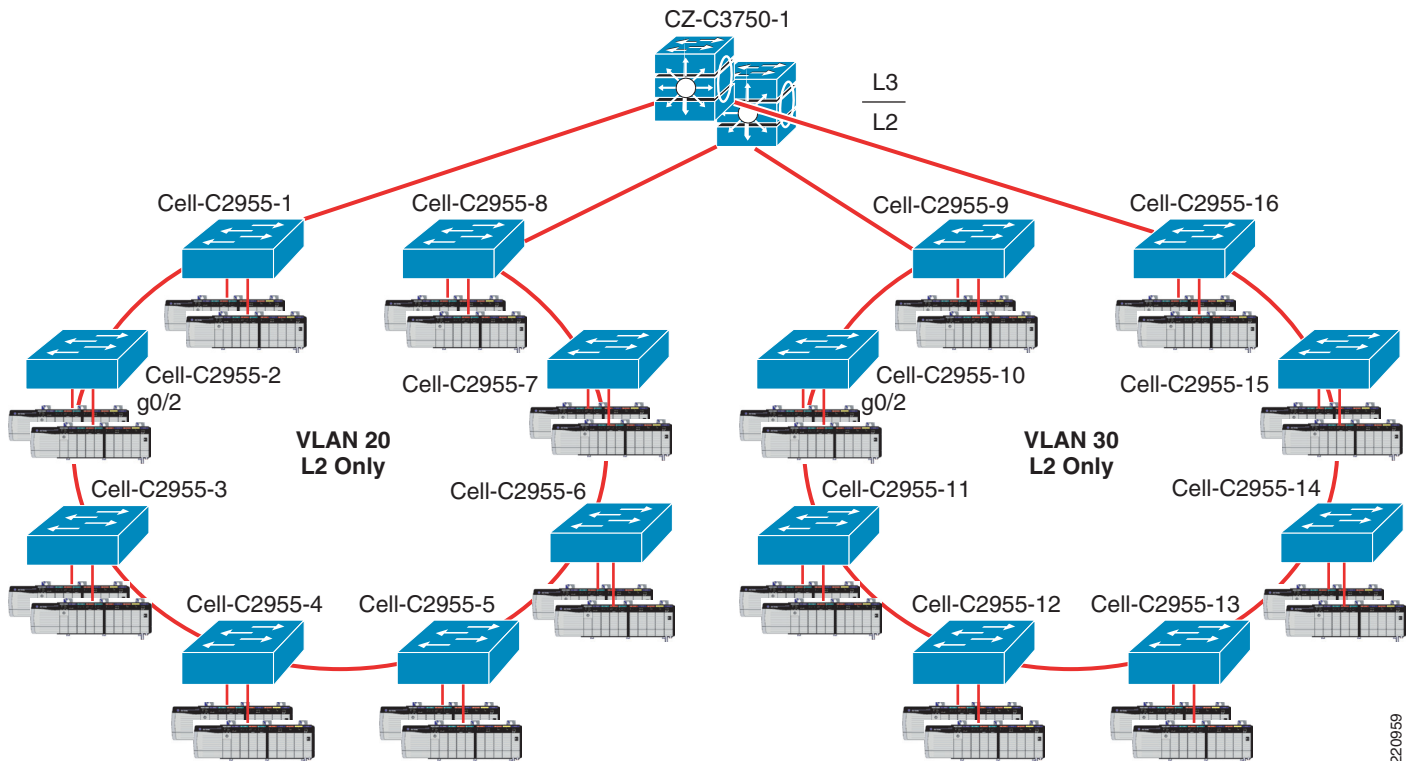
VLAN Highlights of Ring Topology

Following are VLAN highlights of the ring topology:

- All downward-facing FastEthernet (100 Mbps) ports connected to devices are configured as access ports for a single VLAN.

- All uplinks (GigEthernet, 1000 Mbps) ports are connected as dot1q trunks carrying only the VLAN defined above.

- At the top of the ring, the Catalyst 3750 terminates all VLANs in the ring below with L3 SVIs configured.

- The Catalyst 3750 provides inter-VLAN routing functionality.

VLAN Recommendations

The following are VLAN recommendations for EttF phase 1.1:

- Use one VLAN per ring topology for all manufacturing traffic per cell/area zone.

- If network traffic in one ring is consistently above 2500 packets per second (pps), consider dividing communicating PAC, I/O, and HMI groups into other VLANs.

- If non-manufacturing traffic (PC, and so on) must exist in the ring, it should be on a separate VLAN.

- Remove VLAN 1 from trunk ports and assign a new native VLAN. (See Spanning Tree Protocol Design, page 4-7.)

- Configure VTP Mode as "transparent" to avoid operational error because very few VLANs are used.

VLAN Benefits for EttF

In a flat, bridged network, all broadcast packets generated by any device in the network are sent to and received by all other network nodes. The ambient level of broadcasts generated by the higher layer protocols in the network, known as *broadcast radiation*, typically restricts the total number of nodes that the network can support. In extreme cases, the effects of broadcast radiation can be so severe that an end station spends all its CPU power on processing broadcasts.

VLANs have been designed to address the following problems inherent in a flat, bridged network:

- Scalability issues of a flat network topology

- Simplification of network management by facilitating network reconfigurations

VLANs offer the following features:

- Broadcast control—Just as switches isolate collision domains for attached hosts and forward only appropriate traffic out a particular port, VLANs refine this concept further and provide complete isolation between VLANs. A VLAN is a bridging domain, and all broadcast and multicast traffic is contained within it.

- Security—VLANs provide security in the following two ways:

  - High-security users can be grouped into a VLAN, possibly on the same physical segment, and no users outside of that VLAN can communicate with them.

  - Because VLANs are logical groups that behave like physically separate entities, inter-VLAN communication is achieved through a router. When inter-VLAN communication occurs through a router, all the security and filtering functionality that routers traditionally provide can be used because routers are able to look at Layer 3 information. In the case of nonroutable protocols, there can be no inter-VLAN communication. All communication must occur within the same VLAN.

- Performance—The logical grouping of users allows, for example, a control engineer making intensive use of a networked PAC to be assigned to a VLAN that contains just that engineer and the I/O devices he or she needs. The work of the engineer does not affect the rest of the engineering group, which results in improved performance for the engineer (by being on a dedicated LAN) and improved performance for the rest of the engineering group (whose communications are not slowed down by the engineer using the network).

- Network management—The logical grouping of users, divorced from their physical or geographic locations, allows easier network management. It is no longer necessary to pull cables to move a user from one network to another. Adds, moves, and changes are achieved by configuring a port into the

appropriate VLAN. Expensive, time-consuming recabling to extend connectivity in a switched LAN environment is no longer necessary because network management can be used to logically assign a user from one VLAN to another.

Spanning Tree Protocol Design

STP Overview

Spanning Tree Protocol (STP) is a Layer 2 protocol designed to run on bridges and switches. Its main purpose is to ensure that loops are avoided when there are redundant paths by deterministically blocking appropriate interfaces. If a link failure occurs in such a network, the STP of choice is responsible for establishing a new path for data traffic. The IEEE specification is 802.1D, which has evolved to include the following:

- Common Spanning Tree
- Per VLAN Spanning Tree (PVST)
- Per VLAN Spanning Tree Plus (PVST+, a Cisco proprietary superset of 802.1D)
- Classic STP (802.1D)
- Multiple Instance Spanning Tree (MISTP/802.1S)
- Rapid Spanning Tree (RSTP/802.1W)

Cisco has a recommended Spanning Tree toolkit that includes the following:

- PortFast—Lets the access port bypass the listening and learning phases
- UplinkFast—Provides 3–5 second convergence after link failure
- BackboneFast—Cuts convergence time by MaxAge for indirect failure
- Loop Guard—Prevents the alternate or root port from being elected unless Bridge Protocol Data Units (BPDUs) are present
- Root Guard—Prevents external switches from becoming the root
- BPDU Guard—Disables a PortFast-enabled port if a BPDU is received
- BPDU Filter—Prevents sending or receiving BPDUs on PortFast-enabled ports

For more information on Spanning Tree, see the following URL:
http://www.cisco.com/warp/public/473/146.html

EttF version 1.1 recommends only RSTP, IEEE 802.1w, which includes the features from the Cisco Spanning Tree toolkit.

STP Configurable Parameters

With RSTP, IEEE 802.1w, Cisco does not recommend making many changes to the default STP settings. Only the bridge priority should be changed unless you have a valid reason for making other changes.

STP parameters include the following:

- Bridge priority—A configurable value to be used as portion of the bridge identifier. This is the first consideration of STP when root bridge determination is taking place.

The default value of the bridge priority is 32768. In root bridge calculation, the bridge with the lowest value is declared the root bridge. If two or more bridges are involved in a tie, the bridge address (MAC) is used as the final determining factor.

- Hello time—The time interval between the transmission of configuration BPDUs by a bridge that is attempting to become the root or is the root.

  The root bridge generates BPDU packets every *HelloTime* seconds, which according to the IEEE standards should be two seconds (2 sec). Each port of the switch has a timer associated with the BPDU information and receiving the BPDUs refreshes this timer.

- MaxAge—Maximum age of received protocol information before it is discarded.

  The information associated with a port is considered to be stale if the timer reaches *MaxAge.* The default MaxAge is twenty seconds (20 sec). When a switch stops receiving BPDUs from its root port and the MaxAge expires, the switch looks for a new root port, from the pool of blocking ports. If no blocking port is available, it claims to be the root itself on the designated ports.

- Forward delay—Time spent by a port in the listening state and the learning state before moving to the learning or forwarding state, respectively. It is also the value used for the aging time of dynamic entries in the filtering database, while received configuration messages indicate a topology change.

  The default value of the *Forward Delay* is fifteen seconds (15 sec).

- Diameter—Maximum number of bridges between any two points of attachment of end stations. Although this is not configurable directly, it can be manipulated by changing the max age variable from above.

  Network diameter can have a profound effect on the operation of STP and the network as a whole because the latency of BPDUs increases as the network grows in diameter. The default value for *Diameter* is seven (7).

- Port cost(s)—Contribution of the path through this port, when the port is the root port, to the total cost of the path to the root for this bridge.

  The cost of each port between a given bridge and the root bridge contributes to the overall path cost. Some of the default values used are as follows.

  - Gig E (1000 Mbps)—4
  - OC-12 (622 Mbps)—6
  - OC-3 (155 Mbps)—14
 - FastE (100 Mbps)—19
  - Ethernet (10 Mbps)—100

More on STP Redundancy

A Layer 2 loop is defined as the existence of two paths between any two Layer 2 devices within a single network. These loops can create many different problems within a network. The problems usually manifest themselves in the form of a "storm" incident. This is the continuous propagation of one or more packets within the network, such as a broadcast storm. When a Layer 2 switch receives a broadcast, it is sent to each of its ports including ports connected to other switches. If a loop exists in the network, it is possible for a switch to process the broadcast endlessly.

In Figure 4-4, Host A sends a broadcast to Switch 1.

Figure 4-4   Layer 2 Loop
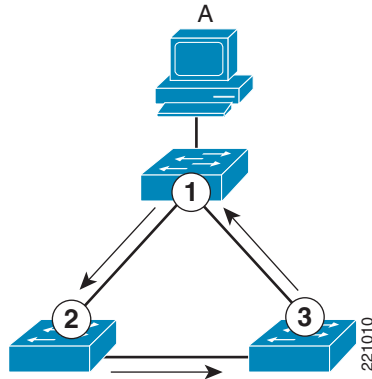
If a loop exists, as shown in Figure 4-4, the broadcast is forwarded to Switch 2, Switch 3, and back to Switch 1. Upon arrival in Switch 1, the broadcast is then resent to Switch 2. This is repeated until a break in the cycle is experienced. The loop is also experienced in the opposite direction and on any other loops present in the network. This loop eventually diminishes network performance. The use of STP prevents Layer 2 loops.

The implementation of STP allows the switches to communicate through BPDUs to form a loop-free topology at Layer 2. During this negotiation process, the switches use the algorithm to decide on the final state of all ports: blocking or forwarding. Blocking strategic ports prevents Layer 2 loops in the network.

In Figure 4-5, through the implementation of STP, Switch 3 is in the blocking state on its port facing Switch 2.

Figure 4-5   Use of STP to Block Loops
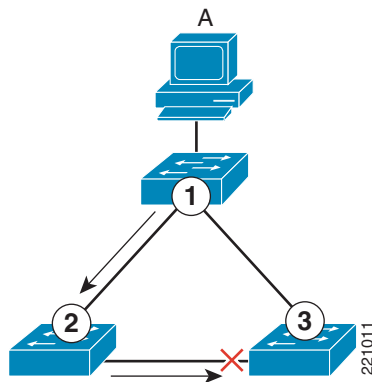
This effectively breaks the Layer 2 loop shown in Figure 4-4. Switch 3 still receives the broadcast; however, it comes from Switch 1. Furthermore, Switch 3 does not forward the broadcast received from Switch 1 to Switch 2 because the port is blocking.

STP Topology for EttF

For EttF version 1.1, the STP topology consists of *n* number of devices in a ring configuration. The purpose of a ring architecture is to achieve a level of redundancy at the lowest possible cabling cost. Furthermore, ring designs are popular in control network environments and are well-aligned with customer expectations. The ring devices are sitting at the access layer in a Cisco three-tier architecture with the distribution layer devices responsible for L2 (downlink) and L3 (uplink) functionality.

For EttF version 1.1, there are essentially 1 to many (1 to *n*) cell/area zones, each constituting a separate L2 STP domain. Only one VLAN is carried per cell/area zone. EttF device traffic flows both intra-cell (within a cell) and inter-cell (across cells). All multicast producer/consumer traffic is confined to within a cell/area zone because of the TTL=1 limitation on multicast traffic. To achieve a more deterministic design, Cisco recommends that a root bridge be chosen rather than letting the network choose one automatically. In Figure 4-6, Device CZ-C3750-1 is configured to be the root bridge.

Figure 4-6   STP Topology for EttF
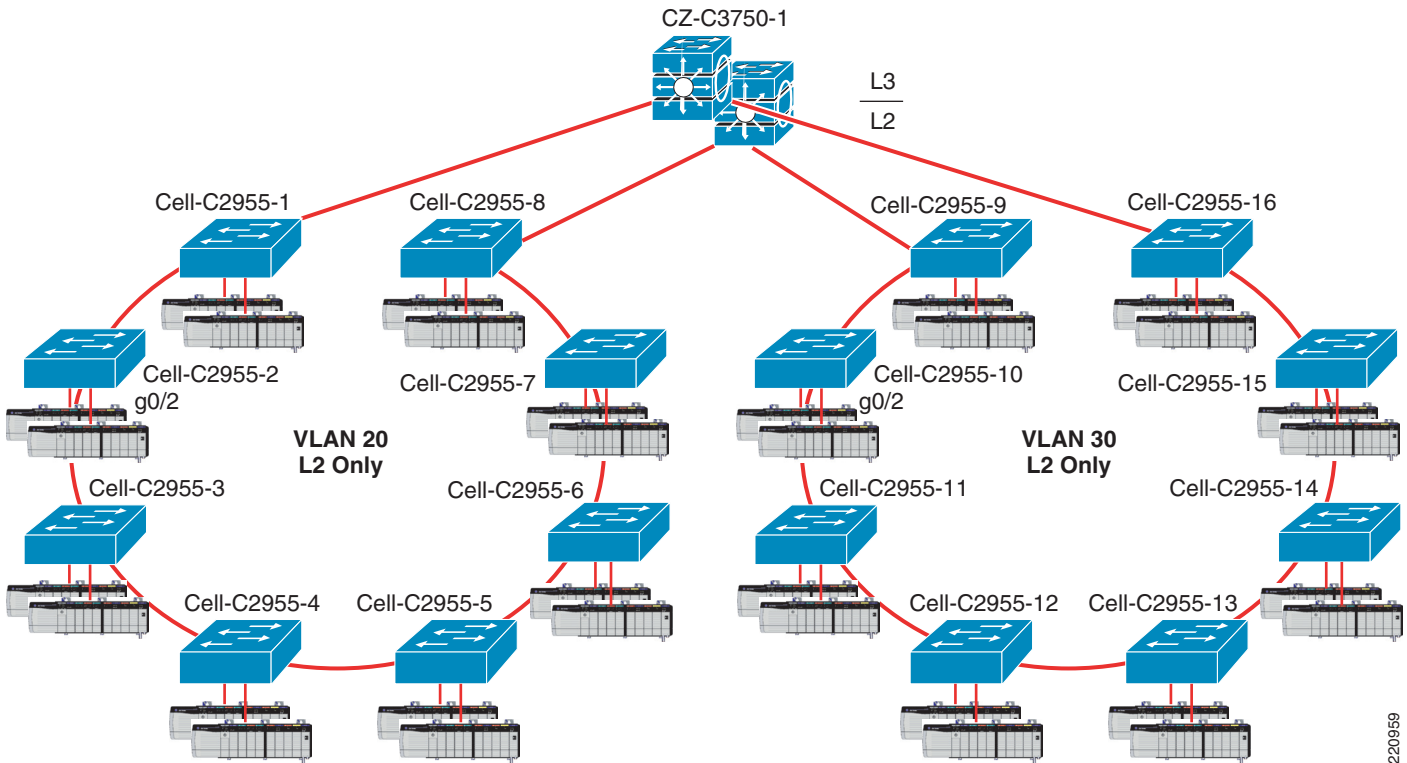
The root bridge can be viewed as the top of the network hierarchy with which every other switch in the network must communicate. The root bridge concept is crucial in determining a loop-free topology at Layer 2. In this determination, each device selects the best path possible towards the root bridge. The result of this is the spanning tree topology. Alternate paths towards the root bridge that would result in Layer 2 loops are in a blocked state.

STP Considerations for the Ring

As a general rule, RPVST+ should be deployed because of its faster convergence and ease of use. Enter the following global command to set the spanning-tree in Rapid-PVST mode:

```
spanning-tree mode rapid-pvst
```

The following sections describe further STP considerations and best practices for a ring topology.

Control Device Placement

Knowing which link is blocking is important in designing where to place devices; for example, a PAC talking to an I/O device. If the control engineer knows that PAC-A will be communicating with I/O-B for the majority of the time, it is advisable to connect them either on the same switch or on adjacent switches that are not STP blocked. This ensures the minimum convergence time if a network failure occurs. This also saves bandwidth and reduces latency because the traffic does not have to traverse the entire ring.

Trunk Ports or Access Ports

On the Cisco 2955, all uplink 1 GE ports should be configured as trunk ports carrying only one VLAN. All FastEthernet (fa0/1–fa0/12) interfaces should be configured as access ports for the particular VLAN carried on the uplink trunk (see Figure 4-7).

Figure 4-7   Trunk Port Configuration
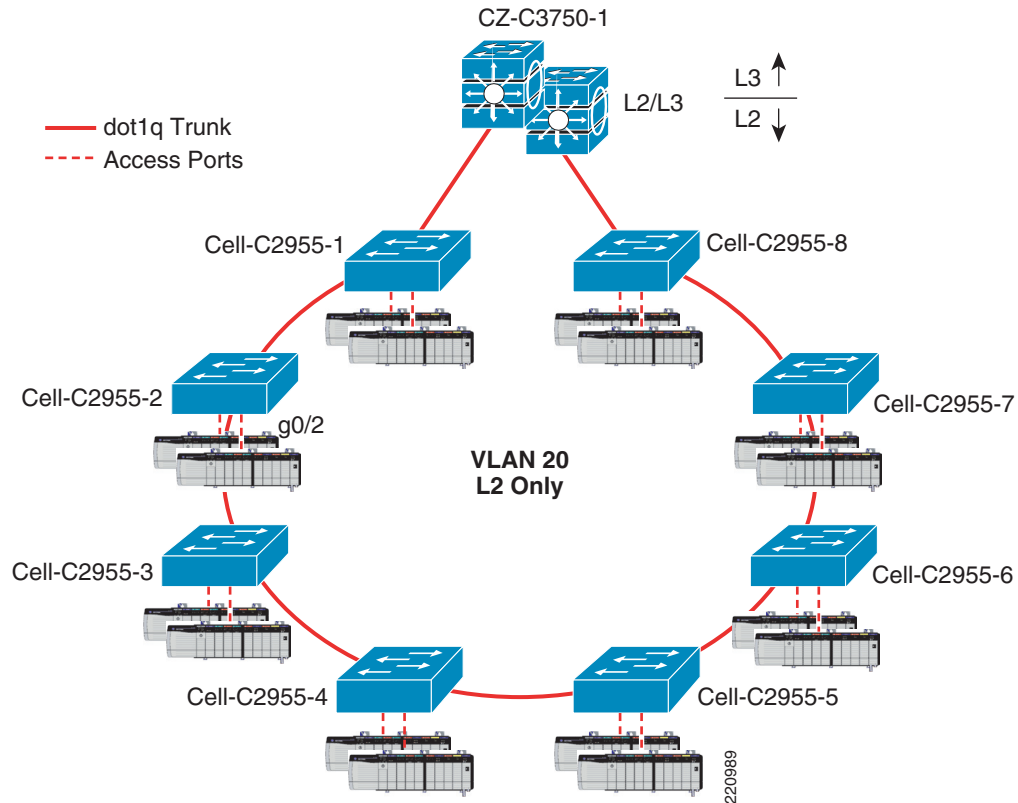
Sample Trunk Configuration

Following is a sample trunk configuration:

```
interface GigabitEthernet0/1
switchport trunk native vlan 20
switchport trunk allowed vlan 20
switchport mode trunk
end
Sample Access Port Configuration:
!
interface FastEthernet0/1
 switchport access vlan 20
 switchport mode access
 spanning-tree portfast
 spanning-tree bpdufilter enable
 spanning-tree bpduguard enable
 ip dhcp snooping limit rate 10
end
```

VLAN 1 Minimization

VLAN 1 has a special significance in Catalyst networks. When trunking, the Catalyst Supervisor Engine always uses the default VLAN, VLAN 1, to tag a number of control and management protocols such as Cisco Discovery Protocol (CDP), VLAN Trucking Protocol (VTP), and Port Aggregation Protocol (PAgP); as well as management protocols such as Simple Network Management Protocol (SNMP), Telnet, Secure Shell (SSH), and Syslog. All switch ports are configured by default to be members of VLAN 1, and all trunks carry VLAN 1 by default. When the VLAN is used in this way, it is referred to as the native VLAN. The default switch configuration sets VLAN 1 as the default native VLAN on the Catalyst trunk ports. You can leave VLAN 1 as the native VLAN, but keep in mind that any switches that run Cisco IOS Software in your network set all interfaces that are configured as Layer 2 switch ports to access ports in VLAN 1 by default. Most likely, a switch somewhere in the network uses VLAN 1 as a VLAN for user traffic.

The main concern with the use of VLAN 1 is that, in general, the Supervisor Engine should not be constantly interrupted by a lot of the broadcast and multicast traffic that end stations generate (user data). Multicast applications in particular tend to send a lot of data between servers and clients. The Supervisor Engine does not need to see this data. If the resources or buffers of the Supervisor Engine are fully occupied as the Supervisor Engine listens to unnecessary traffic, the Supervisor Engine can fail to see management packets that can cause a bridging loop.

VLAN 1 tags and handles most of the control plane traffic. VLAN 1 is enabled on all trunks by default. With larger campus networks, you need to be careful of the diameter of the VLAN 1 STP domain. Instability in one part of the network can affect VLAN 1 and can influence control plane stability and STP stability for all other VLANs. To limit the VLAN 1 transmission of user data and operation of STP on an interface, Cisco recommends doing the following:

- Clear VLAN 1 from the trunk to avoid control plane traffic being part of STP on those links, and allow only VLANs that have data traffic flowing through them:

    Switch(config)#**interface *type slot/port***

    Switch(config-if)#**switchport trunk allowed vlan 20**

- Change the native VLAN from 1 to an arbitrary VLAN that is not in use:

    Switch(config)#**interface *type slot/port***

    Switch(config-if)#**switchport trunk native vlan 999**

To verify, issue the following:

```
Switch#show int trunk

Port          Mode            Encapsulation  Status         Native vlan
Gi0/1         on              802.1q         trunking       999
Gi0/2         on              802.1q         trunking       999

Port      Vlans allowed on trunk
Gi0/1     20
Gi0/2     20

Port      Vlans allowed and active in management domain
Gi0/1     20
Gi0/2     20

Port      Vlans in spanning tree forwarding state and not pruned
Gi0/1     20
Gi0/2     20
```

You should see that the link is successfully trunking and carrying only VLAN 20, and that the native VLAN is something other than 1.

Location of the Root Bridge

The STP root bridge should be forced by setting this bridge to have the lowest priority. In the EttF 1.1 design, Cisco recommends that the Catalyst 3750 be elected as the root bridge because logically it sits at the top of the ring. This makes troubleshooting easier if there is a need to look at the STP state of the ring devices.

```
CZ-C3750-1(config)# spanning-tree vlan 1-1024 priority 8096
```

PortFast on Access Ports

You can use PortFast to bypass normal spanning tree operation on access ports. PortFast speeds up connectivity between end stations and the services to which end stations need to connect after link initialization. The Microsoft DHCP implementation needs to see the access port in forwarding mode immediately after the link state goes up to request and receive an IP address. Some protocols, such as Internetwork Packet Exchange (IPX)/Sequenced Packet Exchange (SPX), need to see the access port in forwarding mode immediately after the link state goes up to avoid get nearest server (GNS) problems.

PortFast Operational Overview

PortFast skips the normal listening, learning, and forwarding states of STP. This feature moves a port directly from blocking to forwarding mode after the link is seen as up. If this feature is not enabled, STP discards all user data until it decides that the port is ready to be moved to forwarding mode. This process can take up to twice the ForwardDelay time, which is 30 seconds by default.

PortFast mode prevents the generation of an STP topology change notification (TCN) each time a port state changes from learning to forwarding. TCNs are normal. However, a wave of TCNs that hits the root bridge can extend the convergence time unnecessarily. A wave of TCNs often occurs in the morning when people turn on their PCs.

The following are PortFast recommendations for EttF 1.1:

- Set STP PortFast to "on" for all enabled host ports connected to either a PAC, I/O device, or HMI.
- Explicitly set STP PortFast to "off" for switch-switch links and ports that are not in use.

STP Limitations

As mentioned earlier, EttF 1.1 implements only RPVST+ because it can achieve faster convergence times than PVST or PVST+ by relying on an active bridge-to-bridge handshake mechanism rather than depending on network-wide timers specified by the root bridge. With standard STP, convergence times are expected in the range of 30–60 seconds, depending on network conditions and timer settings. This is unacceptable in a control zone environment. However, with RPVST+, these numbers can be reduced to sub-second levels depending on conditions (type of failure, number of MAC addresses, traffic load). Also, the number of devices in the ring determine how fast STP converges. For a complete list of testing results, see Appendix A, "Characterization of the EttF Cell/Area Zone Design."

RSTP+ Convergence Process

Figure 4-8 shows the sequence of STP events in the RSTP+ convergence process.

Figure 4-8   **RSTP+ Convergence**

CZ-C3750-1
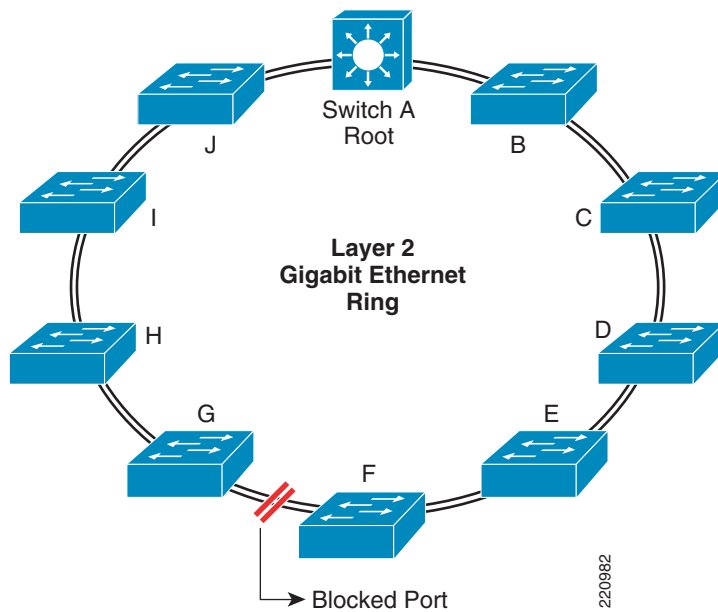
The sequence of STP events that take place given a link failure between Switch A and Switch B is as follows:

1. As soon as the failure occurs, Switch B loses its root port and claims to be the new root bridge.

2. This new root bridge information circulates down in the ring in a clockwise manner, from B to C, C to D, D to E and E to F.

3. When bridge F receives this inferior BPDU (it contains worse information than the one emanating from switch A, the "real" root), it in turn "replies" back to its upstream neighbors (E-D-C-B) to let them know it can still reach root switch A.

4. With RSTP+, there is no MaxAge parameter to depend on before bridge F can react to the reception of the inferior BPDU of E.

5. As soon as bridge F receives that inferior BPDU, it immediately transitions port 2/49 to forwarding from blocking and informs switch E via a "proposal" mechanism that it wants to become its designated bridge.

6. The same mechanism then takes place very rapidly between each pair of switches all the way towards bridge B.

7. Bridge F also initiates a topology change notification when opening up port 2/49 to flush stale learning table information in the other bridges of the ring.

8. The STP network is now converged.

Thus, L2 convergence is defined as the completion of all STP state changes and the completion of updating the MAC addresses in the CAM table as measured by data traffic.

Multicast Design

EtherNet/IP Multicast Traffic Patterns

Although traditional multicast services (usually video feeds) tend to scale with the number of streams, the EtherNet/IP model is implemented as a many-to-many model and scales differently because of a built-in feedback mechanism. In the specification, devices generate data for consumption by other devices. The devices that generate the data are called producers, and the devices receiving the information are called consumers. The data exchange model is therefore referred to as the producer-consumer model. Multicast is more efficient over unicast in that in many cases, multiple consumers want the same information from a particular producer. However, because every consumer of traffic needs to respond with a heartbeat, a significant load of unicast packets is generated on the network.

Most EtherNet/IP devices generate very little data. However, networks with a large number of nodes can generate a large aggregate amount of multicast traffic. If a method to control this is not deployed, this aggregate traffic can swamp some or all of the end devices in the network. This is aggravated by the fact that there is likely unicast traffic (FTP, HTTP, and so on) going to the device. This is even more critical if there are no other mechanisms in place (such as QoS) to prioritize real-time traffic.

In general, end devices can be overwhelmed by the following:

- The port speed is overrun

- More packets are received than the network interface controller can handle

- More packets are received than the host processor can process

If the aggregate data exceeds the port speed (that is, 20 Mbps going to a 10 Mbps configured port), traffic is dropped because of congestion.

If a device on the network can process only 900 packets per second (pps), the network design must ensure that this particular end device does not see more than 900 pps. The following list is an example of a network that produces more than 900 pps. In this example, there are 12 producers. It is assumed that each producer is also a consumer. Each producer is generating only 100 pps, but because all consumers see all producer traffic, they actually see 1200 pps, and because they also are acknowledging at least one stream, they are also sending 100 pps of unicast traffic.

- Number of PACs—12
- RPI—10 Msecs
- Packets of multicast—100 pps
- Size of multicast packets—600 bytes
- Bandwidth of each multicast stream—0.5 Mbps
- Packets of unicast—100 pps
- Size of unicast packets—100 bytes
- Bandwidth of each unicast stream (hello echoes)—0.096 Mbps
- Backplane traffic—7.1 Mbps
- Traffic seen on each port (multicast + unicast)—6.0 Mbps
- Number of packets received—1300 pps

IGMP snooping ensures that only the multicast traffic requested by the particular end device is received on its inbound interface. With IGMP snooping, each end device processes only 200 pps instead of the aggregate 1300.

Figure 4-9 shows an example of multicast with the producer-consumer model.

Figure 4-9   Multicast on Producer-Consumer Model
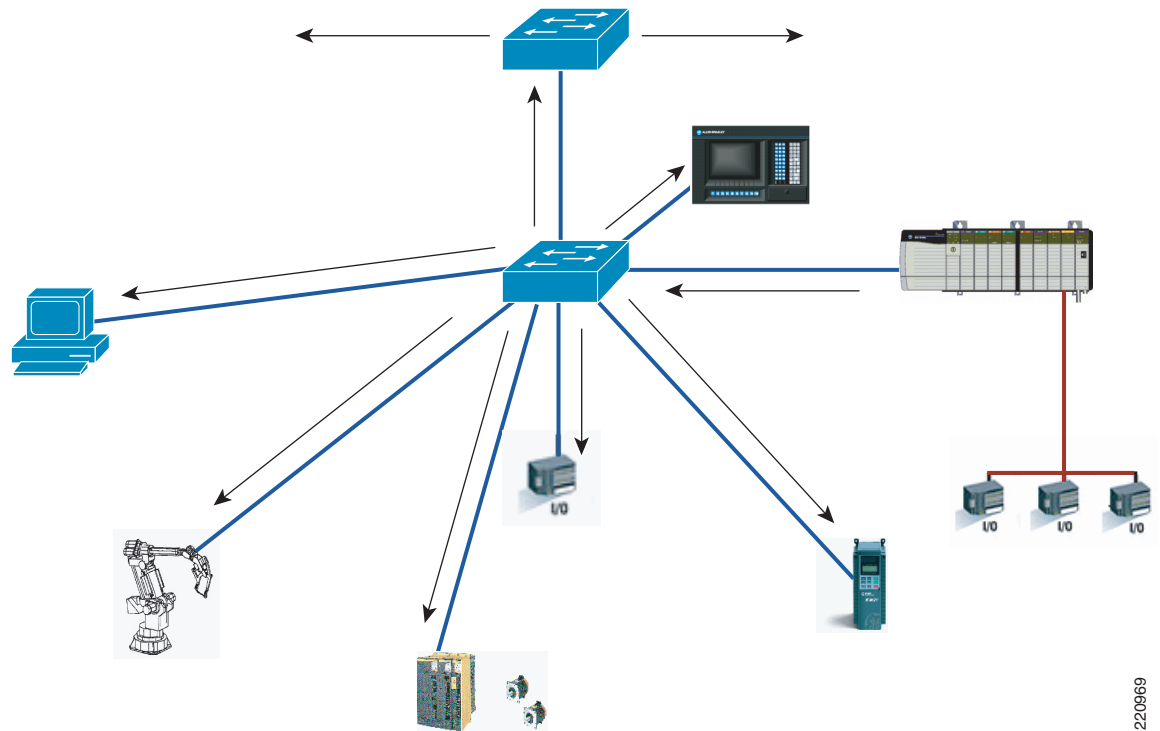
IGMP Snooping

Layer 2 switches can use IGMP snooping to constrain the flooding of multicast traffic by dynamically configuring the multicast traffic to be forwarded to only those access interfaces associated with devices requesting the multicast group. As the name implies, IGMP snooping requires the LAN switch to snoop on the Internet Group Management Protocol (IGMP) transmissions between the host and the router and to keep track of multicast groups and member ports. When the switch receives an IGMP report from a host for a particular multicast group, the switch adds the host port number to the forwarding table entry (see Figure 4-10); when it receives an IGMP leave group message from a host, it removes the host port from the table entry (see Figure 4-11). It also periodically deletes entries if it does not receive IGMP membership reports from the multicast clients.

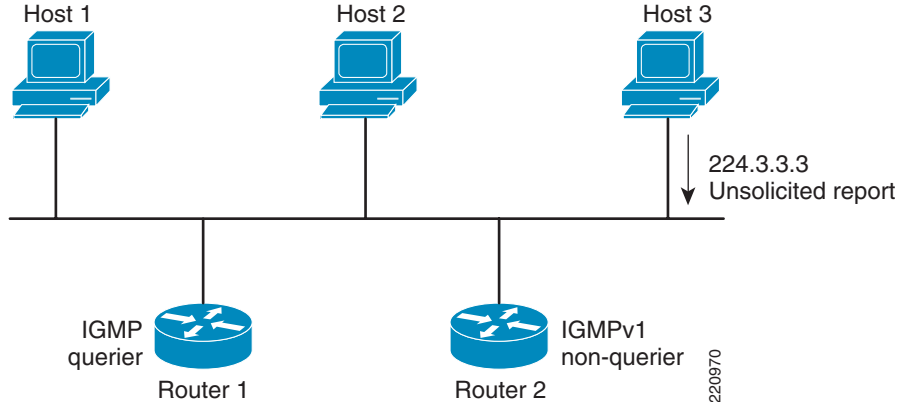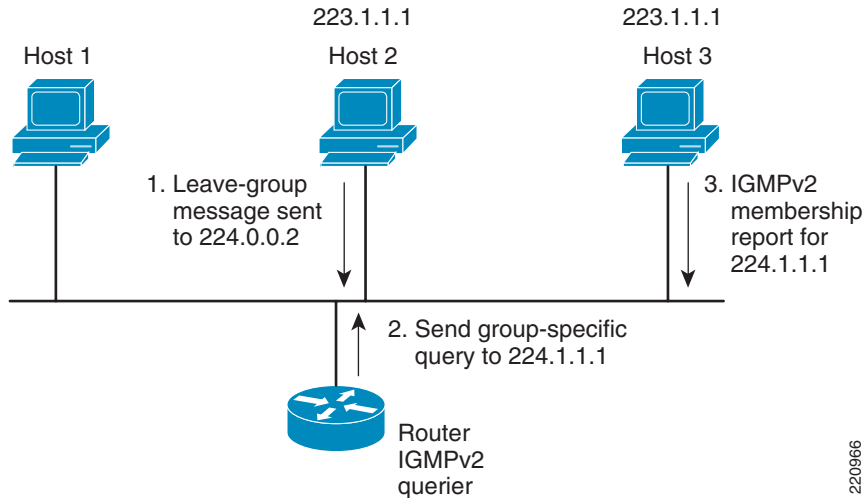*Figure 4-10*   *IGMPv1 Join Process*

Figure 4-11   *IGMPv2 Leave Process*

The multicast router sends out periodic general queries to all VLANs. All hosts interested in this multicast traffic send join requests and are added to the forwarding table entry. The switch creates one entry per VLAN in the Layer 2 forwarding table for each MAC group from which it receives an IGMP join request. It sets a flag for each port that will receive the particular group.

Multicast group membership lists can consist of both user-defined and IGMP snooping-learned settings. If an STP TCN, a port group, or a VLAN ID change occurs, the IGMP snooping-learned multicast groups from this port on the VLAN are deleted, and must be relearned on the next IGMP query message.

When a host connected to the switch wants to join an IP multicast group, it sends an unsolicited IGMP join message, specifying the IP multicast group to join. Alternatively, when the switch receives a general query from the router, it forwards the query to all ports in the VLAN. Hosts wanting to join the multicast group respond by sending a solicited report (join) message to the switch. The switch creates a multicast forwarding table entry for the group if one is not already present. It also adds the interface where the join message was received to the forwarding table entry. The host associated with that interface receives multicast traffic for that multicast group.

When hosts want to leave a multicast group, they can either silently leave by not responding to an IGMP query message, or they can send an IGMPv2 leave message. When the switch receives a leave message from a host, it sends out a MAC-based general query to determine whether any other devices connected

to that interface are interested in traffic for the specific multicast group. The switch then updates the forwarding table for that MAC group so that only those hosts interested in receiving multicast traffic for the group are listed in the forwarding table. If the router receives no reports from a VLAN, it removes the group for the VLAN from its IGMP cache.

I/O or produce tag traffic does not pass through a router. By design, the time-to-live (TTL) parameter is configured in the Rockwell Automation firmware for a value of 1. This value is decremented by any router and then discarded. A value of 1 was selected to avoid attempts to implement I/O control (high-speed) through a slow network device (router) or through a slow network. I/O or produce tag traffic uses multicast messaging.

IGMP Querier and EtherNET/IP Traffic

The purpose of the IGMP snooping with querier is to generate periodic query messages. Consumers of particular multicast groups should respond with an IGMP report message stating that they still want to receive the data stream. If the client is finished with the stream, it does not issue the report message. The Ethernet switch, using IGMP snooping, should prune off that port, which has the same effect as the client issuing an unsolicited IGMP leave message. In the EttF architecture, IGMP snooping is turned on by default in all the switches in the cell/area zone (Catalyst 2955) and the Catalyst 3750. The Catalyst 3750 acts as the IGMP querier. A querier must be configured for each switched virtual interface (SVI) associated with each VLAN. The querier knob can be turned on by configuring **ip pim sparse-dense-mode** in the SVI-Vlan 20 (see Figure 4-12).

Figure 4-12   *EttF Architecture*
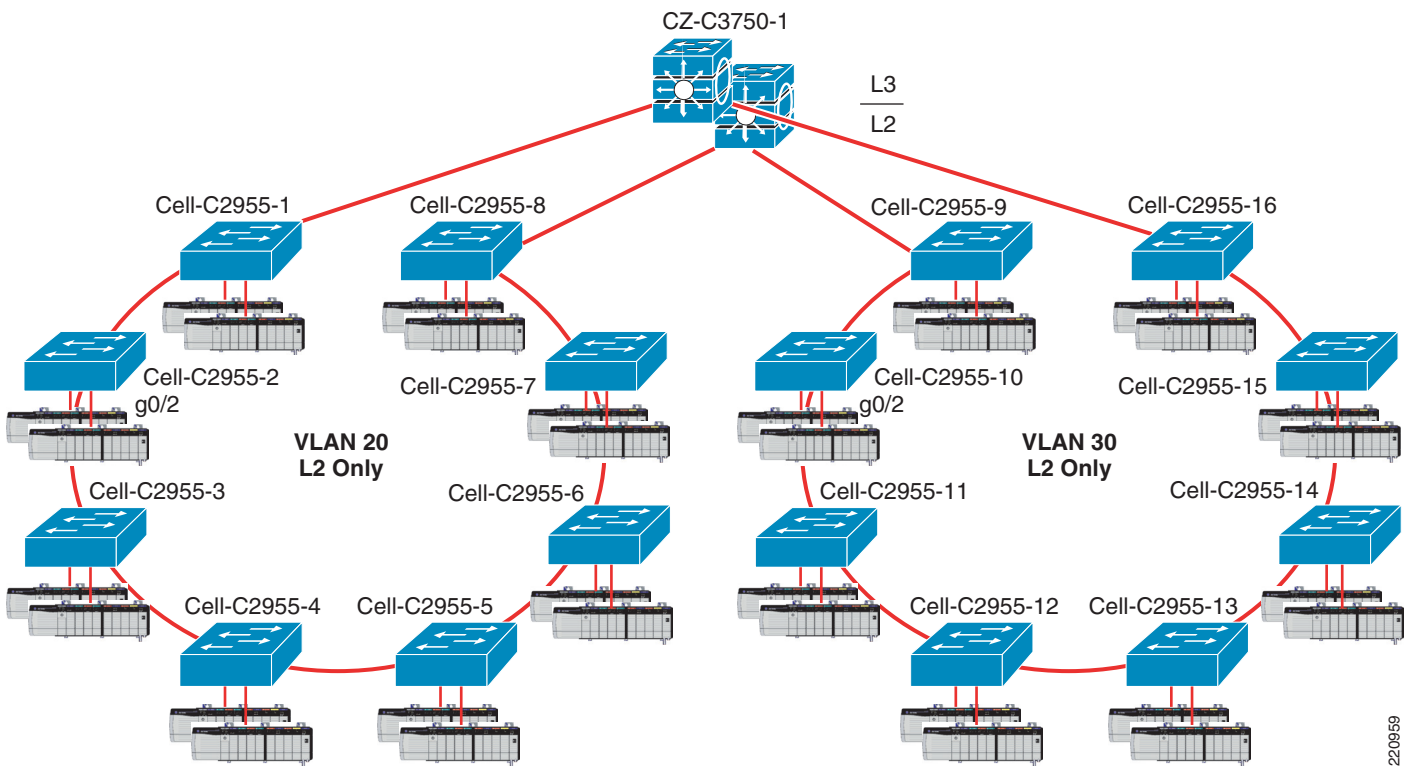
Another option is to configure the global IP IGMP querier address from the CLI as follows: **ip igmp snooping querier address ip_address**. For more information on configuring IGMP snooping with querier, see the following URL:
http://www.cisco.com/en/US/partner/products/hw/switches/ps5532/products_configuration_guide_chapter09186a008081bb8c.html#wp1130762.

IGMP Configurations

Configure IGMP querier on the Catalyst 3750 multicast router (**ip pim sparse-dense-mode**) as follows:

```
C3750-1#
Building configuration...

Current configuration: 87 bytes
!
interface Vlan20
 ip address 10.17.20.1 255.255.255.0
 ip pim sparse-dense-mode
end
```

The **show ip igmp int vlan <>** command verifies the IGMP querying router and the multicast designated router:

```
C3750-1#show ip igmp int vlan 20
Vlan20 is up, line protocol is up
  Internet address is 10.17.20.1/24
  IGMP is enabled on interface
  Current IGMP host version is 2
  Current IGMP router version is 2
  IGMP query interval is 60 seconds
  IGMP querier timeout is 120 seconds
  IGMP max query response time is 10 seconds
  Last member query count is 2
  Last member query response interval is 1000 ms
  Inbound IGMP access group is not set
  IGMP activity: 10 joins, 5 leaves
  Multicast routing is enabled on interface
  Multicast TTL threshold is 0
  Multicast designated router (DR) is 10.17.20.1 (this system)
  IGMP querying router is 10.17.20.1 (this system)
  No multicast groups joined by this system
C3750-1#
```

The multicast groups attached to the Catalyst 3750 (SVI 20) are as follows (**show ip igmp groups vlan 20**):

```
C3750-1#show ip igmp groups vlan 20
IGMP Connected Group Membership
Group Address    Interface             Uptime    Expires   Last Reporter
239.192.21.96    Vlan20                1w0d      00:02:09  10.17.20.164
239.192.21.160   1w1d       00:02:11  10.17.20.164
239.192.24.160   Vlan20                6d01h     00:02:13  10.17.20.164
239.192.24.161   Vlan20                6d01h     00:02:13  10.17.20.164
239.192.24.192   Vlan20                6d02h     00:02:11  10.17.20.164
C3750-1#
```

The output from the **show ip igmp snooping vlan <>** command verifies IGMP snooping as well as the multicast router learning mode. If IGMP snooping with querier is not enabled on the Catalyst 3750, the bottom portion of the output with the multicast router mode is not present:

```
cell-c2955-12#show ip igmp snooping vlan 20
Global IGMP snooping configuration:
-----------------------------------
IGMP snooping              : Enabled
IGMPv3 snooping (minimal): Enabled
Report suppression        : Enabled
TCN solicit query         : Disabled
TCN flood query count     : 2
Last member query interval: 1000

Vlan 20:
--------
IGMP snooping                     : Enabled
Immediate leave                   : Disabled
Multicast router learning mode    : pim-dvmrp
Source only learning age timer    : 10
Last member query interval        : 1000
CGMP interoperability mode        : IGMP_ONLY
cell-c2955-12#
```

The output from the Catalyst 2955 in the cell/area zone (**show ip igmp snooping querier**) verifies the IGMP querier in the Catalyst 3750:

```
cell-c2955-12#show ip igmp snooping querier
Vlan      IP Address      IGMP Version        Port
--------------------------------------------------
20        10.17.20.1      v2                  Gi0/1

cell-c2955-12#
```

Switch Troubleshooting Toolkit

The following are some general troubleshooting techniques on the cell/area zone switches:

1. Use Port Mirroring (SPAN) when troubleshooting and you need to characterize the traffic flow.

   ```
   Switch(config)# monitor session 1 source interface fastethernet0/1
   Switch(config)# monitor session 1 destination interface fastethernet0/2 <- A Sniffer
   would be connected here
   ```

2. Use **show ip traffic** to get a quick snapshot of traffic statistics to check whether there are any skewed data points; for example, excessive broadcasts indicate a broadcast storm.

   Sample output is as follows:

   ```
   IP statistics:
       Rcvd: 98 total, 98 local destination
           0 format errors, 0 checksum errors, 0 bad hop count
           0 unknown protocol, 0 not a gateway
           0 security failures, 0 bad options
          Frags:0 reassembled, 0 timeouts, 0 too big
           0 fragmented, 0 couldn't fragment
       Bcast:38 received, 52 sent
       Sent: 44 generated, 0 forwarded
           0 encapsulation failed, 0 no route
       ICMP statistics:
       Rcvd: 0 format errors, 0 checksum errors, 0 redirects, 0 unreachable
             0 echo, 0 echo reply, 0 mask requests, 0 mask replies, 0 quench
   ```

```
              0 parameter, 0 timestamp, 0 info request, 0 other
        Sent: 0 redirects, 3 unreachable, 0 echo, 0 echo reply
              0 mask requests, 0 mask replies, 0 quench, 0 timestamp
              0 info reply, 0 time exceeded, 0 parameter problem
      UDP statistics:
      Rcvd: 56 total, 0 checksum errors, 55 no port
      Sent: 18 total, 0 forwarded broadcasts
      TCP statistics:
        Rcvd: 0 total, 0 checksum errors, 0 no port
      Sent: 0 total
      EGP statistics:
      Rcvd: 0 total, 0 format errors, 0 checksum errors, 0 no listener
        Sent: 0 total
      IGRP statistics:
      Rcvd: 73 total, 0 checksum errors
        Sent: 26 total
      HELLO statistics:
      Rcvd: 0 total, 0 checksum errors
      Sent: 0 total
      ARP statistics:
      Rcvd: 20 requests, 17 replies, 0 reverse, 0 other
      Sent: 0 requests, 9 replies (0 proxy), 0 reverse
      Probe statistics:
      Rcvd: 6 address requests, 0 address replies
      0 proxy name requests, 0 other
      Sent: 0 address requests, 4 address replies (0 proxy)
          0 proxy name replies
```

3. To troubleshoot link problems (port flapping, errdisable, and so on), see the guide at the following URL:
http://www.cisco.com/en/US/partner/products/hw/switches/ps700/products_tech_note09186a008015bfd6.shtml

4. Verify that CPU usage is below 50 percent average utilization. High CPU can indicate a broadcast storm, multicast data traffic flooding (for example, IGMP snooping configured incorrectly), or an attack.

```
cell-c2955-9#show proc cpu | exc 0.00
CPU utilization for five seconds: 1%/0%; one minute: 0%; five minutes: 0%
 PID Runtime(ms)   Invoked     uSecs  5Sec   1Min   5Min TTY Process
  22    6918340   3448871      2005  0.40%  0.39%  0.40%   0 Calhoun Statisti
  68         60        43      1395  0.90%  0.09%  0.02%   0 Exec
```