Configure Multiple Spanning Tree Protocol

This chapter introduces you to Multiple Spanning Tree Protocol (MSTP) which is one of the variants of Spanning Tree Protocol (STP) and describes how you can configure the MSTP feature.

- Overview of Spanning Tree Protocol, on page 1
- Overview of MSTP, on page 2
- MSTP Support on Cisco NCS 5000 Series Routers, on page 2
- Configuring MSTP, on page 4
- Configuring MSTP BPDU Guard, on page 6
- References for Spanning Tree Protocol, on page 8

Overview of Spanning Tree Protocol

Ethernet is no longer just a link-layer technology used to interconnect network vehicles and hosts. Its low cost and wide spectrum of bandwidth capabilities coupled with a simple plug and play provisioning philosophy have transformed Ethernet into a legitimate technique for building networks, particularly in the access and aggregation regions of service provider networks.

Ethernet networks lacking a TTL field in the Layer 2 (L2) header and, encouraging or requiring multicast traffic network-wide, are susceptible to broadcast storms if loops are introduced. However, loops are a desirable property as they provide redundant paths. Spanning tree protocols (STP) are used to provide a loop free topology within Ethernet networks, allowing redundancy within the network to deal with link failures.

There are many variants of STP; however, they work on the same basic principle. Within a network that may contain loops, a sufficient number of interfaces are disabled by STP so as to ensure that there is a loop-free spanning tree, that is, there is exactly one path between any two devices in the network. If there is a fault in the network that affects one of the active links, the protocol recalculates the spanning tree so as to ensure that all devices continue to be reachable. STP is transparent to end stations which cannot detect whether they are connected to a single LAN segment or to a switched LAN containing multiple segments and using STP to ensure there are no loops.

For more information, see References for Spanning Tree Protocol, on page 8

Restrictions for STP on Cisco NCS 5000 Series Routers

The following restrictions are applicable for STP on Cisco NCS 5000 Series Routers

- The only type of STP that is supported on Cisco NCS 5000 Series Routers is Multiple Spanning Tree Protocol (MSTP).
- Per vlan Spanning Tree(PVST/PVST+/PVRST) is not supported on Cisco NCS 5000 Series Routers.
- Access gateway feature is not supported.

Overview of MSTP

The Multiple Spanning Tree Protocol (MSTP) is an STP variant that allows multiple and independent spanning trees to be created over the same physical network. The parameters for each spanning tree can be configured separately, so as to cause a different network devices to be selected as the root bridge or different paths to be selected to form the loop-free topology. Consequently, a given physical interface can be blocked for some of the spanning trees and unblocked for others.

Having set up multiple spanning tree instances, the set of VLANs in use can be partitioned among them; for example, VLANs 1 - 100 can be assigned to spanning tree instance 1, VLANs 101 - 200 can be assigned to spanning tree instance 2, VLANs 201 - 300 can be assigned to spanning tree instance 3, and so on. Since each spanning tree has a different active topology with different active links, this has the effect of dividing the data traffic among the available redundant links based on the VLAN—a form of load balancing.

MSTP Support on Cisco NCS 5000 Series Routers

Cisco NCS 5000 Series Routers support MSTP, as defined in IEEE 802.1Q-2005, on physical Ethernet interfaces and Ethernet Bundle interfaces.

In addition, the below Cisco features are supported:

- BPDU Guard—This Cisco feature protects against misconfiguration of edge ports.
- Flush Containment—This Cisco feature helps prevent unnecessary MAC flushes that would otherwise occur following a topology change.
- Bringup Delay—This Cisco feature prevents an interface from being added to the active topology before it is ready to forward traffic.

MSTP BPDU Guard

The MSTP BPDU Guard feature protects against misconfiguration of edge ports.

Note In order to enable the MSTP BPDU Guard feature for an interface, the command **portfast bpduguard** must be configured on it.

Port Fast

The Port Fast feature manage the ports at the edge of the switched Ethernet network. For devices that only have one link to the switched network (typically host devices), there is no need to run MSTP, as there is only one available path. Furthermore, it is undesirable to trigger topology changes (and resultant MAC flushes) when the single link fails or is restored, as there is no alternative path.

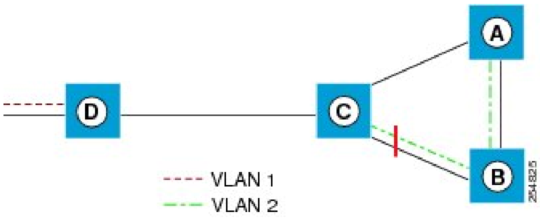
By default, MSTP monitors ports where no BPDUs are received, and after a timeout, places them into edge mode whereby they do not participate in MSTP. When **portfast** is explicitly configured on an interface, MSTP considers that interface to be an edge port and removes it from consideration when calculating the spanning tree. And hence the convergence time for the whole network is improved when **portfast** is configured.

Note MSTP BPDU Guard feature is supported by configuring interfaces in port fast mode. BPDU guard feature will error-disable the port on receiving BPDU packets.

Flush Containment

Flush containment is a Cisco feature that helps prevent unnecessary MAC flushes due to unrelated topology changes in other areas of a network. This is best illustrated by example. The following figure shows a network containing four devices. Two VLANs are in use: VLAN 1 is only used on device D, while VLAN 2 spans devices A, B and C. The two VLANs are in the same spanning tree instance, but do not share any links.

Figure 1: Flush Containment



If the link AB goes down, then in normal operation, as C brings up its blocked port, it sends out a topology change notification on all other interfaces, including towards D. This causes a MAC flush to occur for VLAN 1, even though the topology change which has taken place only affects VLAN 2.

Flush containment helps deal with this problem by preventing topology change notifications from being sent on interfaces on which no VLANs are configured for the MSTI in question. In the example network this would mean no topology change notifications would be sent from C to D, and the MAC flushes which take place would be confined to the right hand side of the network.

Note Flush containment is enabled by default, but can be disabled by configuration, thus restoring the behavior described in the IEEE 802.1Q standard.

Bringup Delay

Bringup delay is a Cisco feature that stops MSTP from considering an interface when calculating the spanning tree, if the interface is not yet ready to forward traffic. This is useful when a line card first boots up, as the system may declare that the interfaces on that card are *Up* before the dataplane is fully ready to forward traffic. According to the standard, MSTP considers the interfaces as soon as they are declared *Up*, and this may cause it to move other interfaces into the blocking state if the new interfaces are selected instead.

Bringup delay solves this problem by adding a configurable delay period which occurs as interfaces that are configured with MSTP first come into existence. Until this delay period ends, the interfaces remain in blocking state, and are not considered when calculating the spanning tree.

Bringup delay only takes place when interfaces which are already configured with MSTP are created, for example, on a card reload. No delay takes place if an interface which already exists is later configured with MSTP.

Configuring MSTP

The different steps involved in configuring MSTP are as follows:

1. Configure VLAN interfaces

```
Router# configure
Router(config)# interface TenGigE0/0/0/2.1001 l2transport
Router(config-subif)# encapsulation dot1q 1001
Router(config)# interface TenGigE0/0/0/3.1001 l2transport
Router(config-subif)# encapsulation dot1q 1001
Router(config)# interface TenGigE0/0/0/14.1001 l2transport
Router(config-subif)# encapsulation dot1q 1001
Router(config)# interface TenGigE0/0/0/2.1021 l2transport
Router(config-subif)# encapsulation dot1q 1021
Router(config)# interface TenGigE0/0/0/3.1021 l2transport
Router(config-subif)# encapsulation dot1q 1021
Router(config)# interface TenGigE0/0/0/14.1021 l2transport
Router(config-subif)# encapsulation dot1q 1021
Router(config-subif)# commit
```

2. Configure L2VPN bridge-domains with the VLAN interfaces configured in the previous step.

```
Router# configure
Router(config)# l2vpn bridge group mstp
Router(config-l2vpn-bg)# bridge-domain mstp1001
Router(config-l2vpn-bg-bd)# int TenGigE 0/0/0/2.1001
Router(config-l2vpn-bg-bd-ac)# exit
Router(config-l2vpn-bg-bd)# int TenGigE 0/0/0/3.1001
Router(config-l2vpn-bg-bd-ac)# exit
Router(config-l2vpn-bg-bd)# int TenGigE 0/0/0/14.1001
Router(config-l2vpn-bg-bd-ac)# exit
Router(config-l2vpn-bg-bd)# exit
Router(config-l2vpn-bg)# exit
Router(config-l2vpn-bg)# bridge-domain mstp1021
Router(config-l2vpn-bg-bd)# int TenGigE 0/0/0/2.1021
Router(config-l2vpn-bg-bd-ac)# exit
Router(config-l2vpn-bg-bd)# int TenGigE 0/0/0/3.1021
Router(config-l2vpn-bg-bd-ac)# exit
Router(config-l2vpn-bg-bd)# int TenGigE 0/0/0/14.1021
Router(config-l2vpn-bg-bd-ac)# commit
```

3. Configure MSTP.

```
Router# configure
Router(config)# spanning-tree mst abc
Router(config-mstp)# name mstp1
Router(config-mstp)# instance 1001
Router(config-mstp-inst)# vlan-ids 1001-1020
Router(config-mstp-inst)# exit
Router(config-mstp)# instance 1021
Router(config-mstp-inst)# vlan-ids 1021-1040
Router(config-mstp-inst)# exit
Router(config-mstp)# int tenGigE 0/0/0/2
Router(config-mstp-if)# exit
Router(config-mstp)# int tenGigE 0/0/0/3
Router(config-mstp-if)# exit
Router(config-mstp)# int tenGigE 0/0/0/14
Router(config-mstp-if)# commit
```

Running Configuration for MSTP

```
!
Configure
/* Configure VLAN interfaces */
interface TenGigE0/0/0/2.1001 l2transport
 encapsulation dot1q 1001
!
interface TenGigE0/0/0/3.1001 l2transport
 encapsulation dot1q 1001
!
interface TenGigE0/0/0/14.1001 l2transport
 encapsulation dot1q 1001

interface TenGigE0/0/0/2.1021 l2transport
 encapsulation dot1q 1021
!
interface TenGigE0/0/0/3.1021
 l2transport
 encapsulation dot1q 1021
!
interface TenGigE0/0/0/14.1021 l2transport
 encapsulation dot1q 1021
!
/* Configure L2VPN Bridge-domains */
l2vpn
 bridge group mstp
  bridge-domain mstp1001
   interface TenGigE0/0/0/2.1001
   !
   interface TenGigE0/0/0/3.1001
   !
   interface TenGigE0/0/0/14.1001
   !
bridge-domain mstp1021
   interface TenGigE0/0/0/2.1021
   !
   interface TenGigE0/0/0/3.1021
   !
   interface TenGigE0/0/0/14.1021
!
/* Configure MSTP */
spanning-tree mst abc
 name mstp1
 instance 1001
  vlan-ids 1001-1020
 !
```

```
instance 1021
  vlan-ids 1021-1040
 !
interface TenGigE0/0/0/2
 !
 interface TenGigE0/0/0/3
 !
 interface TenGigE0/0/0/14
```

Verification for MSTP

The MSTP configuration can be verified using the command **show spanning-tree mst**

```
/* Verify the MSTP configuration */
Router# show spanning-tree mst abc instance 121
Mon Jan 23 12:11:48.591 UTC
Role:  ROOT=Root, DSGN=Designated, ALT=Alternate, BKP=Backup, MSTR=Master
State: FWD=Forwarding, LRN=Learning, BLK=Blocked, DLY=Bringup Delayed

Operating in dot1q mode


MSTI 121:

  VLANS Mapped: 121-130

  Root ID    Priority    32768
             Address     dceb.9456.b9d4
             This bridge is the root
             Int Cost    0
             Max Age 20 sec, Forward Delay 15 sec


  Bridge ID  Priority    32768 (priority 32768 sys-id-ext 0)
             Address     dceb.9456.b9d4
             Max Age 20 sec, Forward Delay 15 sec
             Max Hops 20, Transmit Hold count  6


Interface    Port ID           Role State Designated           Port ID
             Pri.Nbr Cost                 Bridge ID            Pri.Nbr
------------ ------- --------- ---- ----- -------------------- -------
BE1          128.1   10000     DSGN FWD   32768 dceb.9456.b9d4 128.1
Te0/0/0/1    128.2   2000      DSGN FWD   32768 dceb.9456.b9d4 128.2
Te0/0/0/16   128.3   2000      DSGN FWD   32768 dceb.9456.b9d4 128.3
Te0/0/0/17   128.4   2000      DSGN FWD   32768 dceb.9456.b9d4 128.4
```

Configuring MSTP BPDU Guard

This section describes how you can configure MSTP BPDU Guard.

```
Router# configure
Router(config)# l2vpn bridge group bg1
Router(config-l2vpn-bg)# bridge-domain bd1
Router(config-l2vpn-bg-bd)# int TenGigE 0/0/0/7
Router(config-l2vpn-bg-bd-ac)# root
Router(config)# spanning-tree mst m0
Router(config-mstp)# interface tenGigE 0/0/0/7
Router(config-mstp-if)# portfast bpduguard
Router(config-mstp-if)# root
Router(config)# int tenGigE 0/0/0/7 l2transport
Router(config-if-l2)# commit
```

Running Configuration with MSTP BPDU Guard

```
!
Configure
l2vpn
 bridge group bg1
  bridge-domain bd1
   interface TenGigE0/0/0/7
   !
spanning-tree mst m0
 interface TenGigE0/0/0/7
  portfast bpduguard
!
interface TenGigE0/0/0/7
 l2transport
 !
```

Verification for MSTP BPDU Guard

Verify that you have configured MSTP BPDU Guard.

```
/* Verify the MSTP BPDU Guard configuration */
Router# show interfaces tenGigE 0/0/0/7
Wed Nov  9 09:23:56.268 UTC
TenGigE0/0/0/7 is error disabled, line protocol is administratively down
  Interface state transitions: 2
  Hardware is TenGigE, address is 7cad.7425.c8c8 (bia 7cad.7425.c8c8)
  Layer 2 Transport Mode
  MTU 1514 bytes, BW 10000000 Kbit (Max: 10000000 Kbit)
     reliability 255/255, txload 0/255, rxload 0/255
  Encapsulation ARPA,
  Full-duplex, 10000Mb/s, link type is force-up
  output flow control is off, input flow control is off
  Carrier delay (up) is 10 msec
  loopback not set,
  Last link flapped 00:00:49
  Last input 00:00:40, output 00:00:40
  Last clearing of "show interface" counters never
  5 minute input rate 0 bits/sec, 0 packets/sec
  5 minute output rate 0 bits/sec, 0 packets/sec
     38752 packets input, 4611429 bytes, 0 total input drops
     1 drops for unrecognized upper-level protocol
     Received 1 broadcast packets, 38751 multicast packets
              0 runts, 0 giants, 0 throttles, 0 parity
     0 input errors, 0 CRC, 0 frame, 0 overrun, 0 ignored, 0 abort
```

References for Spanning Tree Protocol

This section provides references for STP. For an overview of STP, see Overview of Spanning Tree Protocol, on page 1

STP Operation

All variants of STP operate in a similar fashion: STP frames (known as bridge protocol data units (BPDUs)) are exchanged at regular intervals over Layer 2 LAN segments, between network devices participating in STP. Such network devices do not forward these frames, but use the information to construct a loop free spanning tree.

The spanning tree is constructed by first selecting a device which is the *root* of the spanning tree (known as the root bridge), and then by determining a loop free path from the *root bridge* to every other device in the network. Redundant paths are disabled by setting the appropriate ports into a blocked state, where STP frames can still be exchanged but data traffic is never forwarded. If a network segment fails and a redundant path exists, the STP protocol recalculates the spanning tree topology and activates the redundant path, by unblocking the appropriate ports.

The selection of the root bridge within a STP network is determined by the lowest Bridge ID which is a combination of configured bridge priority and embedded mac address of each device. The device with the lowest priority, or with equal lowest priority but the lowest MAC address is selected as the root bridge.

The selection of the active path among a set of redundant paths is determined primarily by the port path cost. The port path cost represents the cost of transiting between that port and the root bridge - the further the port is from the root bridge, the higher the cost. The cost is incremented for each link in the path, by an amount that is (by default) dependent on the media speed. Where two paths from a given LAN segment have an equal cost, the selection is further determined by the lowest bridge ID of the attached devices, and in the case of two attachments to the same device, by the configured port priority and port ID of the neighboring attached ports.

Once the active paths have been selected, any ports that do not form part of the active topology are moved to the blocking state.

Topology Changes

Network devices in a switched LAN perform MAC learning; that is, they use received data traffic to associate unicast MAC addresses with the interface out of which frames destined for that MAC address should be sent. If STP is used, then a recalculation of the spanning tree (for example, following a failure in the network) can invalidate this learned information. The protocol therefore includes a mechanism to notify topology changes around the network, so that the stale information can be removed (flushed) and new information can be learned based on the new topology.

A *Topology Change* notification is sent whenever STP moves a port from the blocking state to the forwarding state. When it is received, the receiving device flushes the MAC learning entries for all ports that are not blocked other than the one where the notification was received, and also sends its own topology change notification out of those ports. In this way, it is guaranteed that stale information is removed from all the devices in the network.

Variants of STP

There are many variants of the Spanning Tree Protocol:

- Legacy STP (STP)—The original STP protocol was defined in IEEE 802.1D-1998. This creates a single spanning tree which is used for all VLANs and most of the convergence is timer-based.
- Rapid STP (RSTP)—This is an enhancement defined in IEEE 802.1D-2004 to provide more event-based, and hence faster, convergence. However, it still creates a single spanning tree for all VLANs.
- Multiple STP (MSTP)—A further enhancement was defined in IEEE 802.1Q-2005. This allows multiple spanning tree instances to be created over the same physical topology. By assigning different VLANs to the different spanning tree instances, data traffic can be load-balanced over different physical links. The number of different spanning tree instances that can be created is restricted to a much smaller number than the number of possible VLANs; however, multiple VLANs can be assigned to the same spanning tree instance. The BPDUs used to exchange MSTP information are always sent untagged; the VLAN and spanning tree instance data is encoded inside the BPDU.
- Per-Vlan STP (PVST)—This is an alternative mechanism for creating multiple spanning trees; it was developed by Cisco before the standardization of MSTP. Using PVST, a separate spanning tree is created for each VLAN. There are two variants: PVST+ (based on legacy STP), and PVRST (based on RSTP). At a packet level, the separation of the spanning trees is achieved by sending standard STP or RSTP BPDUs, tagged with the appropriate VLAN tag.
- Per-Vlan Rapid Spanning Tree (PVRST)— This feature is the IEEE 802.1w (RSTP) standard implemented per VLAN, and is also known as Rapid PVST or PVST+. A single instance of STP runs on each configured VLAN (if you do not manually disable STP). Each Rapid PVST+ instance on a VLAN has a single root switch. You can enable and disable STP on a per-VLAN basis when you are running Rapid PVST+. PVRST uses point-to-point wiring to provide rapid convergence of the spanning tree. The spanning tree reconfiguration can occur in less than one second with PVRST (in contrast to 50 seconds with the default settings in the 802.1D STP).
- Resilient Ethernet Protocol (REP)— This is a Cisco-proprietary protocol for providing resiliency in rings. It is included for completeness, as it provides MSTP compatibility mode, using which, it interoperates with an MSTP peer.