
LAN Aggregation Through Switch Meshing

Technical Brief

Contents

Introduction - - - - -	page 1
Switch Meshing Fundamentals - - - - -	page 4
How Switch Meshing Operates - - - - -	page 6

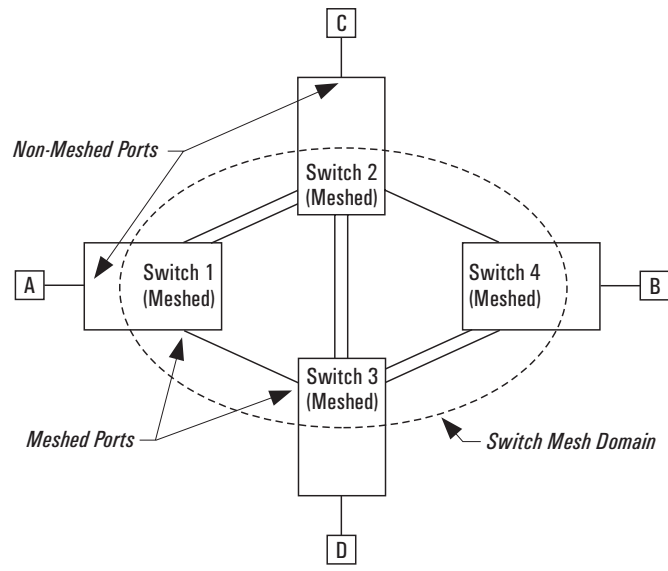
Introduction

Switch meshing is a LAN aggregation technology available in the HP ProCurve Switch 8000M and Switch 1600M for enhancing reliability and performance in these ways:

- Provides significantly better bandwidth utilization than either Spanning Tree Protocol (STP) or standard port trunking.
- Uses redundant links that remain open to carry traffic, removing any single point of failure for disabling the network, and allowing quick responses to individual link failures. This also helps to maximize investments in ports and cabling.
- Unlike trunked ports, the ports in a switch mesh can be of different types and speeds. For example, a 10Base-FL port and a 1Gps port can be included in the same switch mesh.



Figure 1. The HP ProCurve Switch 8000M and HP ProCurve Switch 1600M



The mesh-configured ports in switches 1-4 form a Switch Mesh Domain

Figure 2. Example of Switch Meshing

Finding the Fastest Path. Using multiple switches redundantly linked together to form a *switch mesh domain*, switch meshing dynamically distributes traffic across load-balanced switch paths by seeking the fastest paths for new traffic between nodes. In actual operation, the switch mesh periodically determines the best (lowest latency) paths, then assigns these paths as the need arises. The path assignment remains until the related MAC address entry times out. The mesh sees later traffic between the same nodes as new traffic, and may assign a different path, depending on conditions at the time. For example, at one time the best path from node A to node B in figure 2 is through switch 2. However, if traffic between node A and node B ceases long enough for the path assignment to age out, then the next time node A has traffic for node B, the assigned path between these nodes may be through switch 3 if network conditions have changed significantly.

Because Redundant Paths Are Active, Meshing Adjusts Quickly to Link Failures. If a link in the mesh fails, the fast convergence time designed into meshing typically has an alternate route selected in less than a second for traffic that was destined for the failed link.

Meshing Allows Scalable Responses to Increasing Bandwidth Demand. As more bandwidth is needed in a LAN backbone, another switch and another set of links can be added. This means that bandwidth is not limited by the number of trunk ports allowed in a single switch.

Mesh Domain Size. Up to 12 switches are supported in a switch mesh domain. The following figure shows a meshed backbone where the meshed switch hop count is 3.

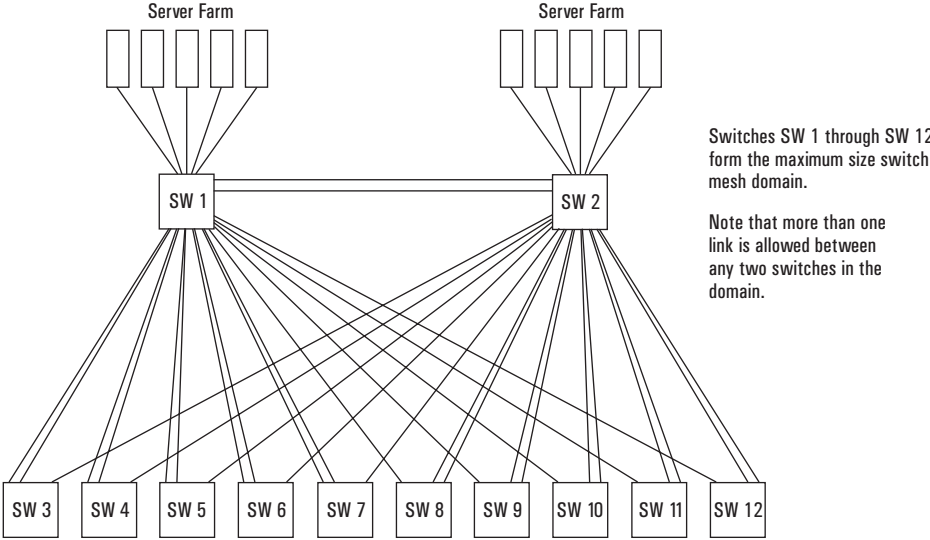


Figure 3. Example of a Backbone Using the Maximum Number of Meshed Switches

Switch Hop Count in the Mesh Domain. A maximum (meshed) switch hop count of five is allowed in the path connecting two nodes via a switch mesh domain topology. In the following illustration, the hop count would be at its maximum if, for example:

- The link between switch 1 and switch 5 failed.
- The lowest latency path was through the outside perimeter links.

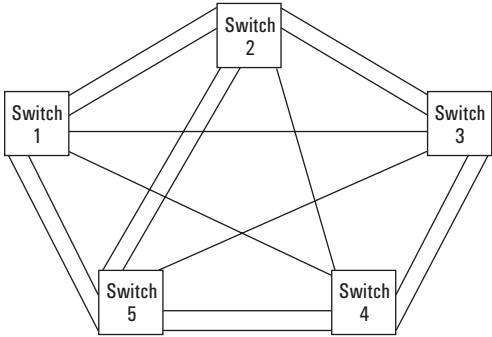


Figure 4. Example of the Maximum Meshed Switch Hop Count

Switch Meshing Fundamentals

Switch Mesh Domain. This is a group of switches exchanging meshing protocol packets. Paths between these switches can have multiple redundant links without creating broadcast storms. A meshed switch can have some ports in the mesh domain and others outside the mesh domain. Meshed links must be point-to-point switch links. Hub links between meshed switches are not allowed. Within any meshed switch, all ports belong to the same switch mesh domain.

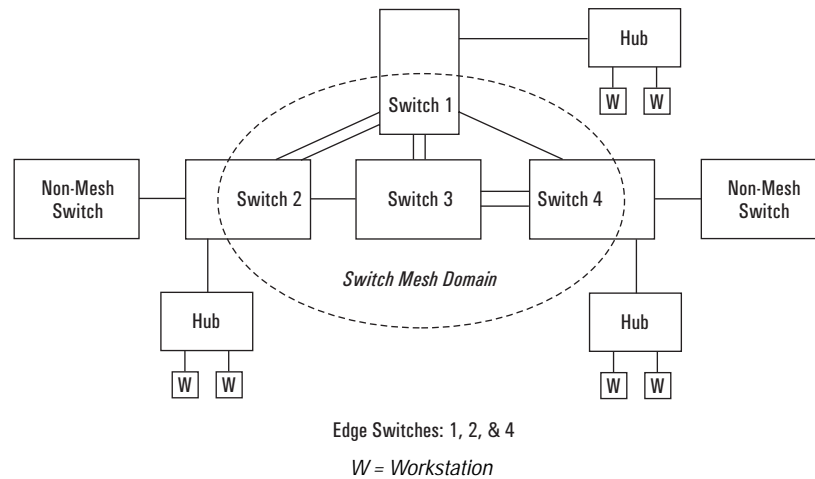


Figure 5. Example of a Switch Mesh Domain in a Network

Edge Switch. This is a switch that has some ports configured for switch meshing (and therefore in the switch mesh domain) and some ports not configured for switch meshing (and therefore outside the domain). See figure 5, above.

Non-Mesh Switch. A switch that has no ports configured for switch meshing and (therefore) is not in the switch mesh domain.

Non-Mesh Port. A switch port that is not configured for switch meshing and (therefore) is not in the switch mesh domain.

Note

Inserting a non-mesh device into the mesh causes the meshed switch port(s) connected to that device to shut down. For this reason, switch mesh domains do not allow the following:

- Hubs
- Switches that are not configured for switch meshing

Also, a switch port that is not configured for meshing (such as an unmeshed port on an edge switch) cannot be connected to a port that is configured for meshing on another switch.

Connecting Multiple Meshed Domains. Connecting two or more switch mesh domains requires separation by either a non-meshed switch or a non-meshed link. For example:

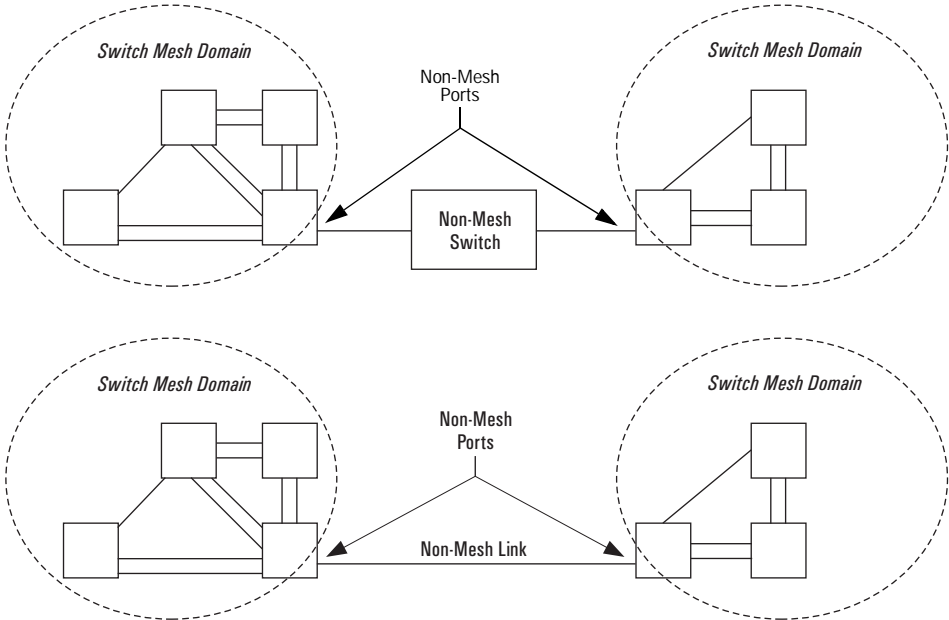


Figure 6. Example of Multiple Meshed Switch Domains Separated by a Non-Mesh Switch or a Non-Mesh Link

Bringing Up a Switch Mesh Domain: When a meshed port detects a non-meshed port on the opposite end of a point-to-point connection, the link will be blocked. Thus, as you bring up switch meshing on various switches, you may temporarily experience blocked ports where meshed links should be running. These conditions should clear themselves after all switches in the mesh have been configured for meshing and then rebooted. *To reduce the effect of blocked ports during bring-up, configure meshing either before installing the meshed switches in the network or before connecting the cables between the meshed ports.*

How Switch Meshing Operates

In a switch mesh domain, traffic is distributed across the available paths with an effort to keep latency the same from path to path. The path selected at any time for a connection between a source node and a destination node is based on these latency and throughput cost factors:

- Outbound queue depth, or the current outbound load factor for any given outbound port in a possible path
- Port speed, such as 10Mbps versus 100Mbps; full-duplex or half-duplex
- Inbound queue depth, or how busy is a destination switch in a possible path
- Increased packet drops, indicating an overloaded port or switch

Paths having a lower cost will have more traffic added than those having a higher cost. Alternate paths and cost information is discovered periodically and communicated to the switches in the mesh domain. This information is used to assign traffic paths between devices that are newly active on the mesh. This means that after an assigned path between two devices has timed out, new traffic between the same two devices may take a different path than previously used.

Flooded Traffic

Broadcast and multicast packets will always use the same path between the source and destination edge switches unless link failures create the need to select new paths. (Broadcast and multicast traffic entering the mesh from different edge switches are likely to take different paths.) When an edge switch receives a broadcast from a non-mesh port, it floods the broadcast out all its other non-mesh ports, but sends the broadcast out only those ports in the mesh that represent the path from that edge switch through the mesh domain. (Only one copy of the broadcast packet gets to each edge switch for broadcast out of its nonmeshed ports. This helps to keep the latency for these packets to each switch as low as possible.)

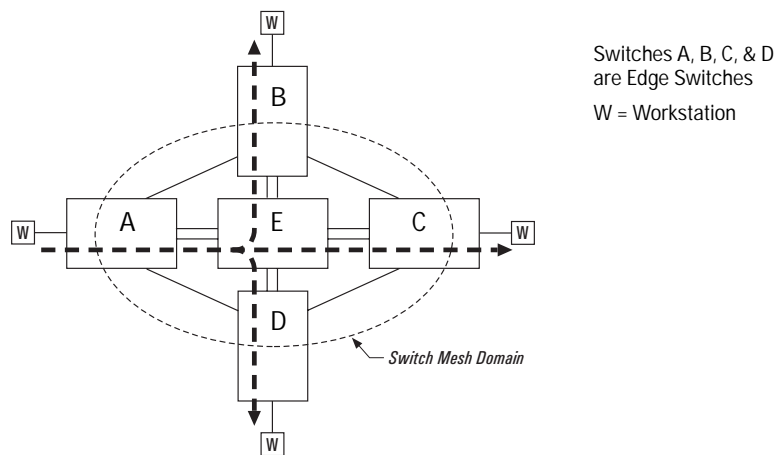


Figure 7. Example of a Broadcast Path Through a Switch Mesh Domain

Any meshed switches that are not edge switches will flood the broadcast packets only through ports (paths) that link to separate edge switches in the controlled broadcast tree. The edge switches that receive the broadcast will flood the broadcast out all non-meshed ports. Some variations on broadcast/multicast traffic patterns include:

- If multiple VLANs are configured, a broadcast path through the mesh domain leads only to ports that are in the same VLAN as the device originating the broadcast. (See “802.1Q VLANs in Meshed Switches” on page 10).
- Broadcast control features such as layer 3 (Automatic Broadcast Control) proxy replies and RIP/SAP filtering will reduce broadcast traffic by preventing the need for many of the broadcasts from crossing the switch mesh domain in the first place.

Unicast Packets with Unknown Destinations

A meshed switch receiving a unicast packet with an unknown destination does not flood the packet onto the mesh. Instead, the switch sends a query on the mesh to learn the location of the unicast destination. The meshed edge switches then send 802.2 test packets through their non-meshed ports. When the unicast destination is found and reported, the unicast packet is then forwarded through the mesh to its destination. By increasing the Address Age Interval in the switch console’s System Information screen, you can cause the switch address table to retain device addresses longer. Because the switches in a mesh exchange address information, this will help to decrease the number of unicast packets with unknown destinations, which will improve latency within the mesh. Also, in an IP environment, it is recommended that you configure meshed switches with their own IP addresses. This makes the discovery mechanism more robust, which contributes to decreased latency.

Spanning Tree Operation with Switch Meshing

As recommended in the IEEE 802.1Q VLAN standard, the HP Switch 8000M and the HP Switch 1600M use single-instance STP. Thus, these switches do not distinguish between VLANs when identifying redundant physical links. Using STP with several switches and no switch meshing configured can result in unnecessarily blocking links and disrupting traffic within a VLAN. For example:

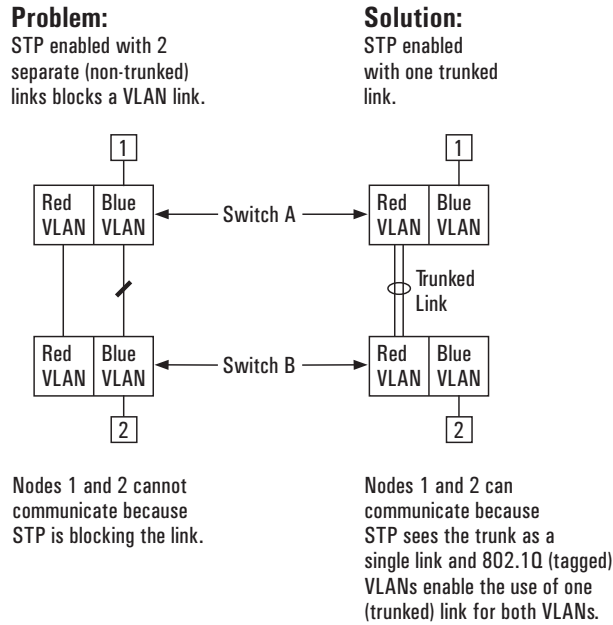


Figure 8. Example Using STP Without and With Switch Meshing

If you enable STP in a network that includes switch meshing, you should enable STP on the meshed switches as well as the non-meshed devices. STP sees a meshed domain as a single path. However, on edge switches in the domain, STP will manage non-meshed redundant links from other switches. For example:

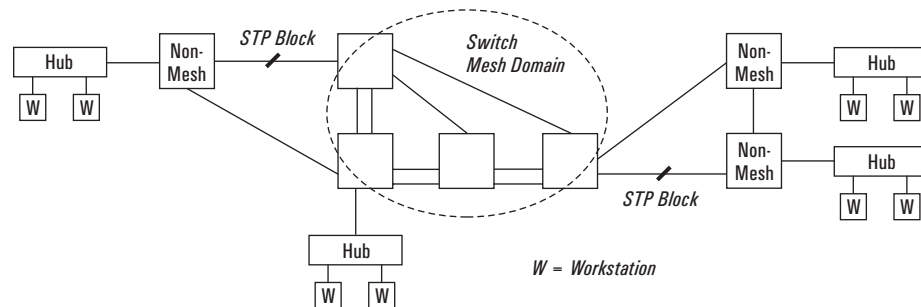


Figure 9. Connecting a Switch Mesh Domain to Non-Meshed Switches

STP should be configured on non-mesh devices that use redundant links to interconnect with other devices or with multiple switch mesh domains. For example:

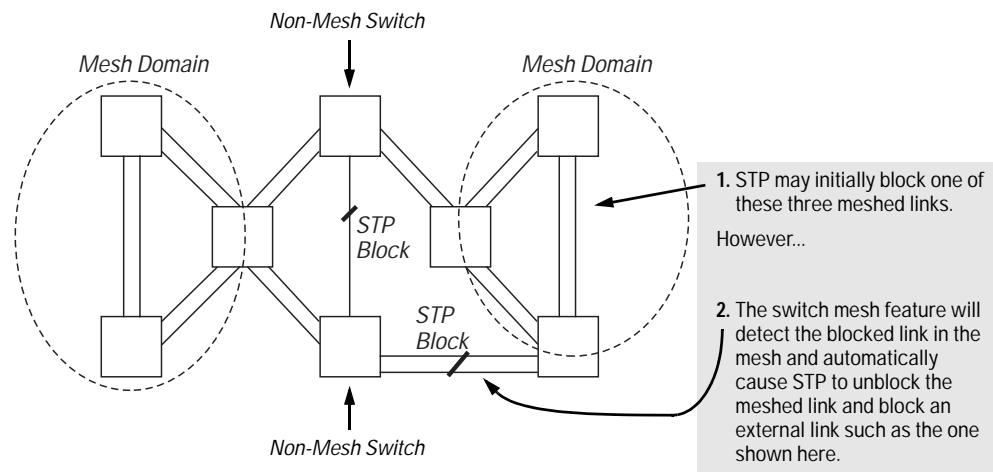


Figure 10. Interconnecting Switch Mesh Domains with Redundant Links

In the above case of multiple switch meshes linked with redundant trunks there is the possibility that STP will temporarily block a mesh link. This is because it is possible for STP to interpret the cost on an external trunked link to be less than the cost on a meshed link. However, if this condition occurs, the meshed switch that has a blocked link will automatically increase the STP cost on the external (non-meshed) link to the point where STP will block the external link and unblock the meshed link. This process typically resolves itself in approximately 30 seconds.

Filtering/Security in Meshed Switches

Because paths through the mesh can vary with network conditions, configuring filters on meshed ports can create traffic problems that are difficult to predict, and is not recommended. However, configuring filters on nonmeshed ports in an edge switch provides you with control and predictability.

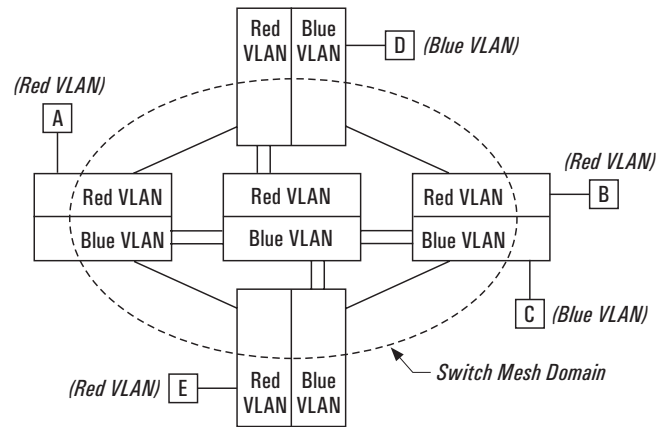
IP Multicast (IGMP) in Meshed Switches

Like trunked ports, the switch mesh domain appears as a single port to IGMP. However, unlike trunked ports, IGMP protocol and multicast traffic may be sent out over several links in the mesh in the same manner as broadcast packets.

802.1Q VLANs in Meshed Switches.

In a network having a switch mesh domain and multiple VLANs configured, all VLANs must be configured on each meshed switch, even if no ports on the switch are assigned to any VLAN. (The switch mesh is a member of all VLANs configured on the network.)

When VLANs are configured and activated, the mesh is seen as a single entity by each VLAN. All ports in the mesh domain are members of all VLANs and can be used to forward traffic for any VLAN. However, the non-mesh ports on edge switches that allow traffic to move between the mesh and non-meshed devices belong to specific VLANs and do not allow packets originating in a specific VLAN to enter non-meshed devices that do not belong to that same VLAN. (It is necessary to use a router to communicate between VLANs.) For example, in the following illustration, traffic from host A entering the mesh can only exit the mesh at the port for hosts B and E. Traffic from host A for any other host (such as C or D) will be dropped because only hosts B and E are in the same VLAN as host A.



All ports inside the mesh domain are members of all VLANs.

Figure 11. VLAN Operation with a Switch Mesh Domain

Using Automatic Broadcast Control in a Switch Mesh Environment

To avoid duplicate replies, switch meshing does not allow ABC proxy replies from within a switch mesh. However, an edge switch can learn of the existence of a device on the other side of the mesh and provide a proxy reply to inquiries. For this reason, it is recommended that you configure ABC in all meshed switches.

Requirements and Restrictions

- **Mesh Support Within the Domain:** All switches in the mesh domain, including edge switches, must support the HP switch meshing protocol. As of Summer, 1998, this includes the HP ProCurve Switch 8000M and HP ProCurve Switch 1600M models. For further information, contact your HP-authorized LAN dealer or refer to HP's Network City website at:
http://www.hp.com/go/network_city
- **Connecting Mesh Domains:** To connect two separate switch meshing domains, you must use non-meshed ports. (The non-meshed link can be a port trunk or a single link.) Refer to figure 6 on page 5.
- **Multiple Links Between Meshed Switches:** Multiple mesh ports can be connected between the same two switches to provide higher bandwidth.
- **Automatic Broadcast Control:** To use ABC with switch meshing, all edge switches in the mesh must have ABC enabled. Also, proxy replies from the switch are not sent out meshed ports.

Technical information subject to change without notice.
© Copyright 1998, Hewlett-Packard Company

June 1998

