ethernet alliance

# Provider Backbone Transport Overview

Version 1.0 December 2007


Authors:

David Allen, Nortel
Paul Bottorff, Nortel
Harpreet Chadha, Extreme Networks
Peter Lunk, Extreme Networks
David Martin, Nortel

# Introduction

Carrier Ethernet has seen tremendous growth over the last few years as service providers roll out new high bandwidth services to their customers. The applications for these services include E-Line and E-LAN business services as well as a growing trend toward using Ethernet in wireless backhaul networks. Essential features needed in Ethernet networks to support carrier applications are: scalability, reliability, guaranteed QoS, traffic management, service management and support for TDM services. Current carrier Ethernet standards, including the IEEE 802.1ah Provider Backbone Bridging (PBB) and IEEE 802.1ag Connectivity Fault Management (CFM) are addressing scalability, reliability and service management, providing new alternatives for metro networks that are based on Ethernet forwarding, simplicity and cost curves.

This paper discusses a new technology called Provider Backbone Transport (PBT) which can be employed within the service provider domain of a Provider Backbone Bridged Network (PBBN) to allow configuration of resilient, traffic engineered, SLA driven, point-to-point Ethernet trunks facilitating guaranteed QoS. These PBT trunks allow carriers to leverage their transport infrastructure experience and management systems for connection-oriented circuits, while operating in parallel with traditional connection-less Ethernet in the PBBN. PBT trunks can use the CFM integrity check toolkit in conjunction with route diverse protection paths, as well as network load balancing, to provide performance guarantees and sub-50ms protection switching.

# Provider Backbone Transport Overview

In a standard PBBN, traffic engineering is limited as a consequence of the IEEE 802.1Q Multiple Spanning Tree Protocol (MSTP) control plane which controls the population of the bridge filtering tables. However, the underlying data plane of the IEEE 802.1Q/802.1ad/802.1ah bridge relay function has no inherent limiting characteristic to prevent full traffic engineering. For example, the IEEE 802.1aq Shortest Path Bridging (SPB) project is improving link utilization by replacing the MSTP control plane with a shortest path tree control plane.
PBT is a method for providing full traffic engineering of point-to-point

Ethernet paths within a PBBN. To accomplish this, PBT operates side by side with the MSTP or SPB control plane using either an external management or control plane which populates the bridge filtering tables of the component Backbone Edge Bridges (BEB) and Backbone Core Bridges (BCB) relays by creating static filtering table entries (see Figure 1 below).

The ability of PBT to utilize an external management or control plane is facilitated by IEEE 802.1ah because the backbone Media Access Control (MAC) addresses within the PBBN are all managed by the service provider, allowing them to be readily discovered and identified in the service provider's topology. The external PBT management or control plane is responsible for maintaining and controlling all the topology information to support the point-to-point unidirectional Ethernet Switched Paths (ESPs) across the PBBN.

The PBT topology can co-exist with the existing active MSTP or SPB topology by allocating separate Backbone Virtual LAN Identifier (B-VID) ranges to PBT, MSTP, or SPB. Alternatively, PBT could simply be assigned the entire B-VID space and standalone. The PBT management or control plane then takes ownership of its assigned range of B-VIDs from the BEBs and BCBs comprising the PBBN.

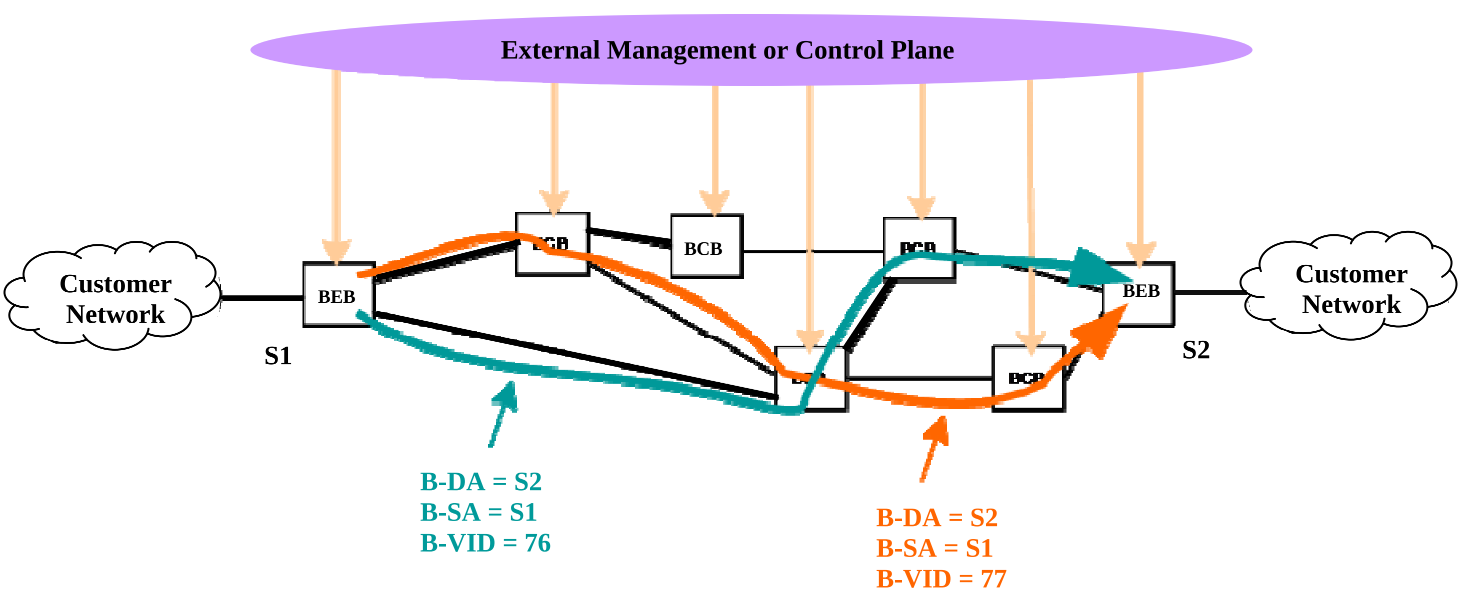


**Figure 1— PBT Network**

The PBT management or control plane forms a topology of backbone destination address (B-DA) rooted trees. For each <B-DA, B-VID> tuple configured by PBT an independent tree is maintained (see Figure 2 below). An ESP is routed by PBT from the leaf BEB (Backbone Source Address, B-SA) along a tree selected by the B-VID to the destination root BEB (B-DA). For a single B-DA, the number of separate routing trees may be up to the number of B-VIDs allocated for PBT. Those B-VIDs may be reused for every B-DA, providing a maximum total number of routing trees equal to the number of B-DAs multiplied by the number of B-VIDs allocated for PBT.

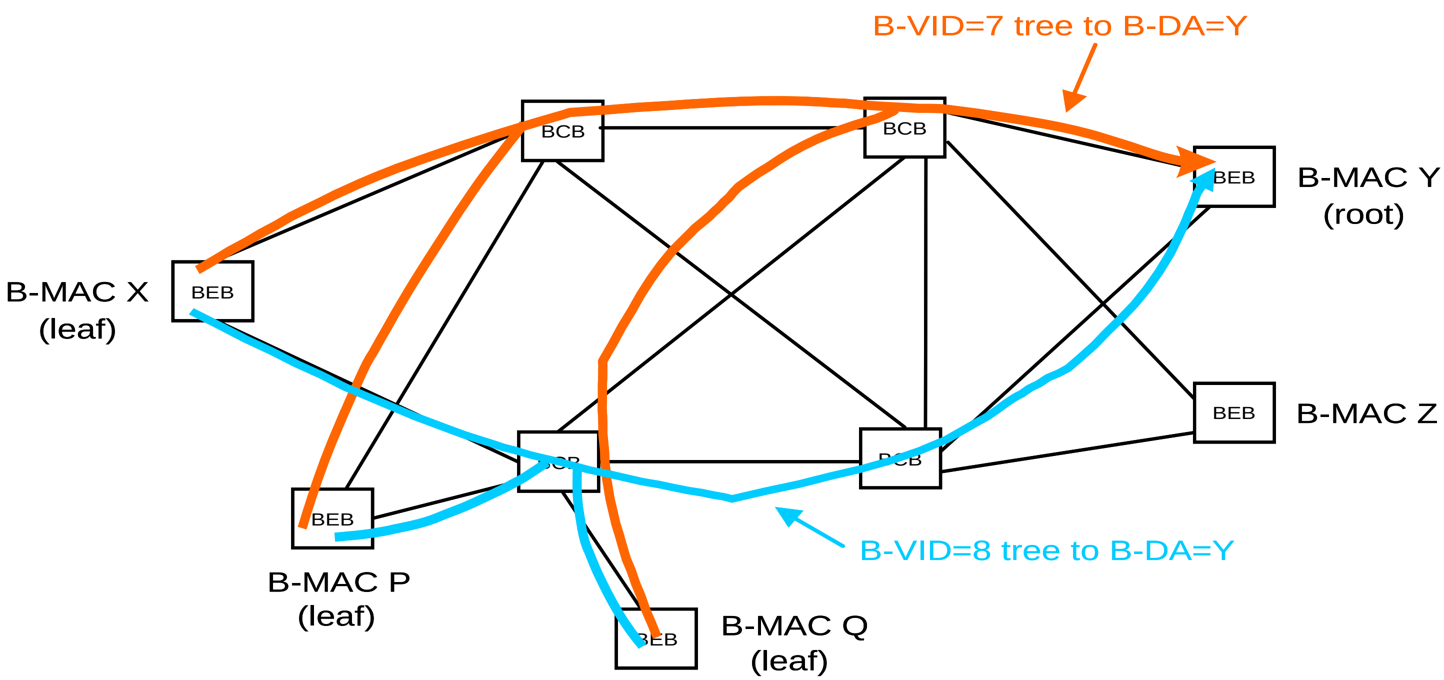


**Figure 2— PBT Tree Topology**

The trees maintained by PBT for routing ESPs do not have to span all the BEBs in a PBBN since they only require connectivity to the source leaves (B-SAs) which have ESPs to a specific backbone destination root (B-DA). For example, in Figure 2 neither tree includes BEB Z. Each tree may connect to as many source BEBs as desired, the only limit being implementation imposed table sizes.

The PBT management or control plane may use any algorithm desired to select the path for a routing tree, thereby providing complete route selection freedom. The PBT management or control plane also manages the bandwidth of all

ESPs along each routing tree. For each leaf BEB which is part of a routing tree maintained by the PBT management or control plane, PBT will maintain a routing tree which provides a co-routed reverse path from the root BEB (B-DA) to the leaf BEB (B-SA). The B-VID used in this reverse ESP does not have to be the same one used for the forward ESP. For example, in Figure 3 below $ESP_{Y7}$ follows the B-VID 7 tree to BEB Y, while in the reverse direction $ESP_{X10}$ follows the B-VID 10 tree back to BEB X.

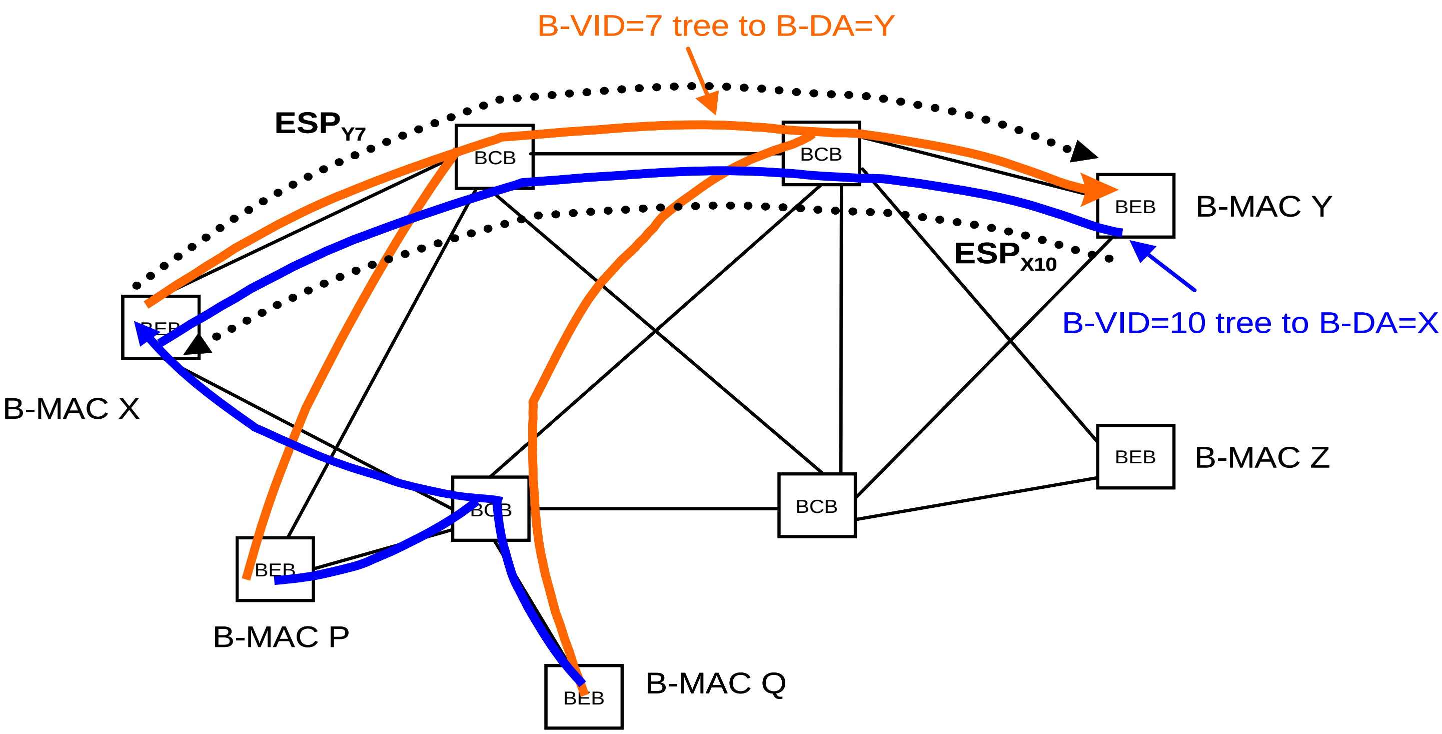


**Figure 3— PBT ESP Pair Over Routing Trees**

The relay function used by PBT is standard Independent VLAN Learning (IVL) bridge forwarding, using a <B-DA, B-VID> tuple look-up for filtering and forwarding decisions, where only the semantics of the tuple are different. The combined <B-DA, B-VID> tuple is treated as a single 58-bit address, where 12 bits are the B-VID and 46 bits are the B-DA (i.e., of the 48 bits in a DA, one bit is used to indicate individual or group address and one bit is used to indicate locally or globally administered address, leaving 46 bits). This fact allows PBT to consider the B-VID part of the address as a path selector to the B-DA, rather than as a B-VLAN ID, allowing up to 4094 unique routing trees to any single B-

DA (i.e., of the $2^{12}$ = 4096 values, 0 and 4096 are reserved, leaving 4094 values). Typically only a small number of B-VIDs are needed for PBT since it is normally not necessary to have even tens of alternate paths to a single destination. In Figure 1 above, two paths are configured to reach S2. The two paths are differentiated by using a different B-VID in combination with the same B-DA. Across a PBT network, the theoretical maximum number of routing trees is about $2^{58}$.

PBT requires no B-VID translation. The forwarding of frames over ESPs is easily achieved by most existing Ethernet bridging equipment without significantly re-specifying the hardware or management. Only three changes are necessary to 802.1Q to support PBT forwarding. Those modifications are:

1. The B-VID address space must be divided between MSTP, SPB and PBT. Spanning tree or multiple spanning trees are used only for the non-PBT-assigned B-VID range.

2. It must be possible to discard frames with unknown unicast destination addresses, group addresses and broadcast traffic, rather than flooding them on PBT B-VIDs. As there is no loop free topology for the assigned PBT B-VID range, flooding would result in unbounded looping and replication.

3. The PBT port state must be set to forwarding only and not learning. B-SA learning is not required, and may interfere with the external management or control plane population of the forwarding tables. For this reason B-SA learning is disabled for the assigned B-VID range.

When PBT is operated on a PBBN where MSTP or SPB is run in parallel, the PBT traffic on Connection Admission Controlled (CAC'd) ESPs requires a higher priority than the best effort spanning tree traffic. Also, when engineering such a mixed network, reserve bandwidth must be set aside for the best effort traffic to allow for spanning tree reconfiguration.

The PBT approach has several useful properties:

The use of a global path identifier (i.e., the <B-DA, B-VID> tuple) for forwarding, with no intermediate label-like translation, is inherently more robust than alternatives. Any mis-configuration or forwarding table error resulting in the deviation of a frame from the intended path will self-identify immediately. There is no possibility of collision with other identifier spaces that can mask the fault. This also suggests that minimal changes are required to existing CFM constructs to successfully instrument ESPs.

The ability to explicitly route and pin paths across the network can be combined with connection admission control and 802.1Q class-based queuing in order to provide guaranteed per-path QoS. The connection admission control function can be enforced by the external management or control plane without any changes to existing Ethernet bridges.

# Ethernet Switched Paths

Figure 4 below shows an example PBBN running PBT. The PBBN B-VID space has been partitioned between MSTP and PBT, with B-VIDs 7 and 8 allocated to PBT. PBT has taken over the port forwarding state machines in the BEBs and BCBs for the allocated B-VID range and controls frame forwarding by adding static entries to the filtering databases of those bridges. MSTP operates normally in parallel to PBT on the other B-VIDs.

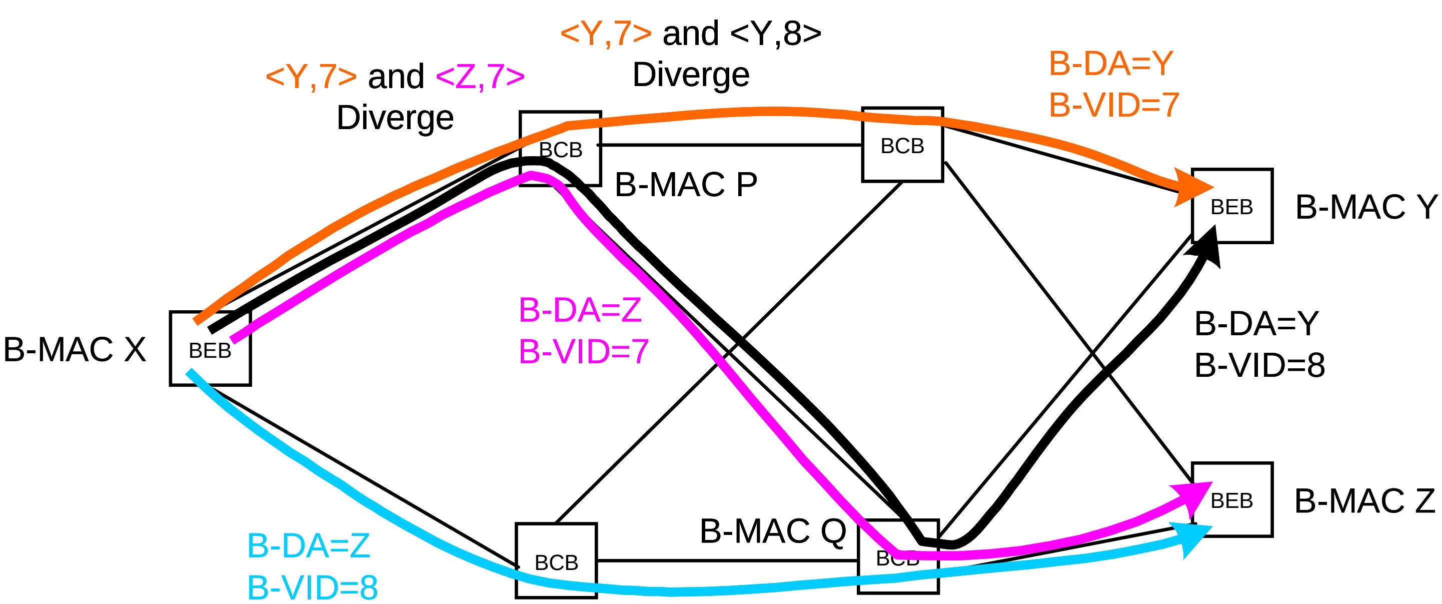


**Figure 4— - PBBN with B-VIDs 7 and 8 allocated for PBT**

As explained above, the B-VIDs allocated to PBT are not treated as B-VLAN IDs, instead they are considered individual path identifiers for one of a maximum of ‘n’ possible routing trees (where ‘n’ is the number of PBT-allocated B-VIDs) to the destination B-DA. B-VIDs in the allocated range may be re-used for other routing trees within the PBBN using different B-DA values (i.e., as long as the <B-DA, B-VID> tuple is unique).

Figure 4 illustrates the complete route freedom of configured forwarding in the PBBN bridges. In the example, a total of 4 ESPs use 2 B-VIDs to forward traffic to 2 B-DAs. At node 'P' above, despite collisions in both the B-DA and the B-VID space, the forwarding properly resolves because the BCB uses the B-DA and B-VID bound together to establish route uniqueness. At node ‘P’ the orange and violet ESPs diverge even though they have the same B-VID because they are addressed to different B-DAs. Likewise at node ‘P’ the orange and black ESPs diverge even though they have the same B-DAs because they have different route selector B-VIDs.

# Fault Management for PBT

The IEEE 802.1ag CFM standard defines three mechanisms: Continuity Check Messages (CCMs), Loopback Messages/Replies (LBMs/LBRs) and Link Trace Messages/Replies (LTMs/ LTRs). The 802.1ag CCMs use a group address, therefore such frames would be discarded by bridges supporting a PBT trunk (c.f., 802.1Q modification #2 above). To support CCMs on a PBT trunk, a provision must be added for unicast addressed CCMs so they are forwarded just like PBT data frames. The ITU-T Y.1731 standard already specifies unicast addresses for CCMs. This definition will need to be added to 802.1Q for PBT CCM support.

Unicast LBMs and LBRs are used for isolation of link faults. Modifications will be necessary in order to target PBT intermediate maintenance points (i.e., MIPs). The 802.1ag LTMs use a group address, therefore such frames would also be discarded by bridges supporting a PBT trunk. In order to support LTMs on a PBT trunk, a provision must be added for unicast addressed LTMs.

# Protection Switching for PBT

Protection switching is also an important management feature of PBT. PBT may provide 1:1 protection by configuring an alternate path with a different B-VID. Both the working and protection paths are constantly monitored using unicast CCMs. The repetition rate of the transmitted CCMs can be configured to meet the desired failure detection time. For example, a CCM repetition interval of 3.3ms on the source BEB would allow for a similar failure detection time as in SONET/SDH systems (i.e., the destination BEB would declare a failure following the loss of 3 CCMs in a row). A failure of the working path results in the data stream being transferred to the protection path by using the alternate B-VID. Referring to Figure 2, a working path from BEB X to BEB Y could follow the B-VID 7 tree, with an associated protection path along the B-VID 8 tree.

When the external management or control plane selects the protection path, care must be taken to ensure the protection path is route diverse from the working path. In Figure 2, a protection path from BEB Q to BEB Y along the B-VID 8 tree is not route diverse from a working path along the B-VID 7 tree, as both paths share a common link to the first BCB.

# Provider Backbone Bridge Traffic Engineering (PBB-TE)

As a result of industry interest in PBT, a new standards project was launched in early 2007 within IEEE called P802.1Qay, Provider Backbone Bridge Traffic Engineering (PBB-TE). The scope of this new project includes the changes to 802.1Q forwarding, the CFM extensions, and support for 1:1 protection switching.

# Conclusion

PBT decouples the 802.1Q data plane from the control plane, allowing an external management or control plane to establish ESPs by populating bridge forwarding tables. The forwarding function uses the standard IVL bridge <B-DA, B-VID> tuple, with a different semantic which identifies a B-DA rooted tree. This provides for massive scalability of paths across a PBT network.

External management or control plane support for connection admission control in conjunction with existing 802.1Q class-based queuing can be used to provide guaranteed QoS.

With only minimal extensions, CFM can operate over PBT and be leveraged for use in 1:1 protection switching of route diverse ESPs. Such enhancements are already underway in the IEEE PBB-TE project.

The traffic management, guaranteed QoS and resiliency provided by PBT will allow carriers to build Ethernet networks capable of delivering premium services and generating new revenue streams, complementing current best effort service offerings, over a single PBBN. In an effort to accelerate the adoption of this new technology, Nortel, Extreme Networks and others are participating in the Carrier Ethernet Ecosystem. This ecosystem brings together industry hardware and software vendors as well as component and tool suppliers to create interoperable Carrier Ethernet solutions.

# Glossary

| | |
|---|---|
| B- | Backbone- (e.g., B-DA, B-SA, B-MAC, B-VID) |
| BCB | Backbone Core Bridge |
| BEB | Backbone Edge Bridge |
| CAC | Connection Admission Control |
| CCM | Continuity Check Message |
| CFM | Connectivity Fault Management |
| DA | Destination Address |
| ESP | Ethernet Switched Path |
| IVL | Independent Virtual LAN Learning |
| LAN | Local Area Network |
| LBM | Loopback Message |
| LBR | Loopback Reply |
| LTM | Linktrace Message |
| LTR | Linktrace Reply |
| MAC | Media Access Control |
| MIP | Maintenance domain Intermediate Point |
| MSTP | Multiple Spanning Tree Protocol |
| PBB | Provider Backbone Bridge |
| PBBN | Provider Backbone Bridge Network |
| PBB-TE | Provider Backbone Bridge Traffic Engineering |
| PBT | Provider Backbone Transport |
| QoS | Quality of Service |
| SA | Source Address |
| SLA | Service Level Agreement |
| SPB | Shortest Path Bridging |
| VID | Virtual LAN Identifier |