> BUSINESS MADE **SIMPLE**

PAR for Provider Backbone Transport

Paul Bottorff pbottorf@nortel.com
David Martin
Glenn Parsons
Dinesh Mohan
July, 2006

NORTEL

Agenda

> Reference:http://www.ieee802.org/1/files/public/docs2006/ah-bottorff-pbt-v1-0506.pdf

> PBT Overview & Value Proposition

> Services supported with PBT

> Connection management with Protection using OAM (802.1 ag)

> What Would the IEEE Need To Do to Support PBT

> PAR Proposal

Carriers Need To

- Build packet based infrastructure allowing efficient support of packet data, voice, and video applications
- Minimize capital costs to allow major build outs
- Maximize facility utilization while providing high quality service
- Minimize operational costs
- Verify service level agreements
- Minimize backward compatibility issues

Filling Ethernet Gaps

> Traffic engineering requires the ability to route traffic in diverse ways to allow full utilization of carrier facilities

> Traffic engineering requires the ability to perform constraint based route management and admission control for service guarantees

> Protection requires the ability to recover some services very quickly. Many networks requirement less than 50 msec from the time of the failure to complete restoration.

> Protection must support traffic engineering, must be able to protect the full QoS guarantees, must be constantly monitored.

Pt-Pt and MPt have different requirement

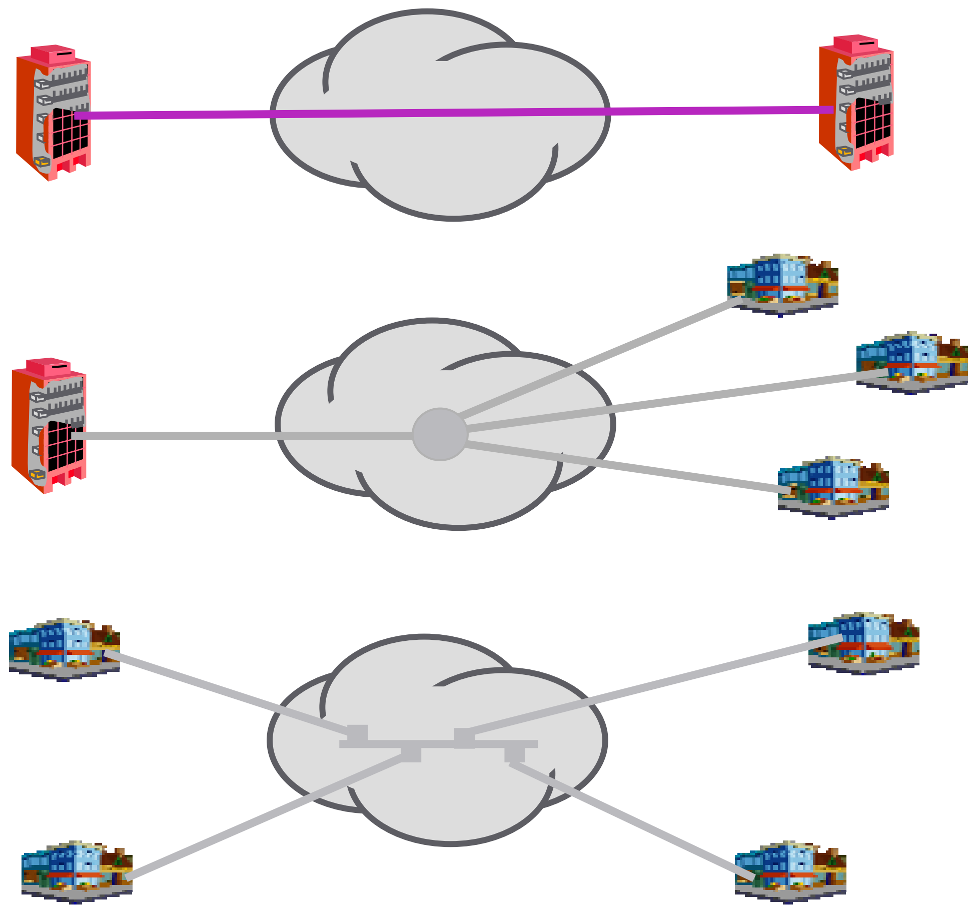
> Most current demand for traffic engineering is for engineered Pt-Pt services used for various types of leased line and trunk replacement.

> The advanced work in progress at MEF on metrics for multipoint traffic engineering will probably result in different network requirements than classic Pt-Pt traffic engineering.

> The current 802.1ah/ad models allow engineering enough multipoint circuits using by managing B-VLANs for metro video distribution.

Focus On Traffic Engineerging for E-LINE

MEF Ethernet Virtual Connections (EVCs)



E-LINE
Router Mesh

Pt-Pt, Like
Duplex Ethernet
Any-to-any

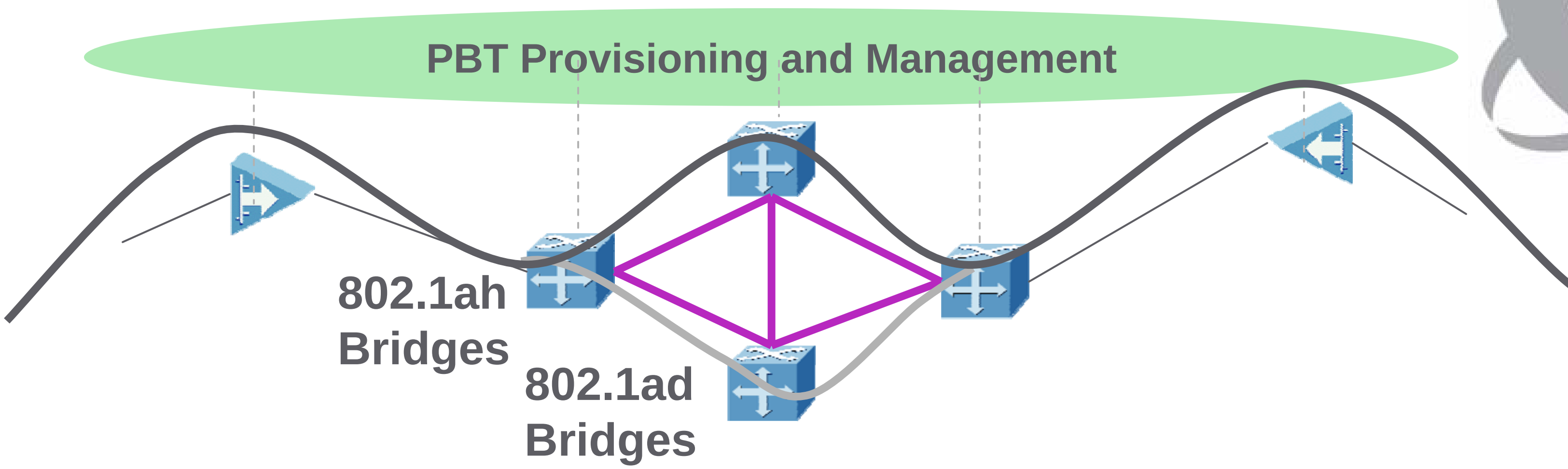
E-TREE
Hub & Spoke

Pt-MPt, Like
EPON Ethernet,
Root-to-Leaf and
Leaf-to-Root

E-LAN
Multi-Site

MPt, Like VLAN,
Any-to-any

A Provisioned P2P Ethernet Transport



PBT is a variation on Provider Backbone Bridging which allows carriers to provision engineered and protected Pt-Pt service instance.

PBT operates by adding configured routes to a nearly standard Provider Backbone Bridged Network. The PBT provisioning and management system allows a carrier to provision point-to-point trunks and services within the Ethernet network. Each trunk is identified by a 16 bit VLAN ID and a 96 bit source/destination address pair.

PBT Basic Concepts

> Divide the B-VID address space between conventional 802.1ah PBBN B-VLANs and PBT.
  - PBBN must operate in Independent VLAN Learning (IVL) mode
  - The number of B-VIDs used for PBT must be at least 2
  - B-VIDs not assigned to PBT operate as normal

> Turn off learning and broadcasting on all PBT B-VIDs
  - On PBT B-VIDs replace flooding of unknown frames with discarding for unknown frames
  - On PBT B-VIDs replace multicasting/broadcasting of frames with discarding of multicast/broadcast frames

> Use a provisioning/management system to configure the Bridge forwarding tables for PBT B-VIDs
  - These are accessed through the bridge MIB

> Each PBT circuit is composed of a working and a protection path
  - The working an protected paths use different B-VIDs to access the same backbone MAC address

> Manage co-routed bundles of PBT backbone circuits using 802.1ag
  - Requires unicast CC messages not currently implemented
  - Management must operate on both the working and protected paths

What do you get?

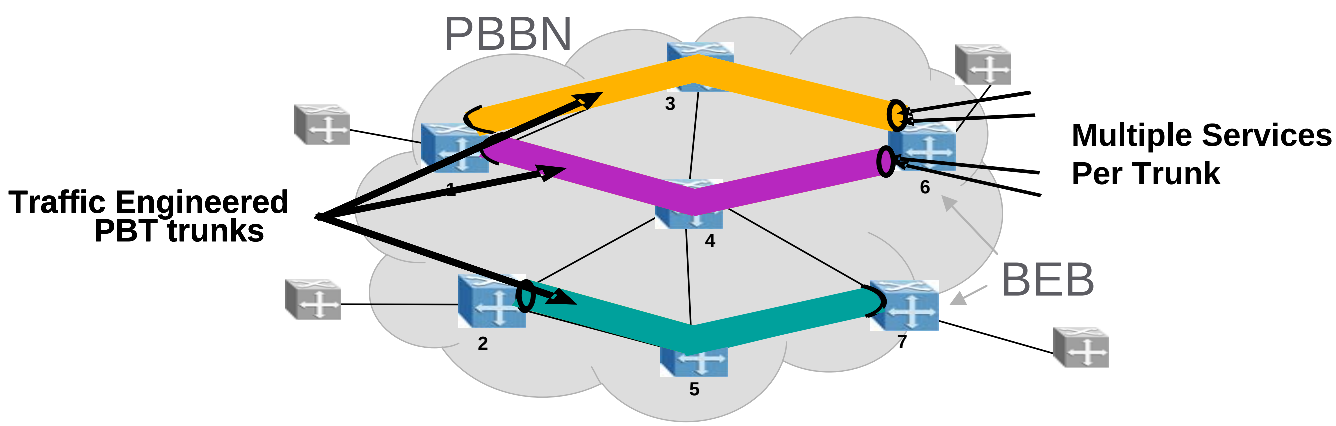
> Complete route selection freedom for PBT P2P trunks
  - Each P2P trunk may be along a different spanning tree
  - Provisioning systems may use shortest path, constraint based, manual placement, or any other route algorithm which assures loop free paths.
  - Each P2P trunk may use a different routing strategy.

> Load may be calculated for each P2P trunk and allocated to each physical link, port, and switch.

> Protection paths are pre-determined to allow rapid failover.

> Both working and protection trunks are constantly monitored.

Provider Backbone Transport (PBT) Values



> PBB provides customer-carrier isolation (encapsulation), mpt, pt-mpt, and pt-pt services based on B-VLAN tunnels routed by MSTP, and defines a management domain

> PBT is a feature added to a PBBN supporting engineered pt-pt trunks. These trunks are used in place of B-VLANs to carry pt-pt, pt-mpt, or mpt services.
- Removes constraint of following MSTP topology for path engineering
- Provides bandwidth management allowing traffic engineering over path
- Any number of 802.1ah services may be carried over a PBT trunk

> Both PBB and PBT use 802.1Q MAC address based relays.

> Leverage Emerging standards for:
- Provider Isolation – PBB (802.1ah)
- OAM & Protection(802.1ag & ITU-T Y.1731/G.8031)
- PWE Carriage (IETF Dry Martini)
- Management system or GMPLS for provisioning
- PBT is currently on the living list at ITU-T SG15 PBT architecture description G.pbt

PBT Values

- > **Traffic Engineering**
  - Control of routing
  - Admission control / policing
- > **Connectivity monitoring (IEEE 802.1ag)**
  - Strong service management when coupled with Traffic Engineering & Service Assurance
- > **Stronger resiliency and Protection (w/ ITU-T Y.1731 / G.8031)**
  - Eliminates Spanning Tree for tunnels
- > **Easier management fit with auto-discovery (w/ IEEE 802.1AB)**
  - Fits with current transport operational model
- > **Clear profit/business case for any given service**
  - Can map a service to its path & resources
- > **Simply Scalable**
- > **Reuses existing hardware and standards**
  - Maximizes the potential of today's ethernet forwarding hardware
  - Many IVL switches require only software changes to support PBT!
    12-bit VLAN/Route Discriminator & 48 bit global address

Delivers a lowest cost, dependable, easy-to-manage infrastructure

PBT MAC Forwarding

> PBT frames are forwarded based on B-DA MAC + B-VID using 802.1Q bridge relays just like normal bridge frames.

> The B-DA determines the BEB destination as normal.

> The B-VID determines the route tree for this B-DA MAC.
- No VLANs exist in the PBT domain.
- This is the key conceptual difference. The B-VID does not determine a VLAN, instead B-VIDs select a path for a destination.
- The conceptual change does not change the operation of a 802.1Q relay. It is just conceptual!

> Each B-VID is reused for path selection for each B-DA. The total number of B-VIDs required is the number of independent paths needed to each destination.

> Things which are different than a standard 802.1Q ports and relay.
- PBT must have control of a PBT B-VID address space and the port state for this B-VID address space.
- PBT ports must start up with learning off and forwarding on.
- Unknown or broadcast frames received at any PBT port must be discarded, not flooded.

> The PBT relay scales just as the Ethernet relay. It is possible to have 70 trillion destinations each with as many as 4094 paths.
- Each relay carries a filtering table only as large as the number of PBT trunks passing through its relay.

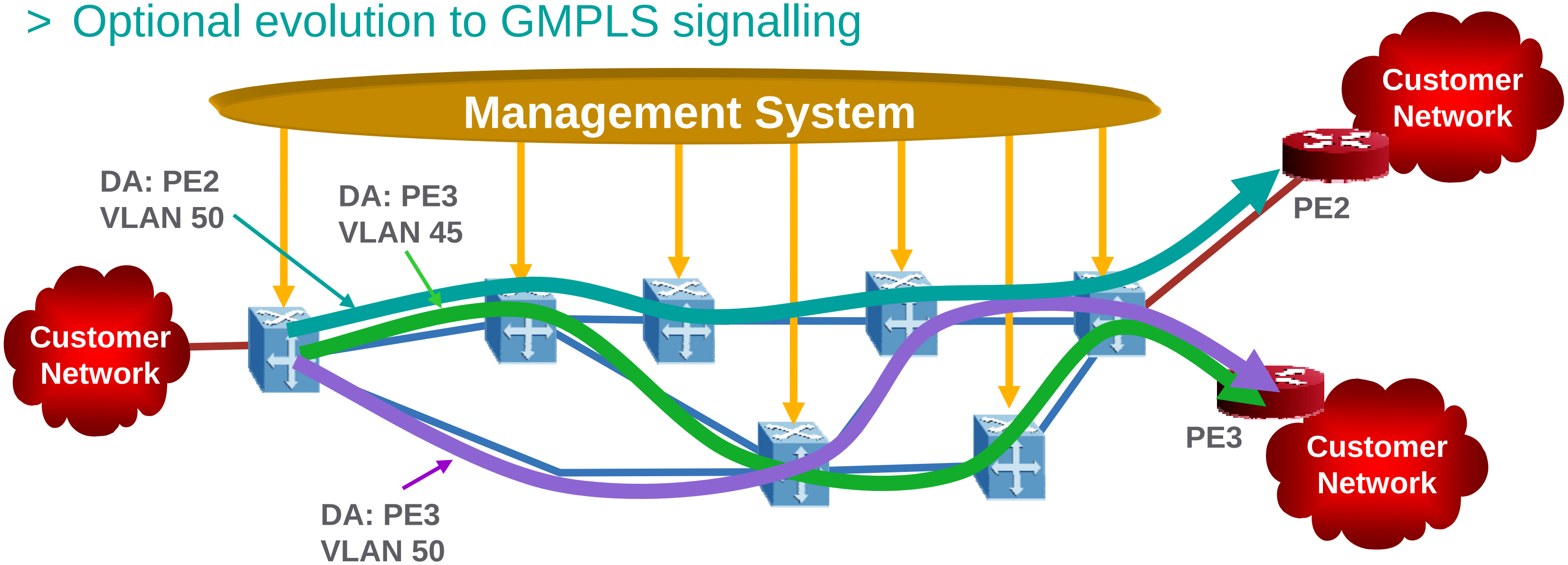
PBT Control and Management

> So we split the B-VIDs, turn off MAC learning, Broadcast Unknown, and STP
  - Use PBB hierarchy to separate customers from the Provider network, and add hierarchical dataplane OAM for instrumentation and protection.

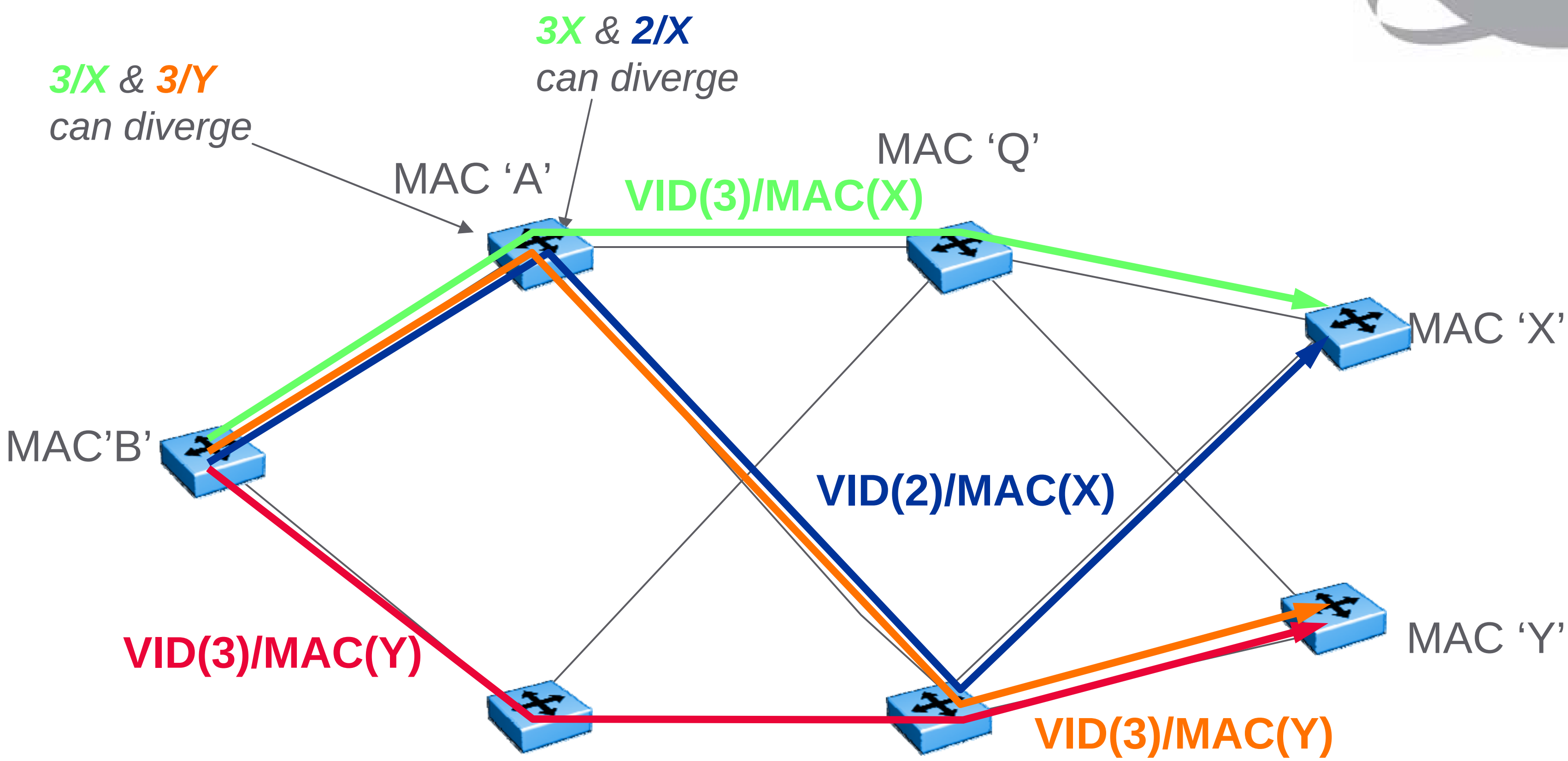
> Place under a Comprehensive Management system
  - Use a base spanning tree to control switches using SNMP
  - Management build complete topology model using auto-discovery base on 802.1AC

> Management sets up connections, populating switch bridging tables:
  - The VLAN tag is no longer a network global : scaling issues are removed;
  - VLAN tags now used to set up per destination alternate paths
  - A range of VLANs can be used for bridging and another range for PBT

> Optional evolution to GMPLS signalling



Dataplane Example



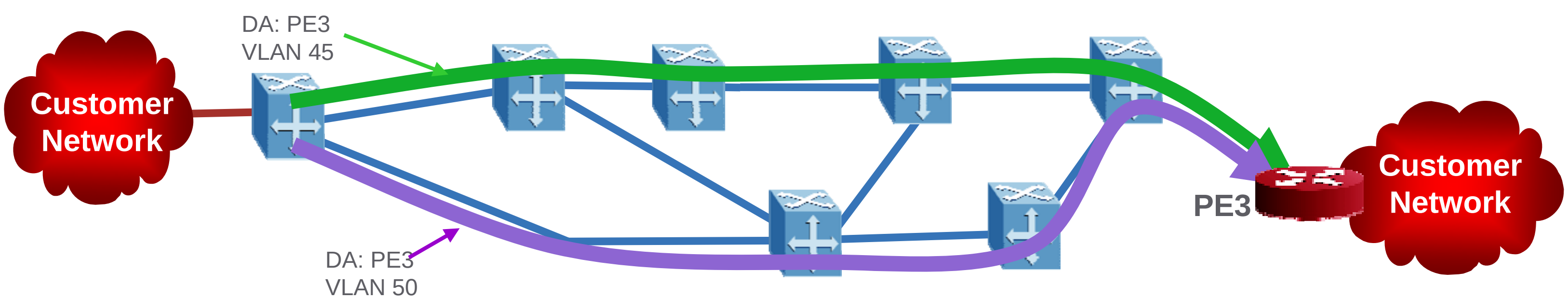
Note that MACs and VIDs can overlap, it is the combination of both that is unique and allows diverse routing

QoS and Resiliency

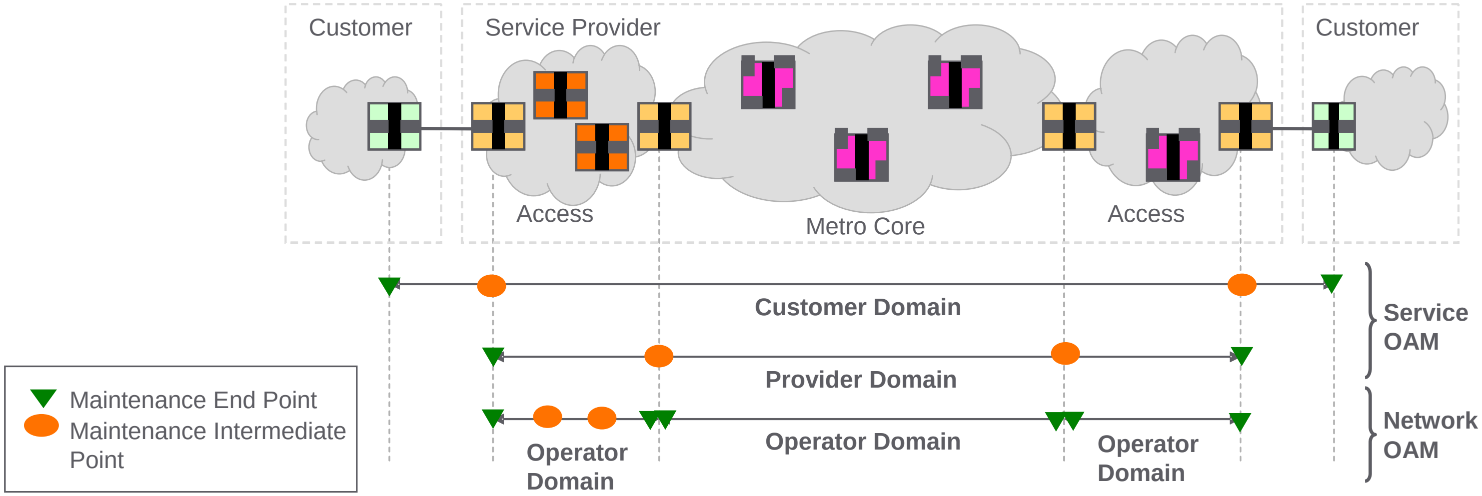
> Bandwidth can be reserved for the tunnel at each end point
  - The management system (or external control plane) does bandwidth allocation for the PBT trunk and each service over the trunk.
  - The bridges just forward frames they do not need any additions

> Ethernet VLAN "p" bits for differentiated services
  - One tunnel can provide per packet CoS
  - Can also support per packet pre-emption for resiliency

> Backup Trunks can be pre-provisioned for redundancy
  - Defined in G.8031 (Ethernet PS coordination) - *ITU SG15/Q9*
  - Ethernet CFM provides fault notification in millisecond time frames
  - Synchronizes PS state at both ends of a path
  - PS type (1+1, 1:1, etc.)
  - Administrative state (what is working, manual switch etc.)
  - Administrative control (force switch, revertive/non-revertive etc.)
  - Primary utility for maintenance operations



PBT OAM Key Principles



> **PBT can reuse all the Ethernet OAM initiatives in the IEEE and ITU**

- Fault detection and notification (IEEE 802.1ag)
  > **CFM hierarchy** (IEEE 802.1ag)
- Service Monitoring and performance (ITU-T Y.1731)
- Resiliency and Protection switching (ITU-T G.8031)
- Link layer discovery (IEEE 802.1AB)

> **Each PBT packet is self identifying**

- Where did it originated (SA MAC)
- Where is it going (DA MAC)
- Which maintenance level is it (CFM)
- What action/functionality does this frame represent.

> **No need to involve an unreliable control plane**

- MPLS OAM relies on control plane
  - Determinism? Scalability?

IEEE 802.1ag & Y.1731 Scope

> Scope

- Y.1731 scope is Ethernet OAM (FM & PM) in Carrier Networks
  - Y.1731 is now an approved recommendation
- 802.1ag scope is CFM (FM sub-set) in Enterprise & Carrier Networks

> Functionality addressed

- Common in 802.1ag and Y.1731:
  - Fault - CCM, LBM/LBR, LTM/LTR
- Additional functionality in Y.1731:
  - Fault - AIS, LCK, TST, APS
  - Performance – CCM (with PM), LMM/LMR, 1DM, DMM/DMR
  - Others - MCC, EXM/EXR, VSM/VSR

> Alignment

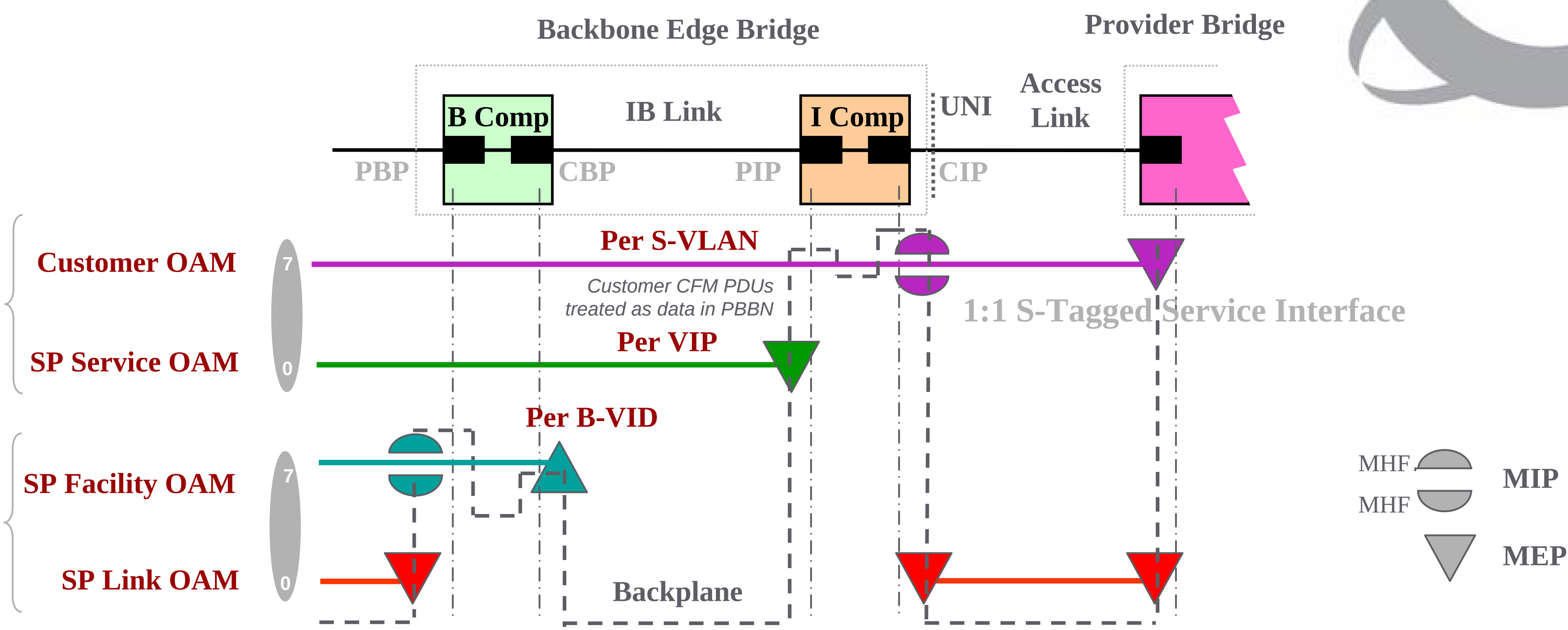
- OpCode space has been shared
  - IEEE 802.1 (values - 0-31, 64-255)
  - ITU-T Y.1731 (values - 32-63) for functionality outside the common subset
- Type codes have been shared
- PDUs for common OpCodes are completely compatible

IEEE 802.1ag and Y.1731 Mapping

> Common Protocol Mapping

Functions	802.1ag Clause	Y.1731 Clause
Continuity Check	20.1	7.1
Unicast Loopback	20.2	7.2.1
Linktrace	20.3	7.3
Protocol Formats	**802.1ag Clause**	**Y.1731 Clause**
CCM	21.6	9.2
LBM	21.7	9.3
LBR	21.7	9.4
LTM	21.8	9.5
LTR	21.9	9.6
Addressing	**802.1ag Clause**	**Y.1731 Clause**
CCM	8.13.11	10.x

PBB & PBT Maintenance Domains



> PBB constitutes:
  - Service Layer (identified by I-SIDs)
  - PBB Tunnels (identified by B-VID)
  - PBB Connections within PBB Tunnels (identified by B-MAC pairs)

> PBT constitutes:
  - PBT Trunks (identified by B-VID + B-MAC)
    - Different services can be carried across PBT Trunks e.g. I-SIDs

While PBB Maintenance Domains constitute both Service and Facility, PBT Maintenance Domains constitute PBT Trunks

OAM Functions for PBT

> Since PBT constitutes provisioned p2p trunks, similar to today's SONET/SDH connections, OAM requirements across PBT trunks are a profile of overall Ethernet OAM (which caters to p2p and mp across transport and enterprise environments). This profile includes:

- Continuity Checks (CCMs)
- Connectivity Verification (Loopbacks)
- Performance Monitoring e.g.
  - Frame Loss
  - Frame Delay
- Protection Switching

> OAM mechanisms required for PBT are already specified, though not entirely within IEEE 802.1, e.g.

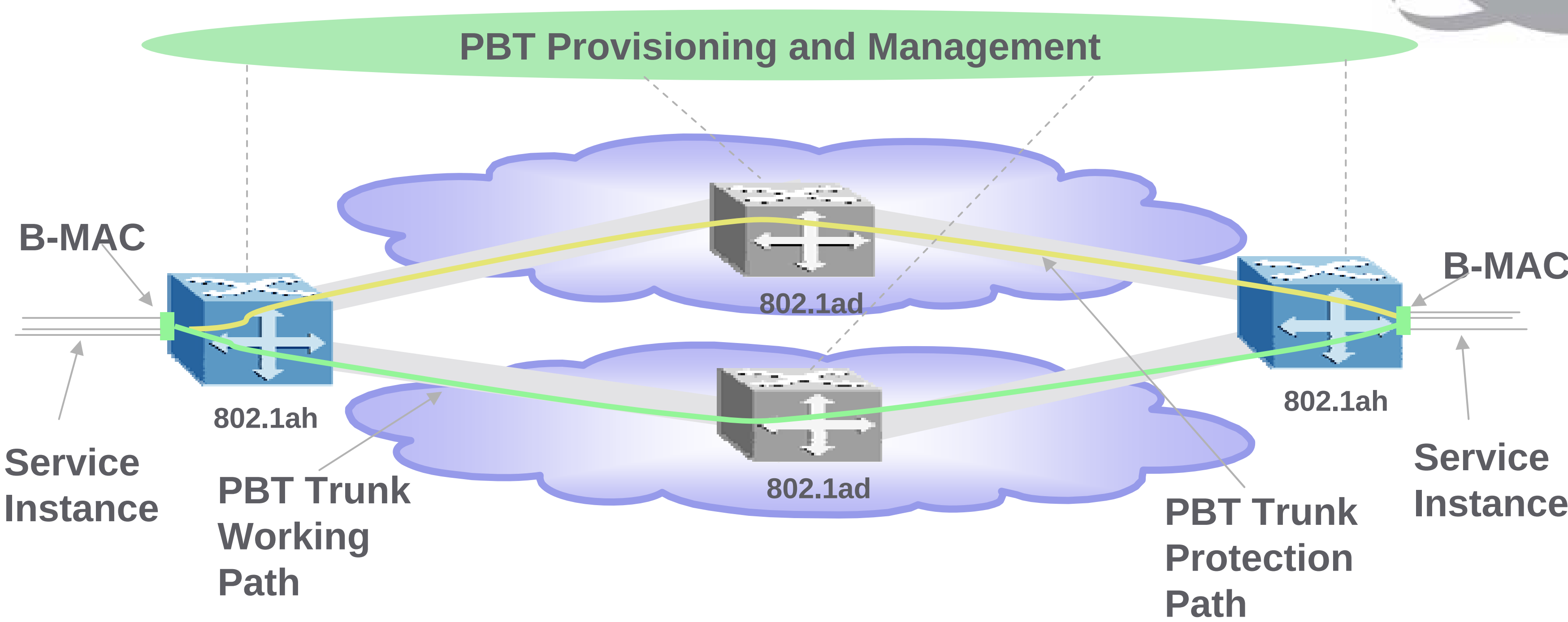
- Unicast CCM – Y.1731
- Unicast Loopback – 802.1ag and Y.1731
- Performance Monitoring (FLR, FD) – Y.1731
- Protection Switching – Y.1731 and G.8031

IEEE 802.1ag enhancements for PBT

> As part of the PBT PAR, enhancements to 802.1ag can be introduced in the following areas:

CFM Enhancements	Specified in	
	Y.1731 Clause	G.8031 Clause
Unicast Addressing for CCM	10.2	
PM – Frame Loss Ratio	8.1, 9, 10	
PM – Frame Delay	8.2, 9, 10	
Protection Switching	7.8, 10.10	11

Protection Switching for PBT Trunks



> Working and protected path use different B-VIDs

> Directed IEEE 802.1ag CC messages are sent over both the Working and Protection paths.

> CC messages are used to determine failure events and cause protection switching.

P802.1ag (CFM) Detects Faults

> CFM CC messages are run constantly on both the working and protection B-VLAN

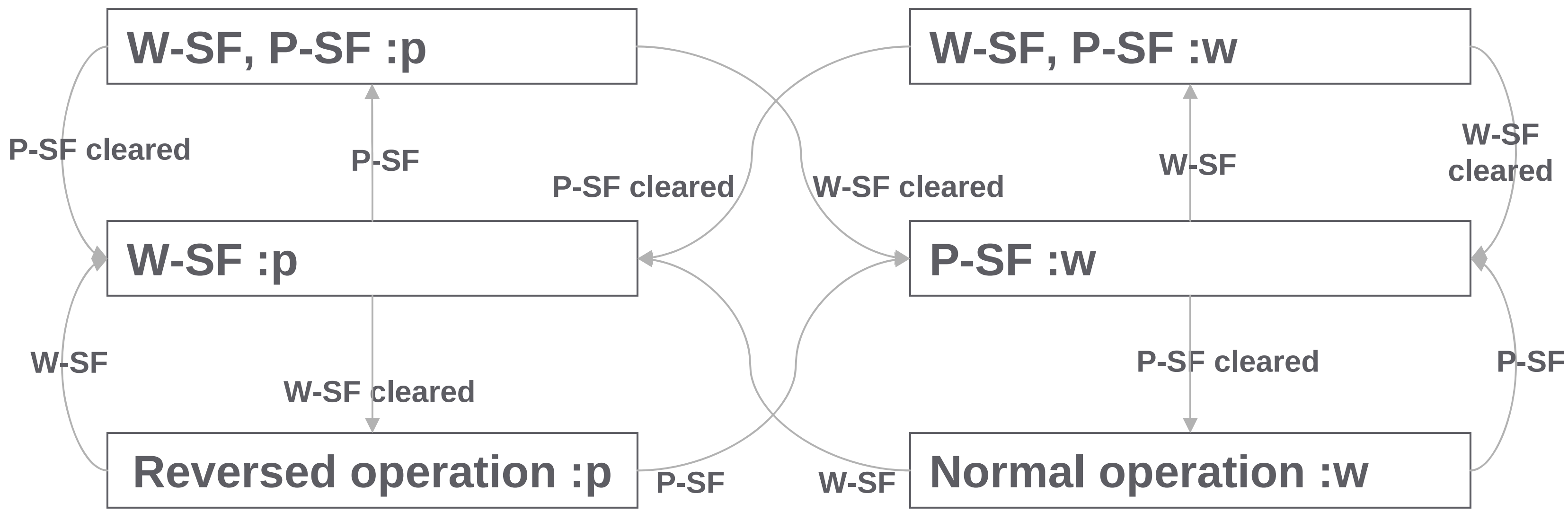
> Failures are detected through CC timeout, AIS detection, or manual intervention

> When a failure is detected the B-Shim moves the services to the protection B-VLAN
  - If the protection B-VLAN is failed the B-Shim protection switching is inhibited.

> Once receivers see traffic on the protection B-VLAN they must discard the old working path traffic to prevent mis-ordering
  - This may result in some loss during the protection switch

> Before the failed B-VLAN is reactivated the NMS must inform all the PBBs involved that the primary is available.
  - Prevents protection switching oscillations

Protection State Machine



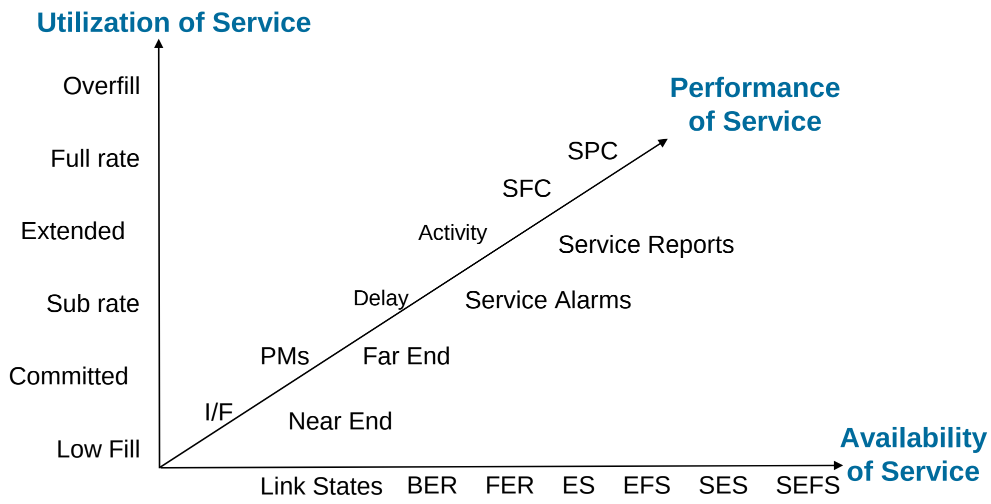
> W-SF state is primary PBT trunk signal fail

> P-SF state is secondary PBT trunk signal fail

> :p state is running on secondary :w state is running on primary

Ethernet SLA Management Features ITU/MEF

1. **Performance of Service**
   a) **Frame Loss Ratio (FLR)** parameter is the number of service frames marked green on a per {VID, Pbits, CoS} basis that are delivered by the Provider network versus the total sent.
   b) **Frame Delay (FD)** Measurement of round trip frame delay by ultiizing the OAM frames as defined in ITU-T Y.1731



   c) **Frame Delay Variation (FDV-Jitter)** Measurement of delay using time stamps of consecutive OAM frames.

2. **Availability of Service**
   a) AoS is currently defined in Y.1731 as the amount of time that the PoS (i.e., FLR, FD, FDV for a given service) is satisfied versus the overall period of time in service.
3. **Utilization of Service**
   a) UoS is a proposed parameter derived from the OUTOCTETS count on a per {VID, P bits, CoS} basis The counter is read periodically (e.g., every second) and binned to some intermediate value (e.g., 1 minute), when an average utilization metric can be calculated
   b) Usage: Tracks bandwidth usage over time, fault detection,

Items in ORANGE are not available in MPLS OAM

Scaling Properties

> The theoretical limit to the number of PBT trunks is 2**58
  - A PBT trunk is addressed by the combination of the B-MAC and B-VID
  - B-MAC provides 2**46 addresses and the B-VID provides 2**12 totaling 2**58
  - For large networks the B-VID address space will be removed over peer E-NNIs reducing the total available address space to the B-MAC address space of 2**46 PBT trunks (~70 trillion)

> Since PBT can distribute the PBT trunks over the network it is possible to scale well beyond the forwarding database limit by selecting routes through parallel switches.
  - Each switch must have a filtering database entry for each PBT trunk supported, therefore the number trunks on any individual switch is limited to the capacity of the filtering database

> Practically the number of PBT trunks is limit by the size of the database which the provisioning system can support and manage.
  - Today's SONET networks have provisioning systems which can support all the trunks currently used in the world, therefore this limitation must not be a barrier to building a world wide PBT network on the same scale as todays SONET network

> Another scaling consideration for PBT is the load generated by CCM messages used to manage PBT trunks
  - CFM at 10 msec. interval is ~ 6.4 Kbytes/sec (100 x 64 byte packets = 6.4 Kbytes/sec)
    - 3.3% of a T1
    - 0.5% of a 10BaseT
  - PBT is a P2P ***trunking*** technology
    - Each trunk would normally have assigned bandwidth greater then 10 Mbits and would carry many services

> The number of PBT trunks supported by a BEB will be limited by the CCM packet processing capability of the BEB.
  - For 1000 trunks with 50 msec protection the total load is about 200,000 f/s.

> The I-SID address space limits any PBBN to handling 2**24 service instances.
  - This can be extended by dividing the network into separate PBBNs connected by E-NNIs as specifed in 802.1ah
  - For peer E-NNI the I-SID is translated over the E-NNI allowing the service address space to indefinitely extended.

PBT Data Relay Modifications

> Define a special MSTID called the PBTID (use 0xFFE) which identifies PBT rather than a MSTI.
  - An MSTIDs not in the MSTI list indicates some protocol other than MSTP that may run in parallel to MSTP (802.1Q-2005 12.12.1 and 8.6.2)

> Allow the FID to MSTID Allocation Table (12.12.2) to allocate a FID to the PBTID. (i.e. FID=0xFFE to PBTID=0xFFE)

> Allow the MST Configuration Table (12.12.3) to allocate VIDs to the PBTID.
  - The PBTID code of 0xFFE in the MST Configuration Table means “this B-VID is not allocated to an MSTI and is available to PBT for use as a route selector“.

> All VIDs allocated to PBT have a port state at each bridge port who’s state is forced to forwarding=on and learning=off (change 8.4)

> Add a static filtering entry type for all unicast addresses, for which no more specific static filtering entry exists (change 8.8.1 bullet a))

> Informative Annex explaining the use of PBT

What Would IEEE Specify?

> Add to or amend 802.1ah

> Things which are required for PBT relay:

- 1)Must provide a method for splitting B-VID address space between different topology protocols
- 2)Must provide PBT states which force the port PBT port state to learning off and forwarding on
- 3)Must provide a feature to disable broadcast and unknown forwarding

> Not required for PBT, but would be nice:

- Provide 802.1ag features for CC, LB, and LT (derived from Y.1731)
  - All CFM frames delivered over PBT must be unicast B-DA
  - For responses frames must a PBT TLV with the reverse B-VID
  - Definitions for PBT trunk group management
- Provide IEEE protection switching features (derived from G.8031)

Title

- **Title:** IEEE Standard for Local and Metropolitan Area Networks---Virtual Bridged Local Area Networks – Amendment ??: - Provider Backbone Transport

Scope Part 1: What the project is

> **Scope:** The scope of this standard is to define feature extensions to 802.1Q-2005 supporting provisioning of scalable, traffic engineered, point-to-point trunks within Provided Backbone Bridged (802.1ah) networks which may be used in addition to B-VLANs for carrying services, while not over taxing the 12 bit B-VID address space.

Scope Part 2: What the project is not

> An 802.1Q bridge with PBT extensions will provide the features necessary to allow an external provisioning or control system to program any desired path through the network. The selection of paths and the management of the bandwidth allocated to the paths is part of the external provisioning or control system and therefore out of the project scope. This project uses MIB management as the method by which external provisioning or control systems program the 802.1Q bridges to create PBT trunks.

Scope Part 3: Required features

> The features specified by this project include support for data forwarding for PBT trunks, CFM CCM messages (see ITU-T Y.1731) on PBT trunks and 1:1 protection switching (see ITU-T G.8031) for PBT trunks. The modifications to 802.1Q needed for PBT data forwarding are a method to support discarding rather than forwarding for broadcast and unknown frames, a method for splitting and configuring the VID space between spanning tree (or shortest path bridging) and an external provisioning system and a method for setting the port state for the VID space allocated to PBT to forwarding and not learning in a way which is compatible and interoperable with Provider Backbone Bridged network (P802.1ah) protocols and equipment. The provisioned, traffic engineered, point-to-point paths will operate on the allocated VIDs simultaneously allowing multiple spanning tree or the new shortest path bridging to manage the topologies of the other VIDs. SNMP management will be supported.

Purpose

> **Purpose:** An essential requirement of provider networks is supporting traffic engineered paths. These paths must not be limited to following a small number of spanning trees or shortest path routes. Instead complete route selection freedom must be allowed. This amendment enables a Service Provider to traffic engineer provisioned point-to-point trunks in a Provider Backbone Bridge network while scaling the number of point-to-point trunks to the limits of the 802 MAC address space.

Need for the Project

> **Need for the Project:** Provider networks rely on the ability to directly control the routing of point-to-point trunks used to transport services. The control of path routing in turn supports traffic engineering for the allocation of bandwidth, assurance of diverse backup path routing, and selection of path performance as required by service level agreements. This project provides essential features used to support direct control of route selection of point-to-point trunks within Provider Backbone Bridged networks, while allowing scaling to the limit of the 802 MAC address space. Despite the demand and initial deployments of point-to-point traffic engineered backbone trunks for carrying customer VLAN traffic, there is currently no interoperability between different vendors, nor a coherent management framework for different techniques. Most major carriers, who will be the users of this standard, are currently deploying point-to-point service networks which need traffic engineering of provisioned point-to-point trunks to meet the demands of transition from existing leased line service.

Stakeholders for the Standard

> **Stakeholders for the Standard:** Developers and users of networking for Provider network environments including networking IC developers, switch and NIC vendors, and users.

>BUSINESS MADE **SIMPLE**

Questions?

NORTEL