802.1aq Shortest Path Bridging Design implications

Don Fedyk
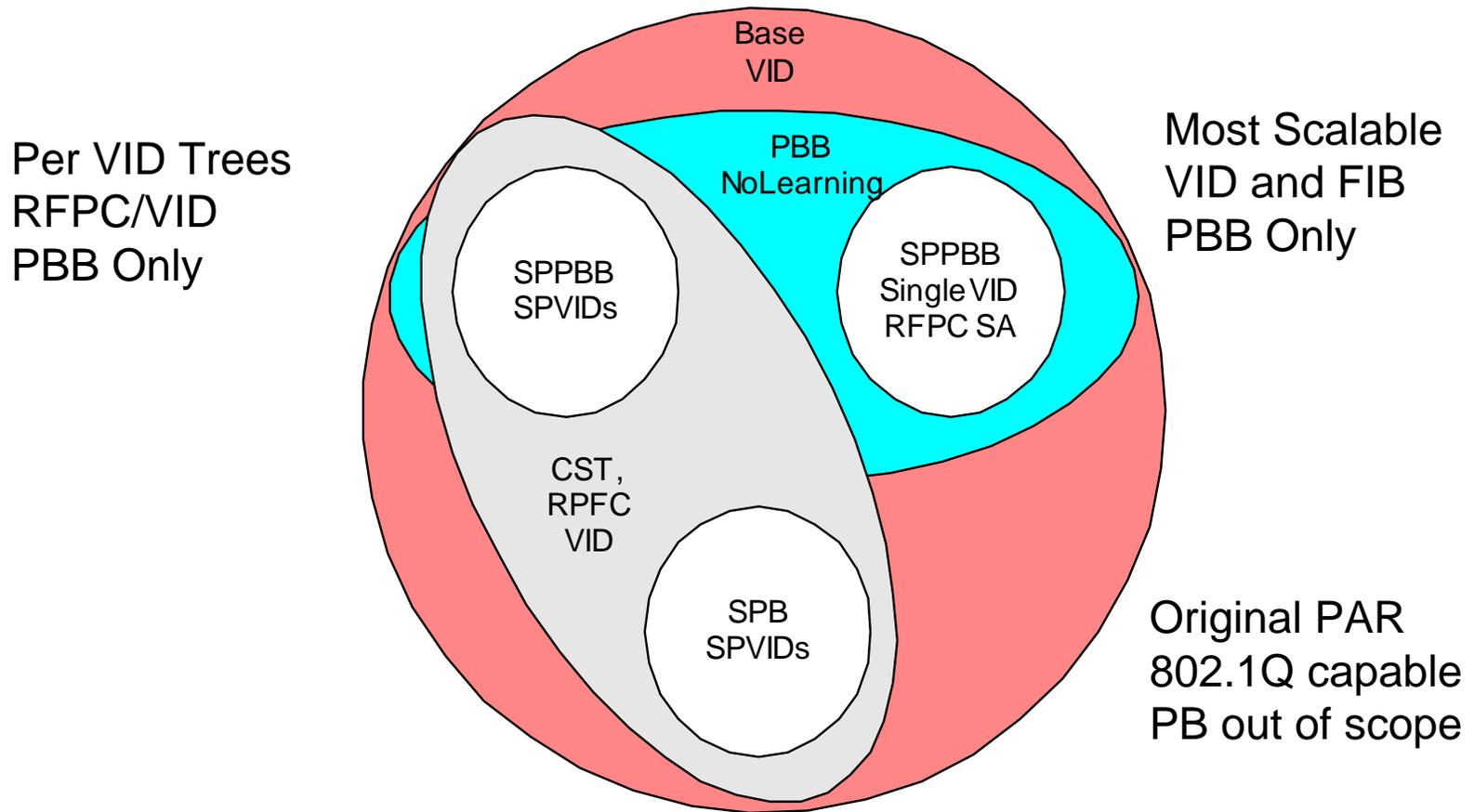Senior Advisor
Nortel
dwfedyk@nortel.com

Design Decisions for 802.1aq

- Best Choice IS-IS as the Link state protocol
- SPB (Shortest Path Bridging) (802.1Q compliant)
  - must use unidirectional VID as SA Proxy, don't own the C-SMAC
  - Solution Attributes
    - Uses VID Trees, one per (edge) bridge, distributed in IS-IS
    - Defines a SPT (Shortest Path Tree) Region, def by "Base VID"
    - SVL learning of unicast forwarding required?
    - Uses VID RPFC – requires VID semantics to be modified to be unidirectional
  - Solution Requirements
    - May Interwork at edges with RSTP, MSTP or Ships in the night with RSTP, MSTP
    - The region may default to a single instance MSTP (associated with the "Base VID") if the VID allocation fails or detects errors

Design Decisions for 802.1aq cont

- SPPBB (Shortest Path Provider Backbone Bridging)
  - May use VID Trees or a Single VID for an SPT Region
    - The region may default to a single instance MSTP if the VID allocation fails or detects errors.
  - Does not allow learning of B-MACs
    - Provider addresses will all be known allows for more efficient flooding (no B-MAC broadcast storms), RPFC, Reduction in forwarding space Shared Forwarding, Efficient Multicast and faster convergence Link State.
  - Uses VID RPFC or SA based RPFC (Single VID) :
    - VID based imposes scaling limits on B-MACs and ECMT
  - Works Ships in the Night with RSTP, MSTP in the B-MAC space.
    - Only Translation is supported at the edges
    - No need to interwork with RSTP or MSTP (learning constraint)

Shortest Path Bridging

Per VID Trees
RFPC/VID
PBB Only

Most Scalable
VID and FIB
PBB Only

Base
VID

PBB
NoLearning

SPPBB
SPVIDs

SPPBB
Single VID
RFPC SA

CST,
RPFC
VID

SPB
SPVIDs

Original PAR
802.1Q capable
PB out of scope

Currently Three Variants Why?

Design Issues for Per VID Trees SPVID

- Uses VID space at a rate of 1 per Node
  - Equal cost paths would cost more VIDs per node per ECMT
- Implications for MMRP
  - MMRP state machine/tree at every node?
  - Multicast interest can be delivered and populated by IS-IS
- Unidirectional VIDs
  - need modifications of other protocols
- Implications on Unicast address forwarding
  - Need VID/Unicast destination B-MAC
- Require VID RPFC
  - Can the VID be ignored for forwarding
   - single Base VID yes, Multiple?

Design Issues for DA Based Trees (Single VID)

- Uses a single VID per instance per domain
- Implications for MMRP
  - MMRP state machine/tree?
  - Multicast interest can be delivered and populated by IS-IS
- Unicast address forwarding
  - Single VID destination B-MAC
- Require SA RPFC

Switching Context of Source Tree Always a VID & DA Context

SPVID

Payload
SPVID
SA
DA

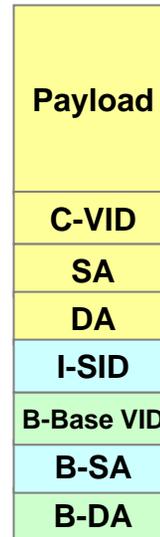| Payload |
|---------|
| C-VID |
| SA |
| DA |
| I-SID |
| B-SPVID |
| B-SA |
| B-DA |
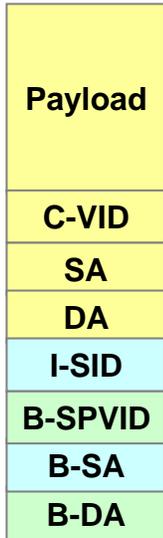
VID Trees

SPVID = Source Tree and VLAN In VID Tree Context

Single VID

Resolve DA to unique destination

Optimal Multicast B-DA based on Multicast Interest (I-SID)

Payload
VID
SA
DA

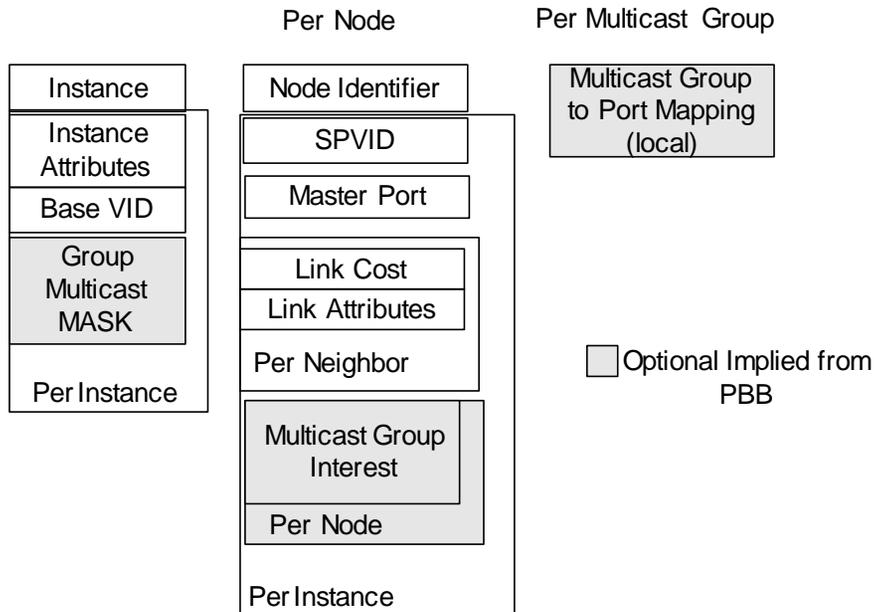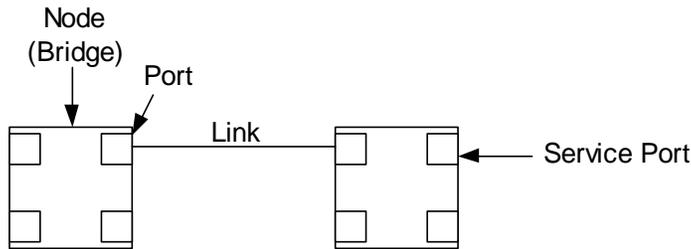| Payload |
|---------|
| C-VID |
| SA |
| DA |
| I-SID |
| B-Base VID |
| B-SA |
| B-DA |

(B-VID + B-DA) = Source Tree and Destination in Source VID Tree Context

VID + DA = Topology and Destination in VID Context

IS-IS Functional Elements
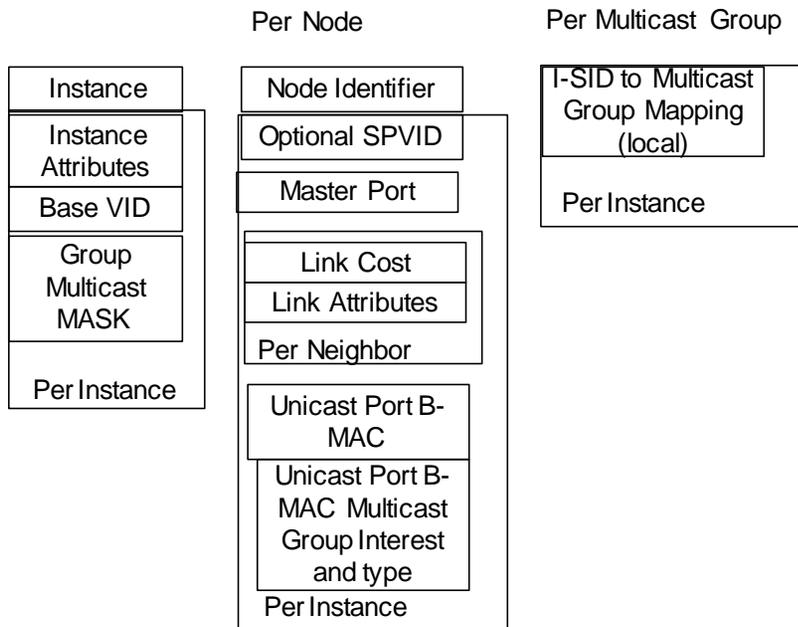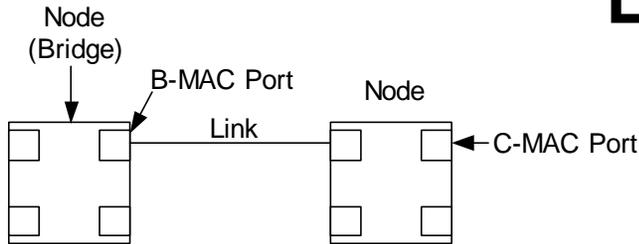# Shortest Path Bridging

Node
(Bridge)

Port

Link

Service Port

Per Node

Per Multicast Group

| Instance |
| Instance Attributes |
| Base VID |
| Group Multicast MASK |
| Per Instance |

| Node Identifier |
| SPVID |
| Master Port |
| Link Cost |
| Link Attributes |
| Per Neighbor |
| Multicast Group Interest |
| Per Node |
| Per Instance |

| Multicast Group to Port Mapping (local) |

Optional Implied from PBB

IS-IS Functional Elements Shortest Path Provider Backbone Bridging

Node
(Bridge)

B-MAC Port

Node

Link

C-MAC Port

Per Node

Per Multicast Group

Instance

Instance Attributes
Base VID

Group Multicast MASK

Per Instance

Node Identifier
Optional SPVID
Master Port

Link Cost
Link Attributes

Per Neighbor

Unicast Port B-MAC
Unicast Port B-MAC Multicast Group Interest and type

Per Instance

I-SID to Multicast Group Mapping (local)

Per Instance

Note When looking at the IS-IS information, the similarities for all three options are striking. There is hope they are not all that different.
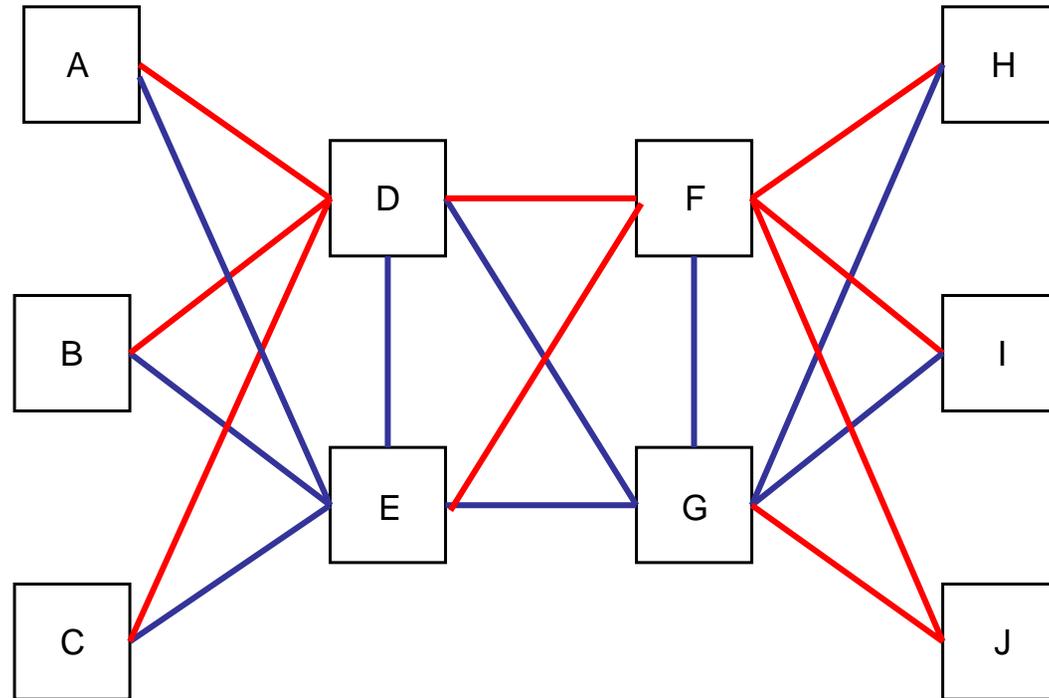
SPPBB Congruency
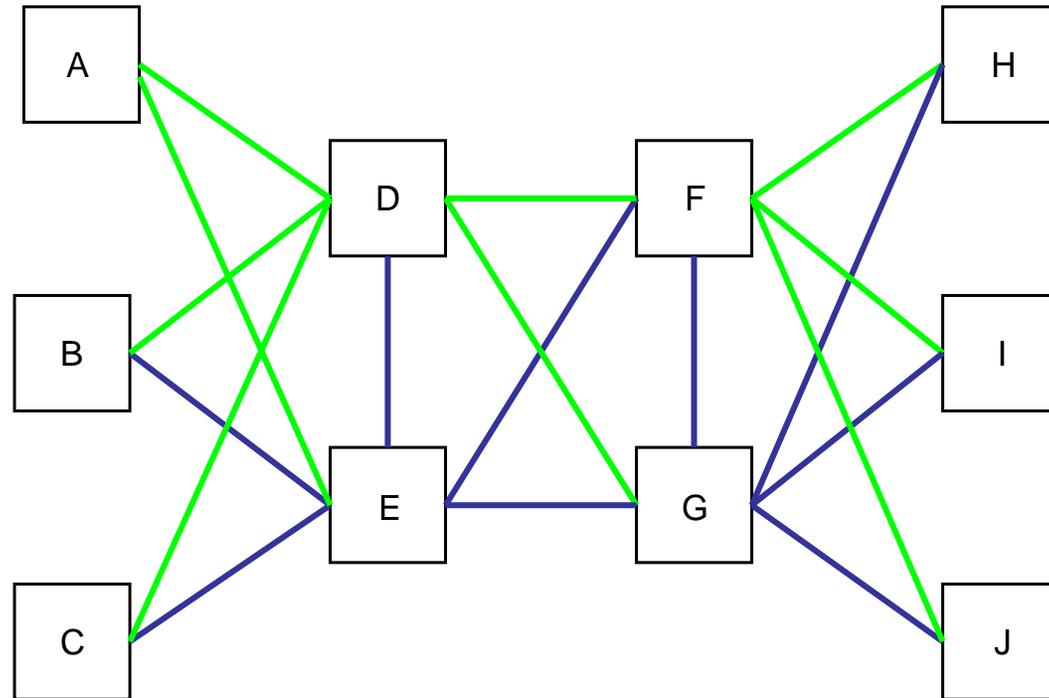
1 MMAC/Multicast Group
1 UMAC/Switch

Shortest Path Tree from A

Create Shortest Path tree from every switch. Also simultaneously
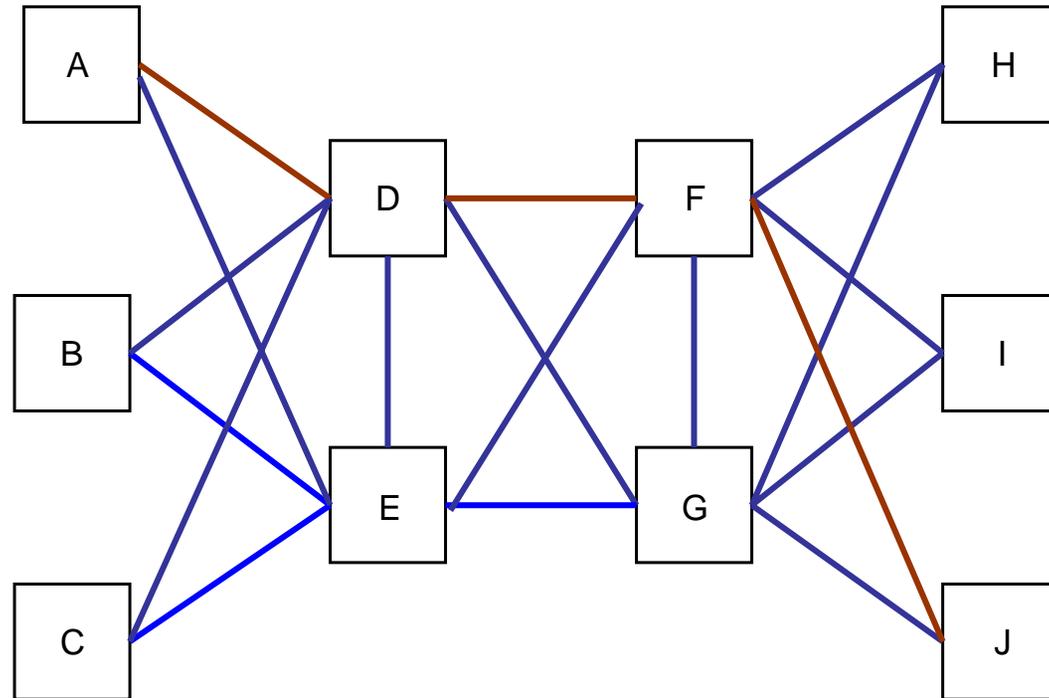**Create a congruent PBB_TE point to point path**

SPPBB Congruency

1 MMAC/Multicast Group
1 UMAC/Switch

Shortest Path Tree from J

Create Shortest Path tree from every switch. Also simultaneously
Create a congruent PBB_TE point to point path
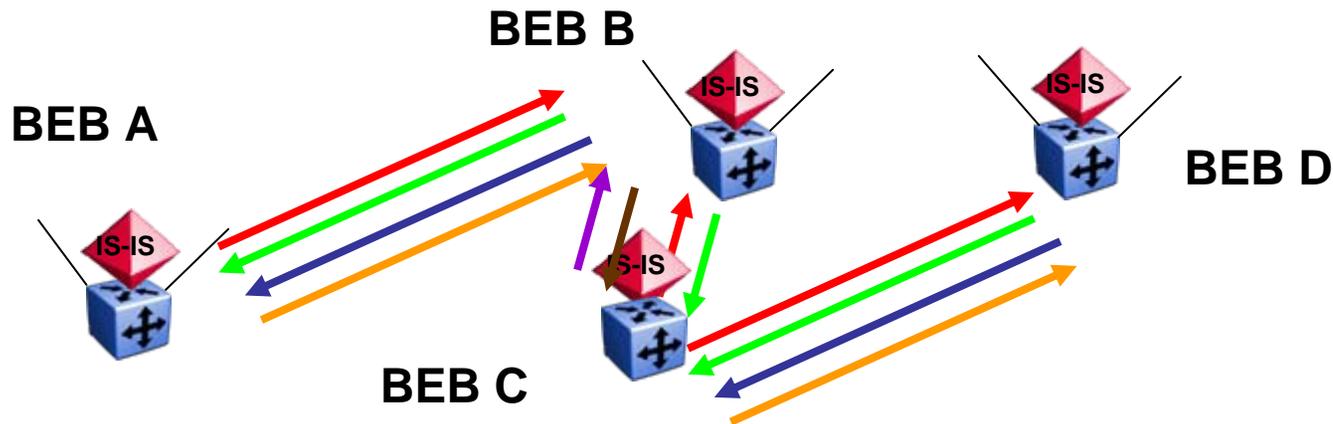
SPPBB Congruency

Congruency
A-J, J-A

1 MMAC/Multicast Group
1 UMAC/Switch

Shortest Path Tree from J

Create Shortest Path tree from every switch. Also simultaneously
**Create a congruent PBB_TE point to point path**

Single VID FIB Zoom In

BEB B

BEB A

BEB D

BEB C

Type		DA	RPFC@B	Multicast Group ID
Unicast A to D	10	UDA-D	USA-A	
Unicast D to A, C to A	10	UDA-A	USA-D/C	
Multicast Shortest path Rooted at A	10	MDA-A	USA-A	Per MGI
Multicast Short Path Tree Rooted at B	10	MDA-D	USA-B	Per MGI

There are 4 paths between nodes (1 bidirectional Unicast) (2 directional multicast/I-SID) 3 forwarding entries relevant to any MGI destination
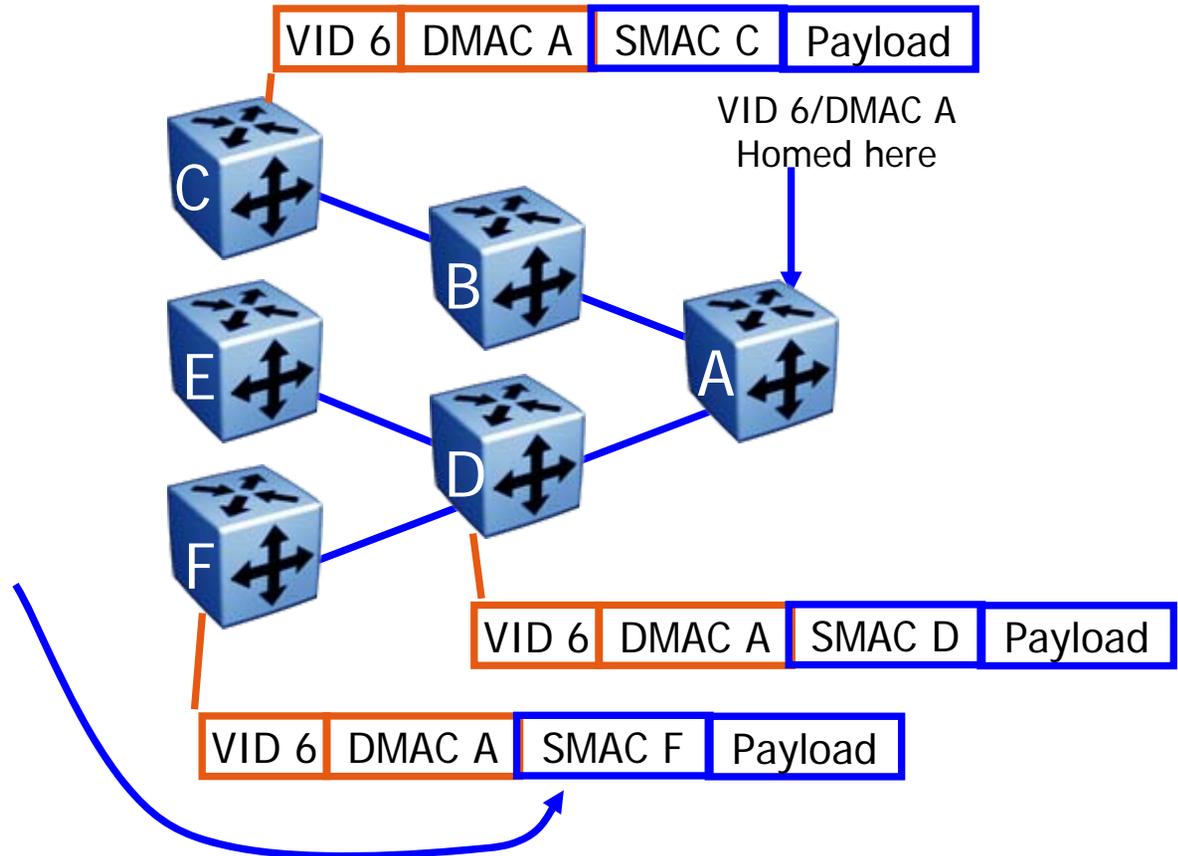
SPVID FIB Zoom In
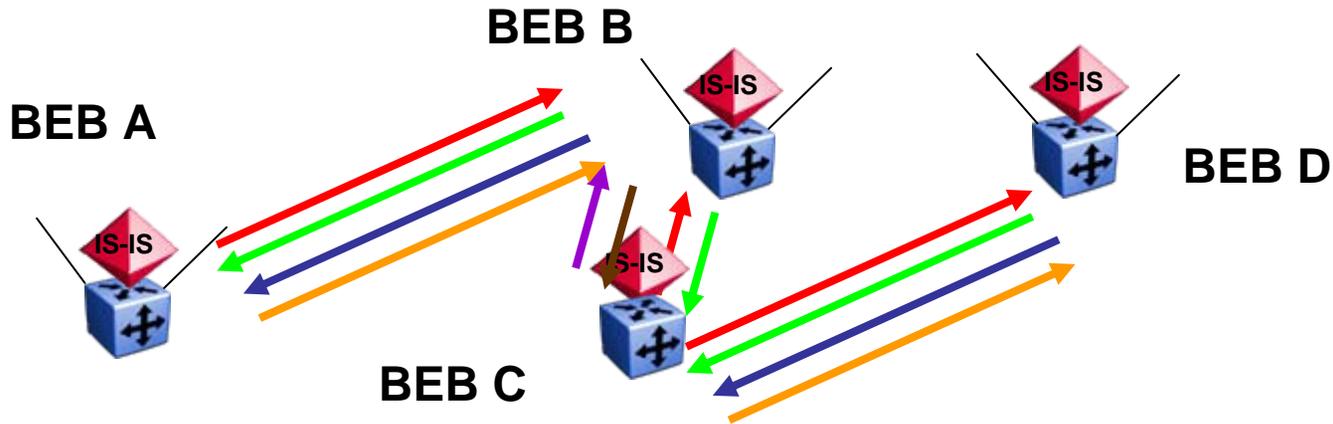
Type	VID	DA	RPFC	Multicast Group ID
Unicast A to D	10	UDA-D	VID 10	
Unicast D to A	20	UDA-A	VID 20	
Multicast Shortest path Rooted at A	10	MDA-A	VID 10	Per VID / MGI
Multicast Short Path Tree Rooted at B	20	MDA-D	VID 20	Per VID / MGI

In this case VID may be used to Multicast or specific Multicast DAs may be installed.

Unicast Shared Forwarding and the PBB-TE

One PBT Label, a single VID+DMAC can be used by multiple sources, providing Order (n) Labels per network.

Additionally, packets can be examined at destination to determine the source.

| VID 6 | DMAC A | SMAC C | Payload |

VID 6/DMAC A
Homed here

| VID 6 | DMAC A | SMAC D | Payload |

| VID 6 | DMAC A | SMAC F | Payload |

Single VID shared forwarding = one entry / (B-VID + DMAC) + any I-SID

SPVID = one entry / (B-VID tree + DMAC) + any I-SID

Control Plane Scope
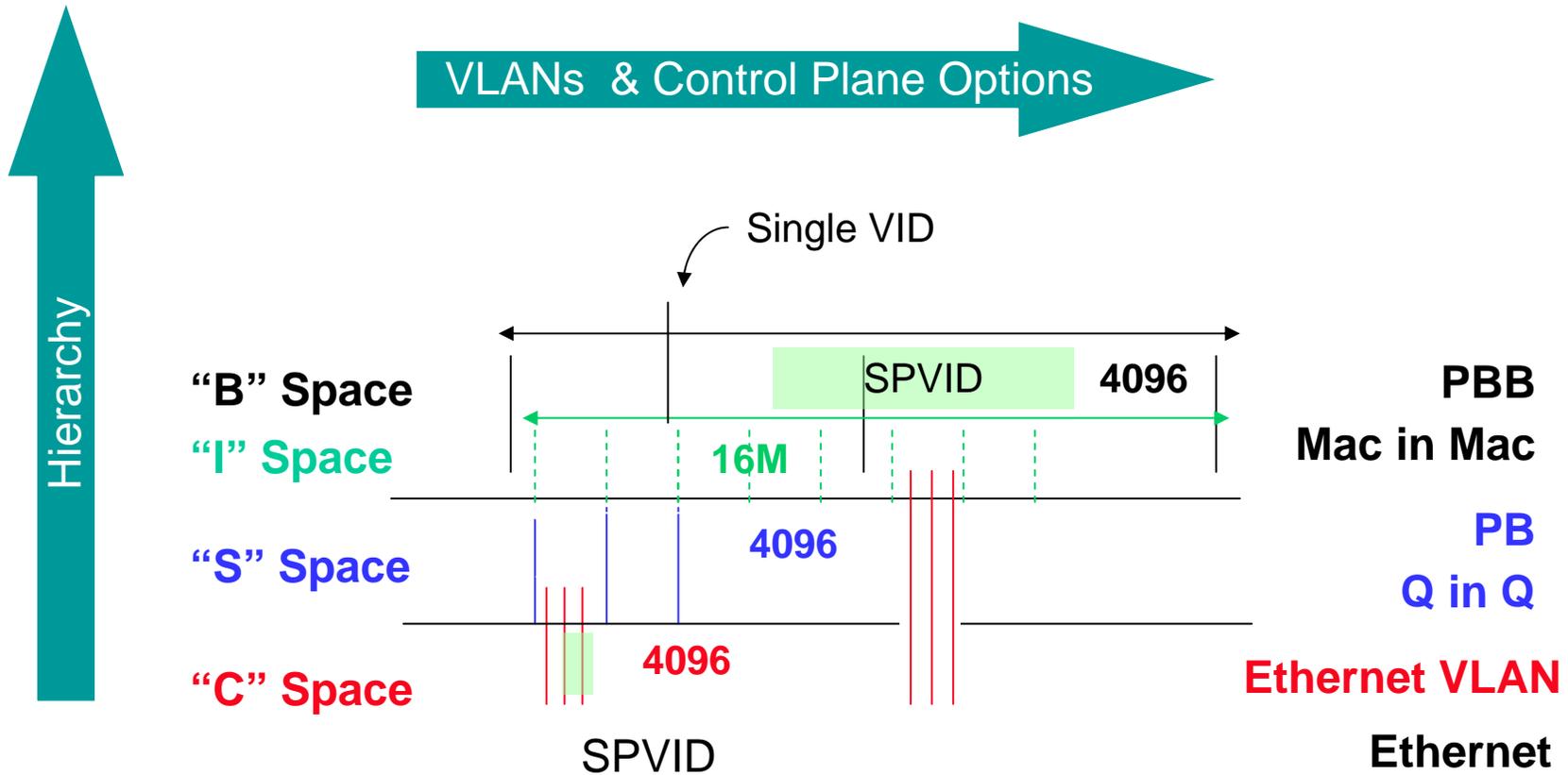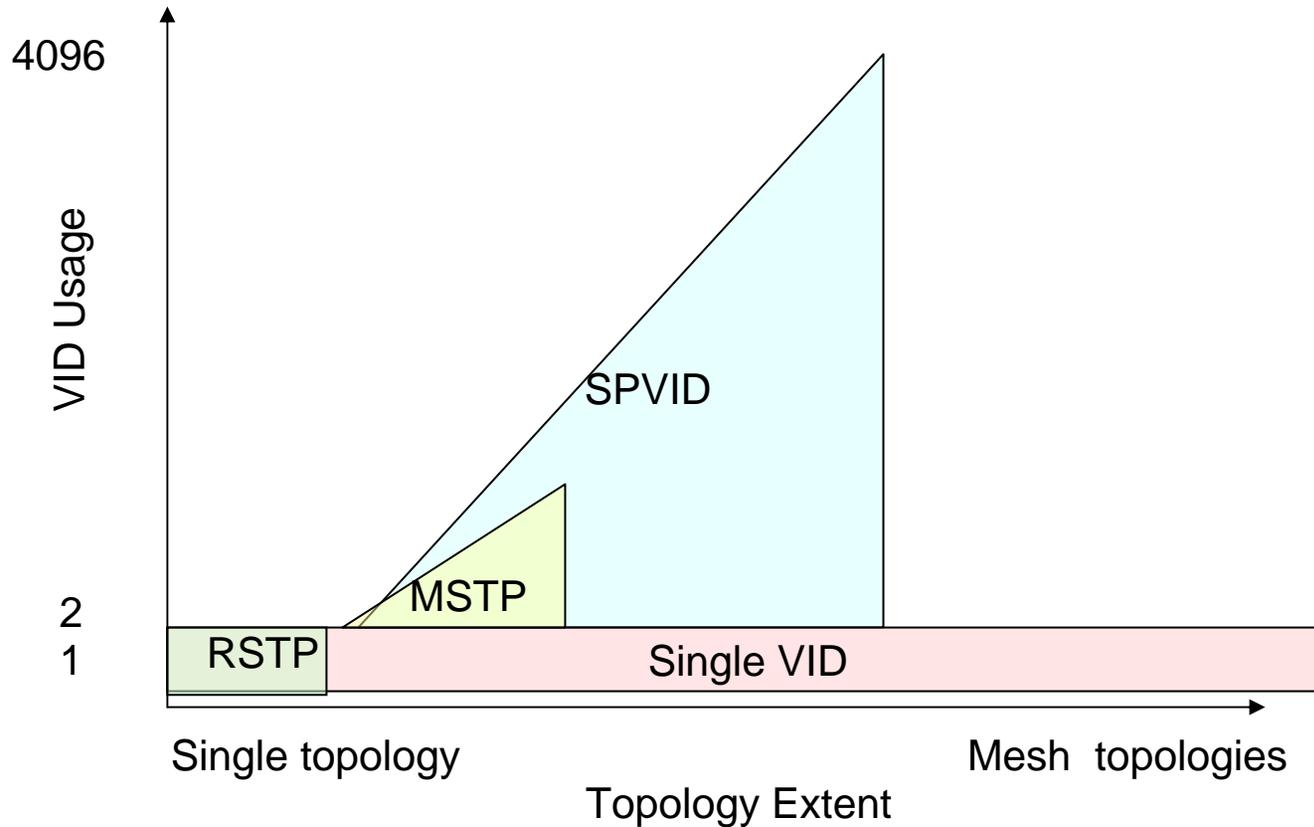

Different Operating Spaces

VLAN Usage and Topology

Topology Extent

(ability to utilize multiple paths)

SPPBB SPVID and Single VID Similarities

- VLAN Topology   Both sport shortest path trees
- VLAN Partitioning   Both use a logical B-VLAN
- Link state topology   Both use Link State
- Mesh Networking   Both support mesh
- Forwarding backwards   Both operate Ships in the night
  compatibility   with other VLANs
- Control plane objects   Similar
- SPT computation   Similar
- Multicast Groups   Both support via IS-IS

SPPBB Differences SPVID vs Single VID

- VLAN Usage   Per VID Tree Vs Single VID / Topology
- Shared Forwarding   Multiple entries per B-DA Vs Single VID one
- RPFC   Per VID Vs Single VID per B-SMAC
- Scalability VIDs   4000 VID trees VS Single VID # of MMACs
- Auto Config   Per Node VID Vs Single Base VID

Next Steps

- Continue Refining Draft
  - D0.4 Available but Rough
- All Spanning Tree - Shortest path bridging implications removed only IS-IS
- PAR issues ?
- Comments