>THIS IS **THE WAY**

Scaling 802.1ah Networks

Paul Bottorff, Michael Chen, Marc Holness, Dinesh Mohan

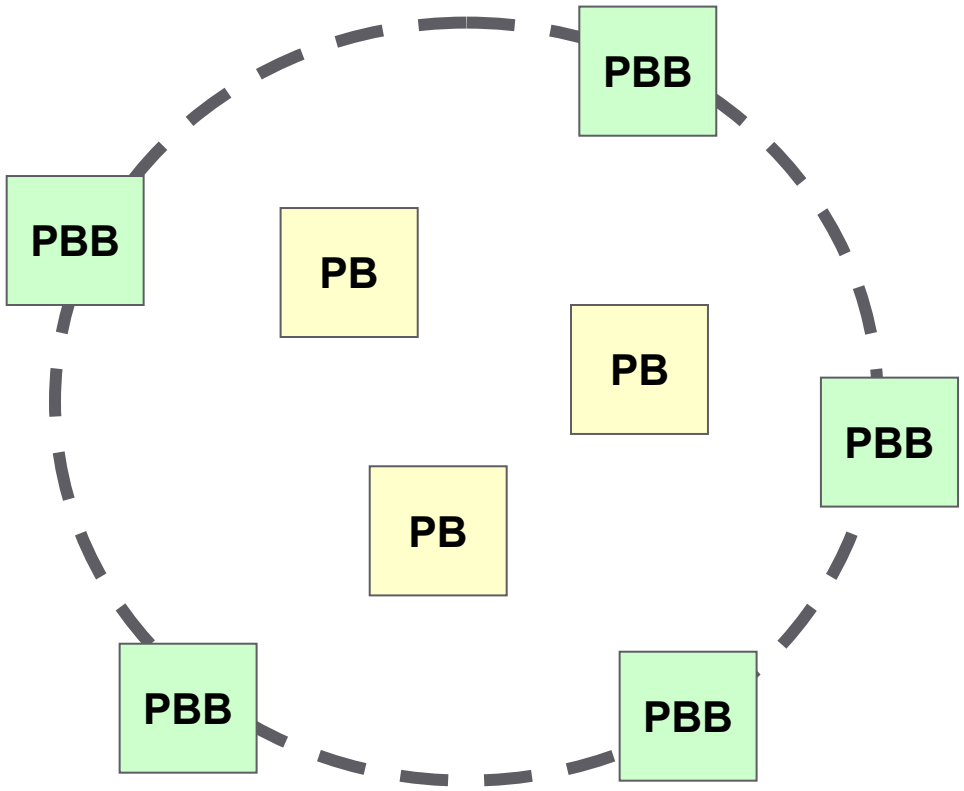
April 2005

>THIS IS NORTEL

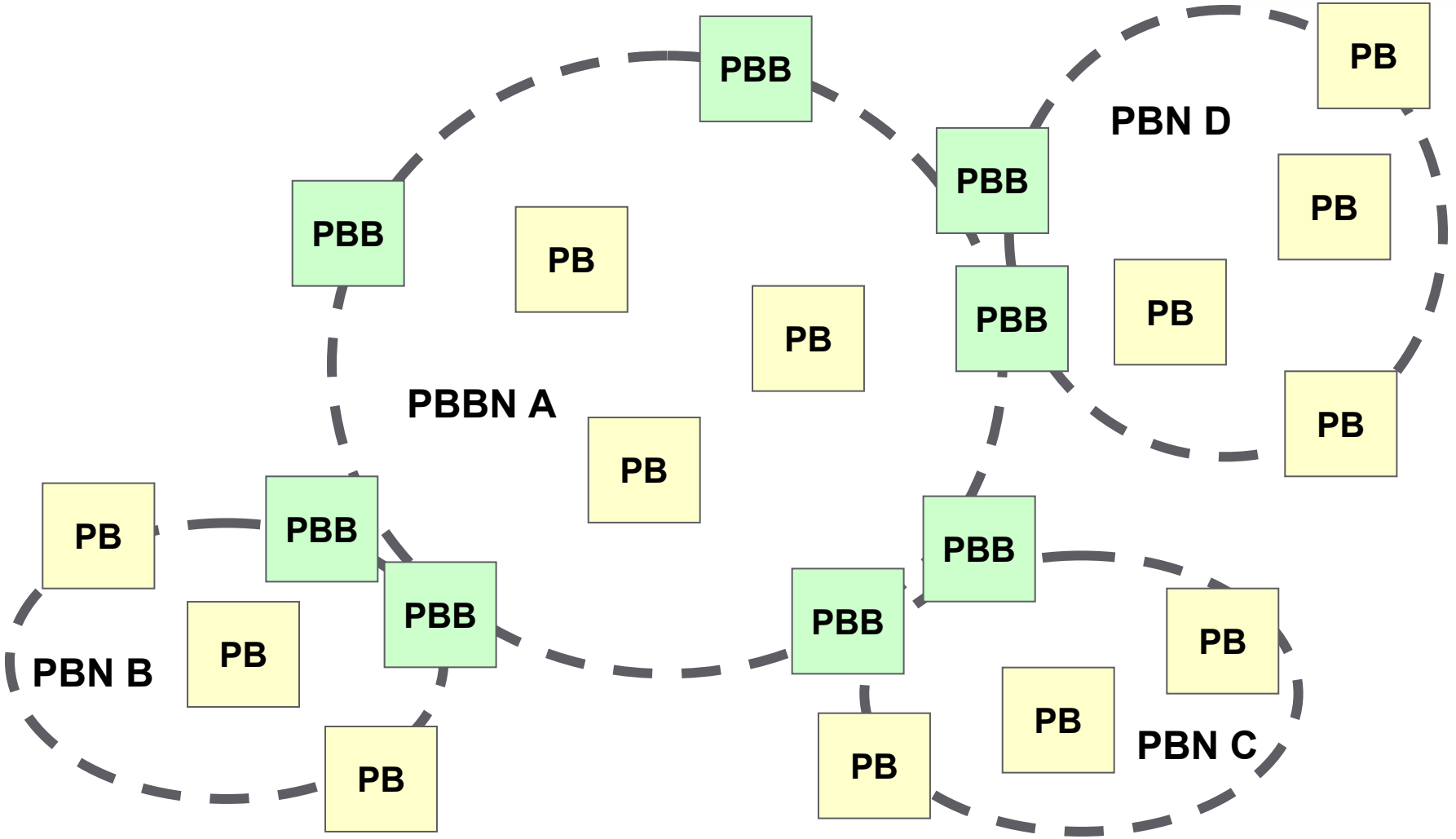
Problem Statement

- Spanning tree scaling is finite
- To make 802.1ah networks scale the 802.1ad and 802.1ah spanning trees must be decoupled, allowing each to operate in independent sub-regions
- In the 802.1ah model spanning tree packets are filtered at the I-Comp to B-Comp boundary
- The spanning trees for each Provider Bridge Network(PBN) are independent from each other and from the Provider Backbone Bridge Network(PBBN)
- The Provider Backbone Bridge must provide a redundant interconnect between the PBN and PBBN spanning tree regions

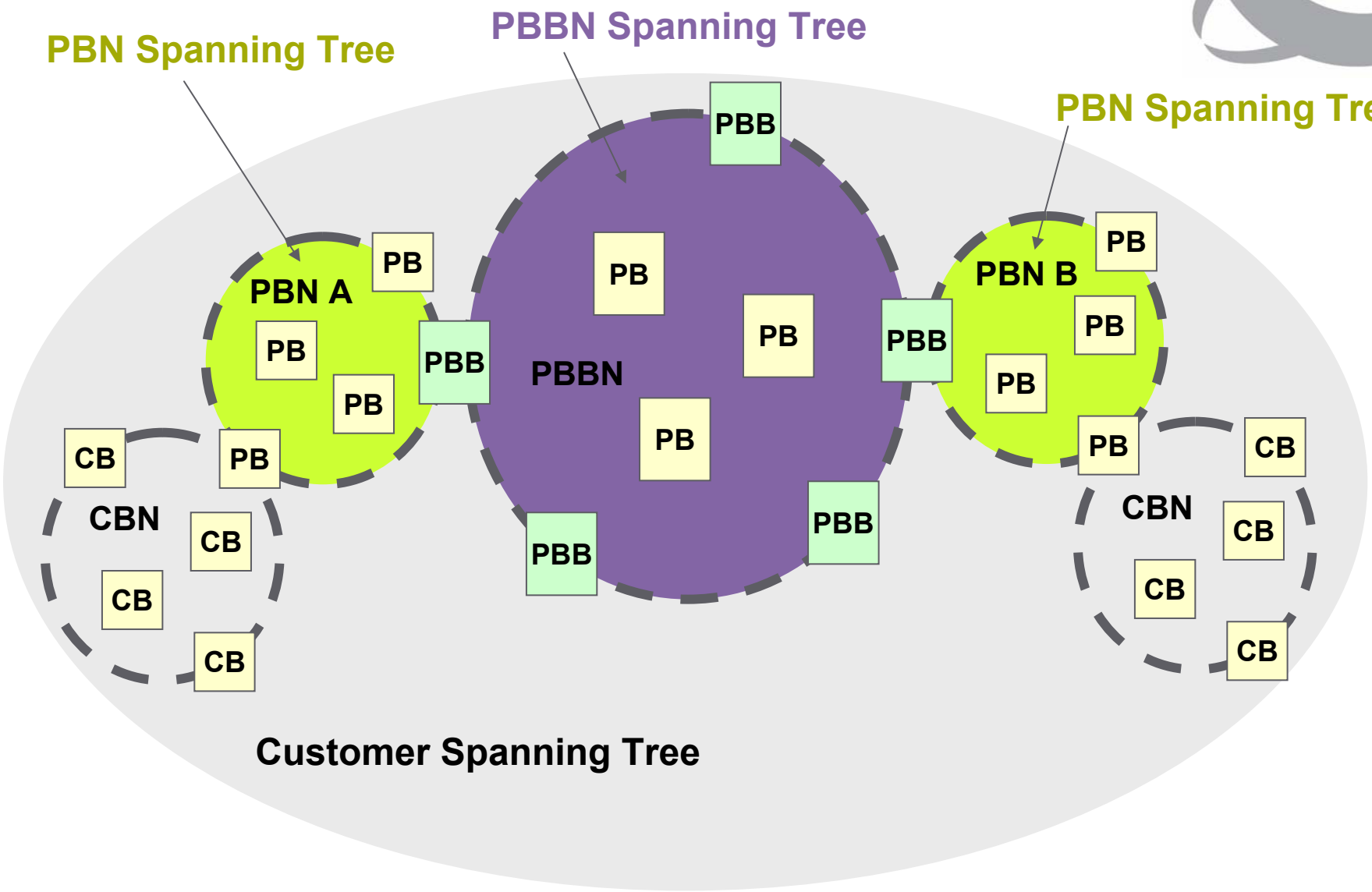
Single 802.1ah Network



Combined 802.1ad and 802.1ah Network



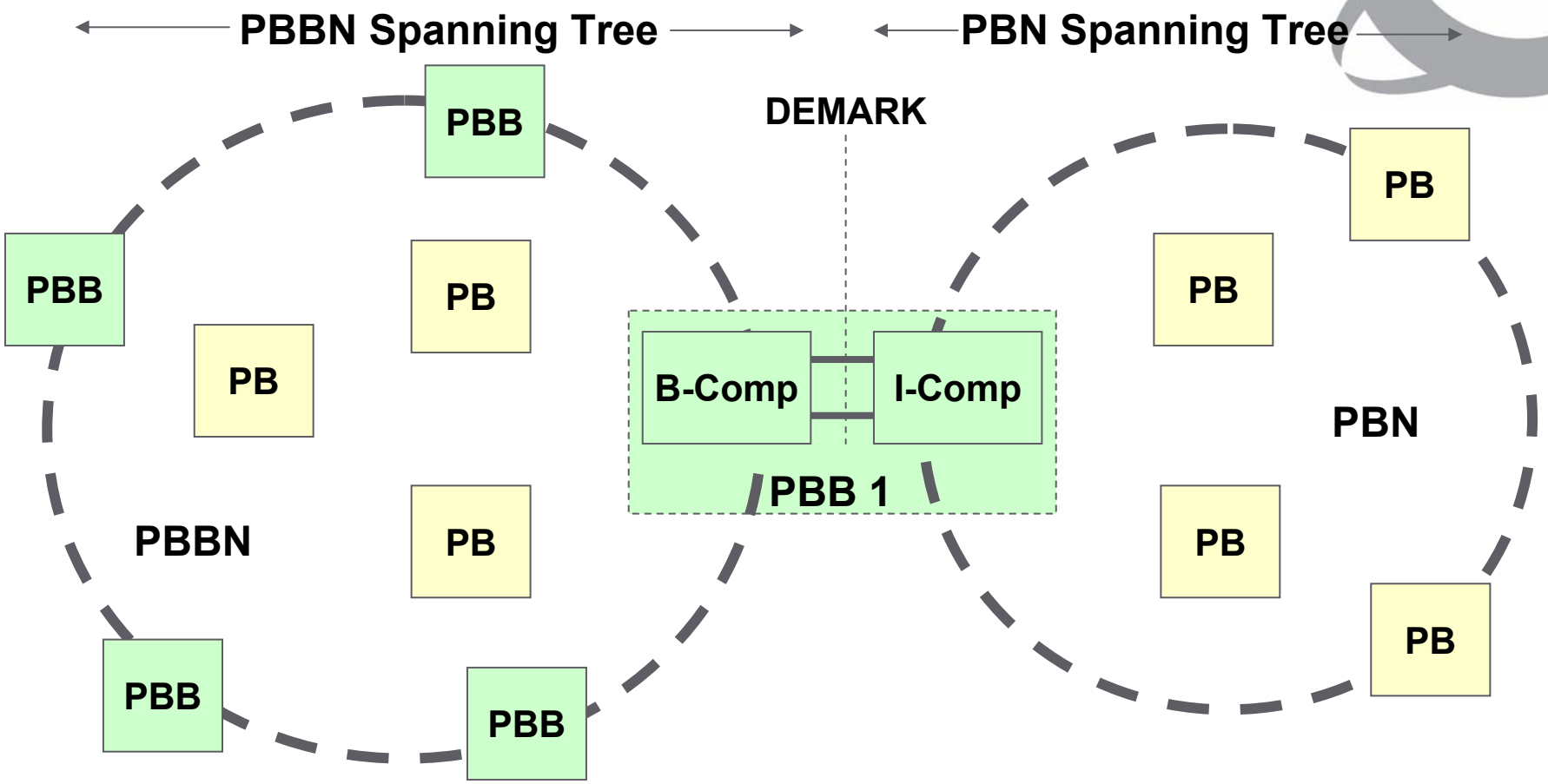
PBN and PBBN Spanning Trees



Topology Assumptions

1. Each PBN and PBBN prevents forwarding loops by running an independent spanning tree
2. Each PBNs connects to other PBNs only through a PBBN
3. Each PBN ensures that no data frames pass through more than one PBB attachment into or out of the PBN.
4. Each PBN ensures that it attaches any given S-VLAN to no more than one PBBN.

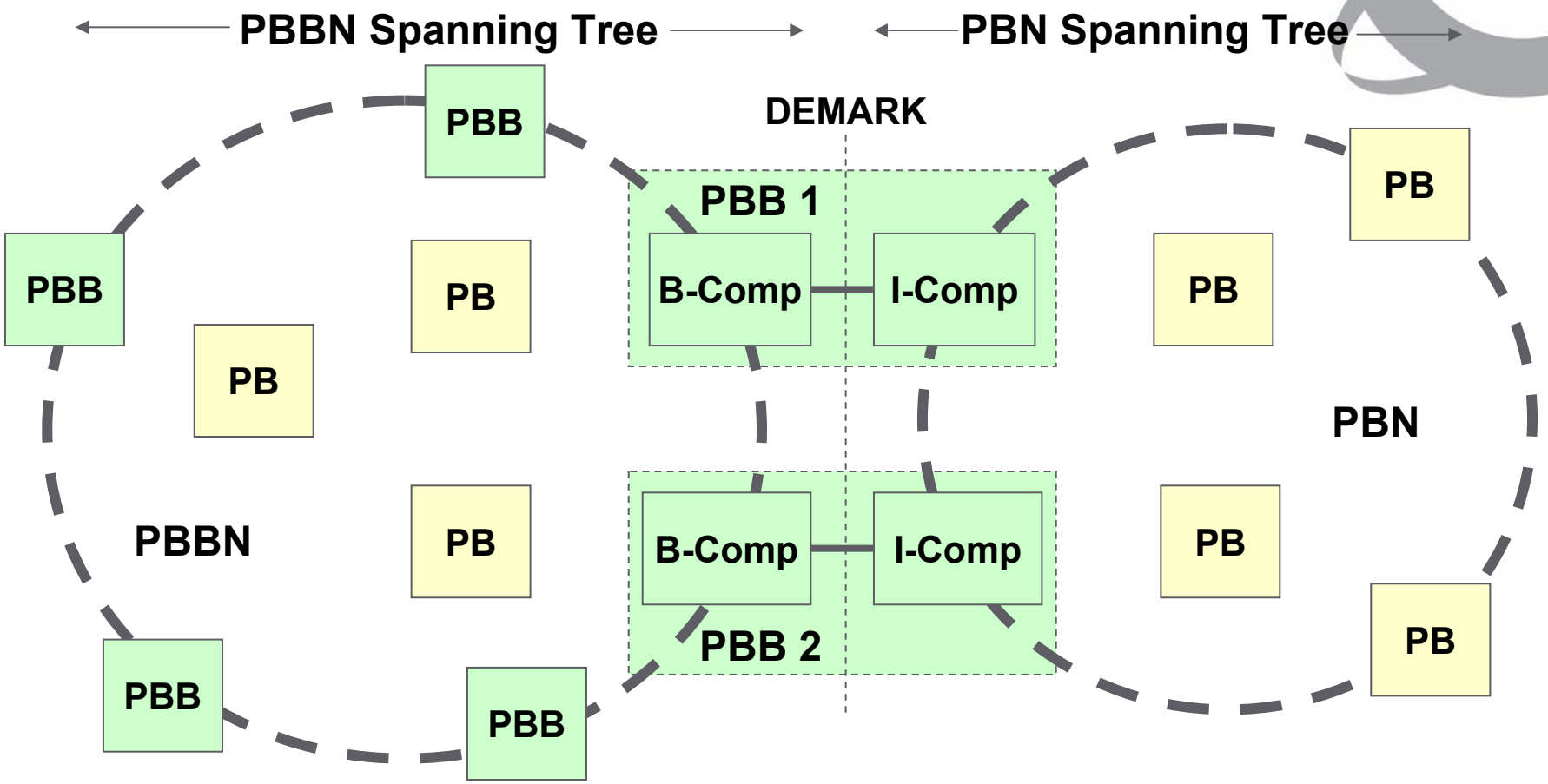
PBB Redundant Interconnect: Class 1



> PBB B-Components participate in the PBBN topology

> PBB I-Components participate in the PBN topology

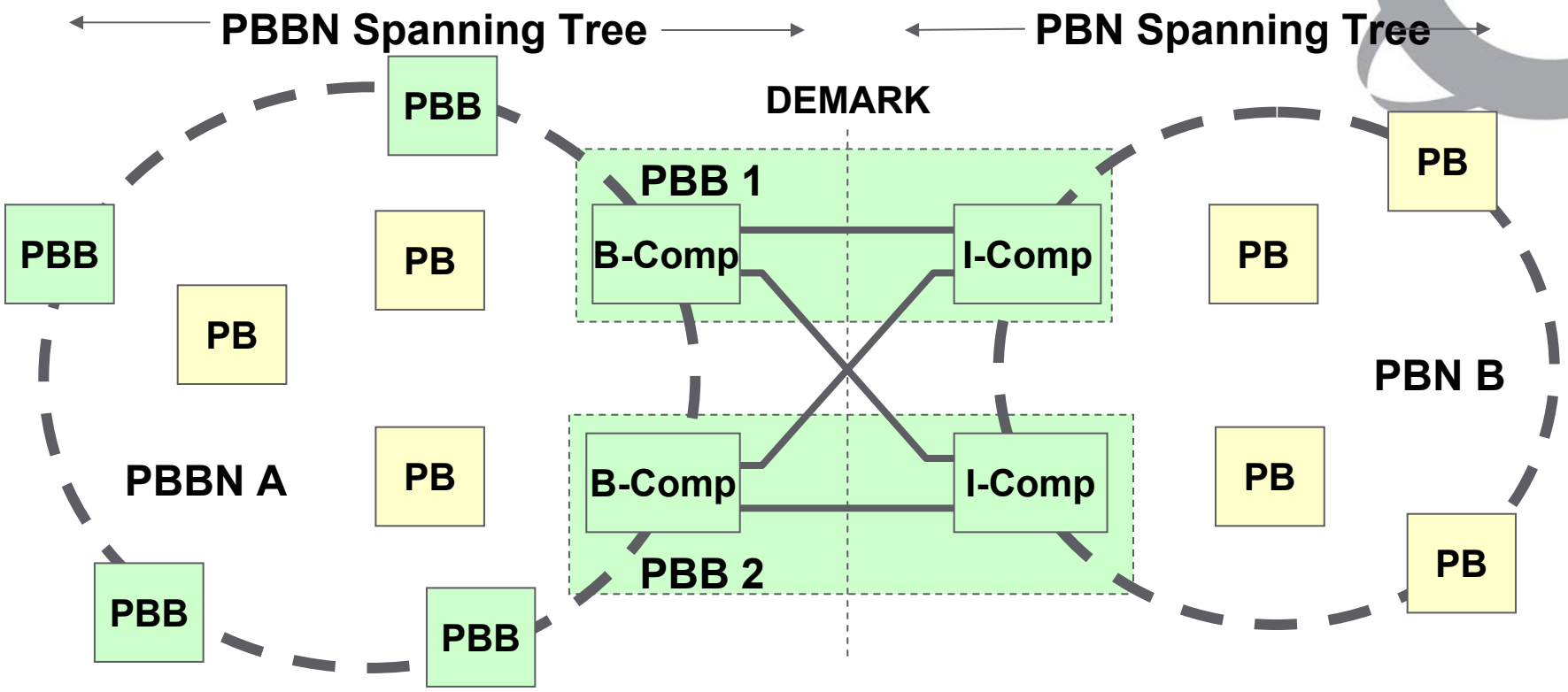
PBB Redundant Interconnect: Class 2



> Redundant PBBs allow recovery from both node and link failures

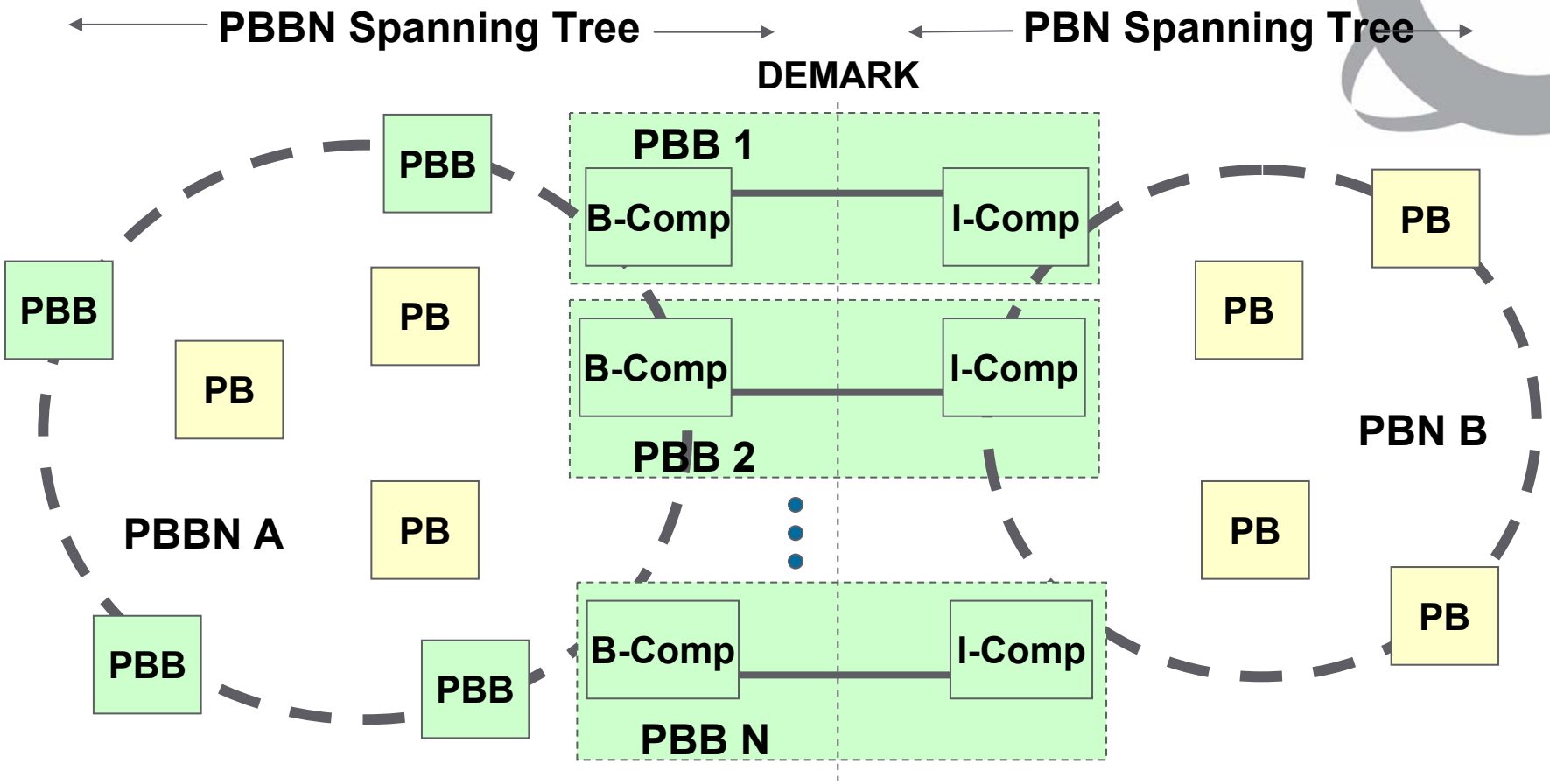
> Multiple links may be used between B-Comp and I-Comp

PBB Redundant Interconnect: Class 3



- Cross connection of PBB1 & PBB2 allows alternate paths through B-Comp or I-Comp
- Advantage is faults in PBBN or PBN may be isolated from affecting the state of the network on the other side of the redundant interconnect
- Disadvantage is the requirement for additional links between PBB1 and PBB2
- Must implement option 1 as a subset since failure of a complete PBB will still require state changes within the network on the other side of the interconnect

Generalized Class 2

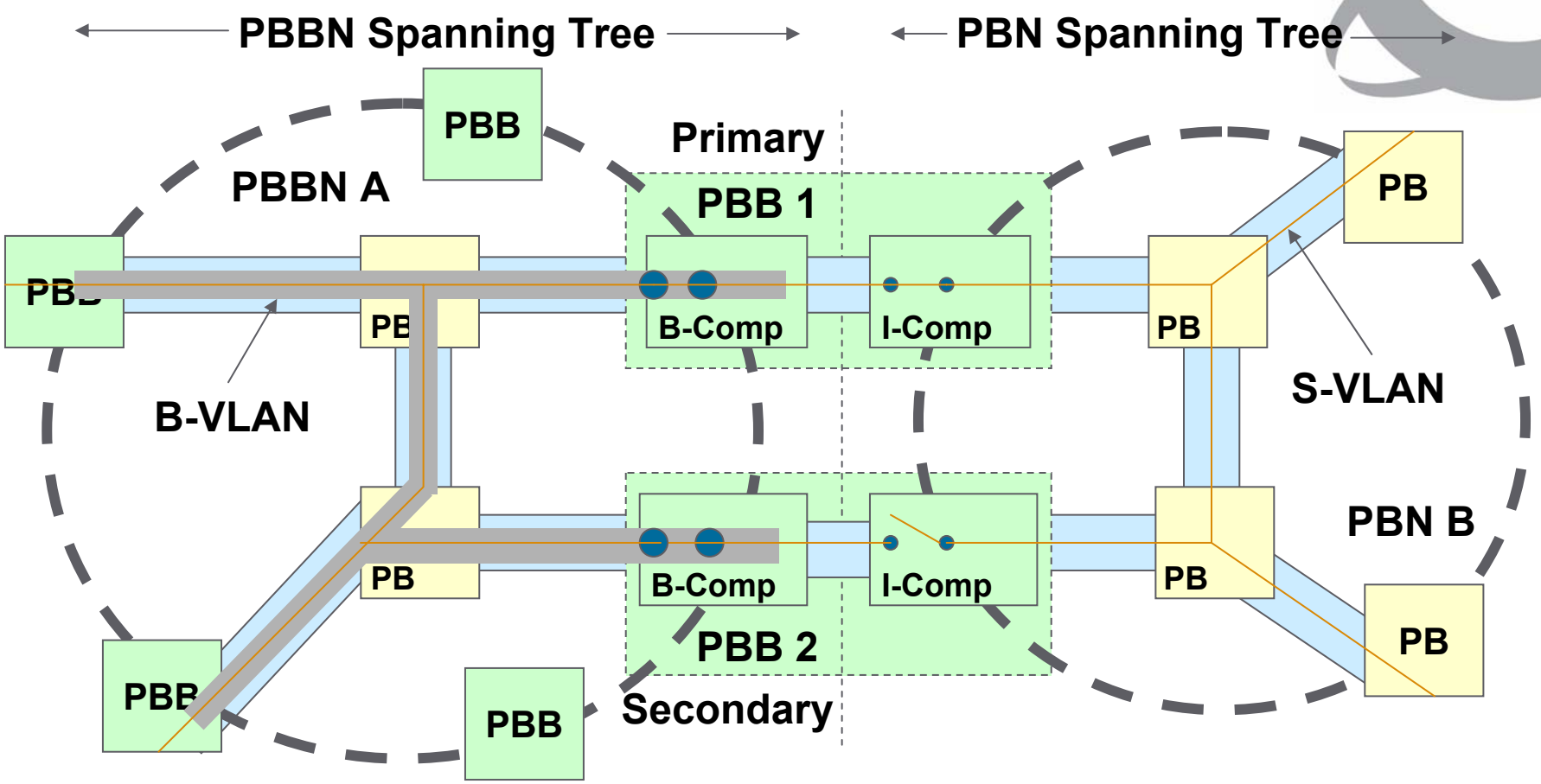


> Both class 2 & 3 may have N redundant PBBs

> Class 2 is a building block for class 3

> This presentation will focus on class 2

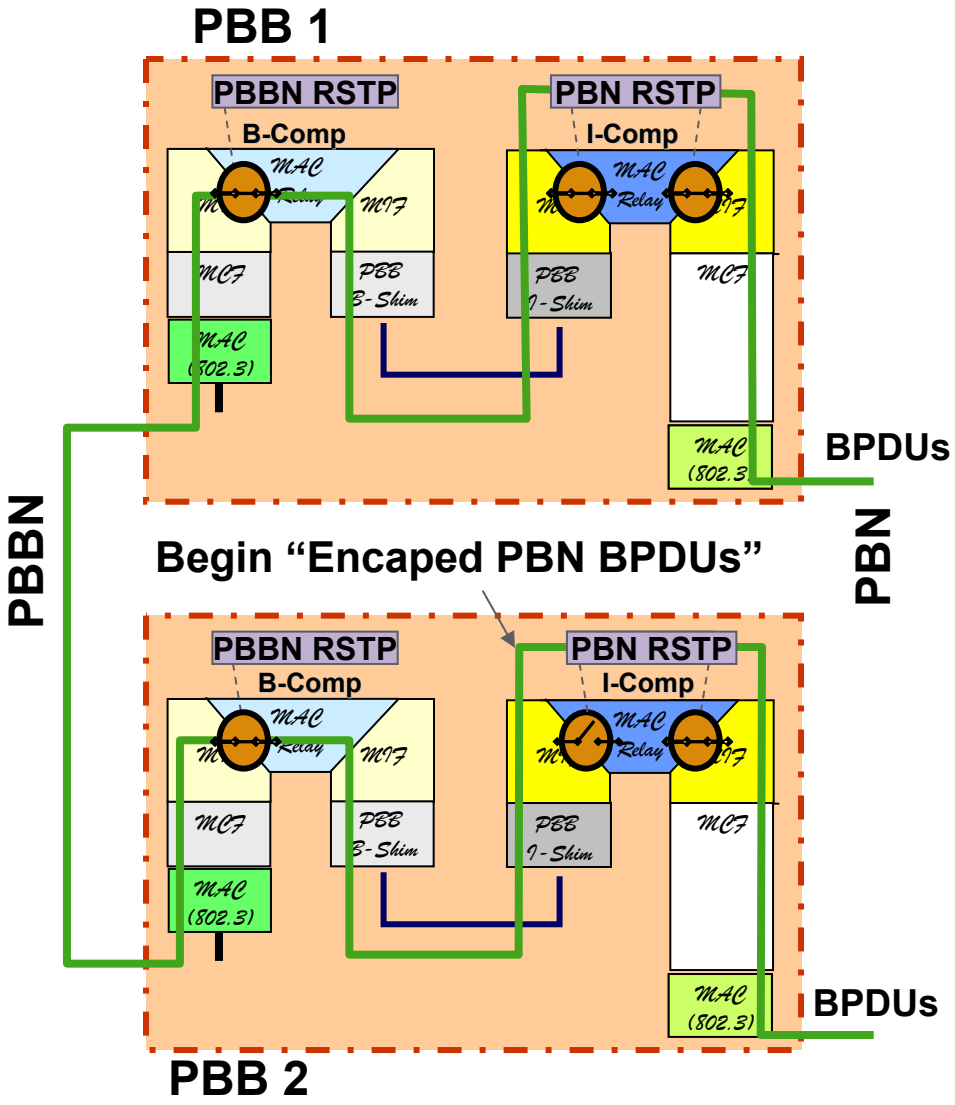
PBB Redundant Interconnect Blocking



> PBBN forwarding may be blocked at the B-Comp

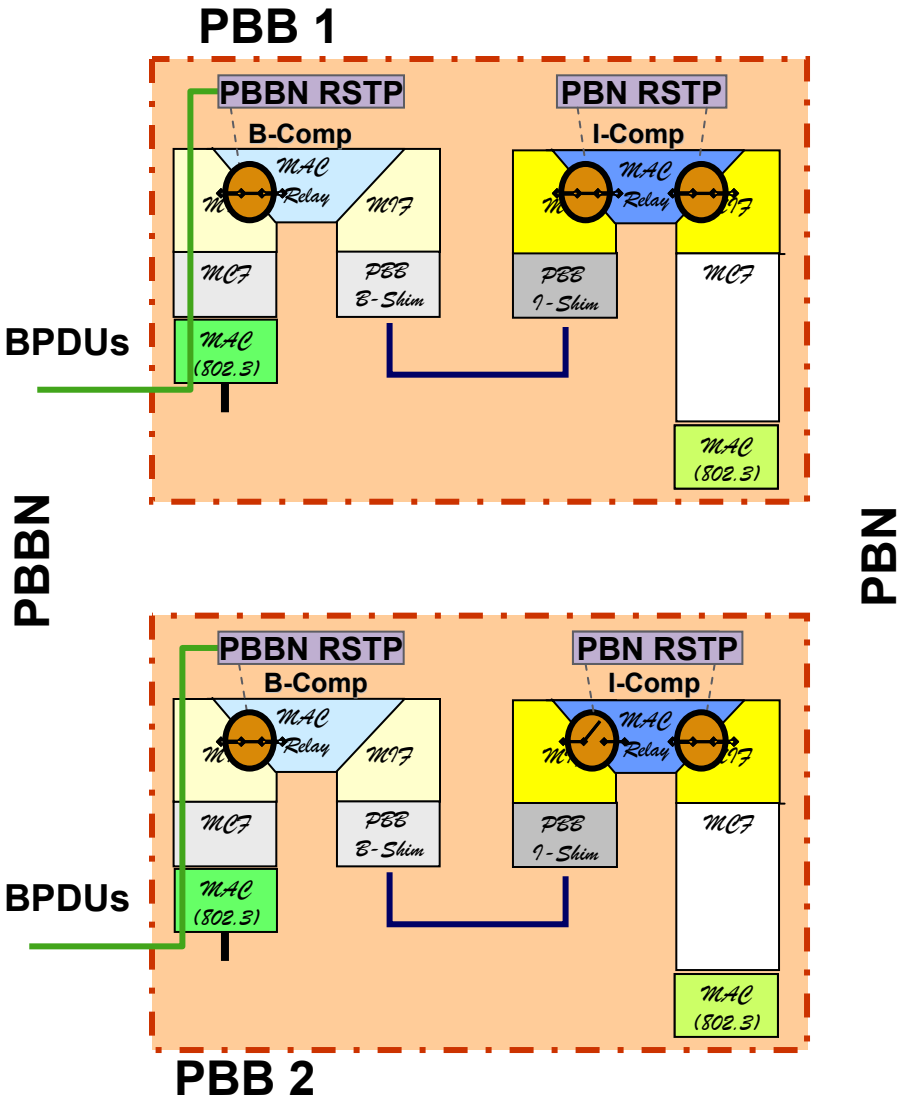
> PBN forwarding may be blocked at the I-Comp

Spanning Tree Handling: Option 1



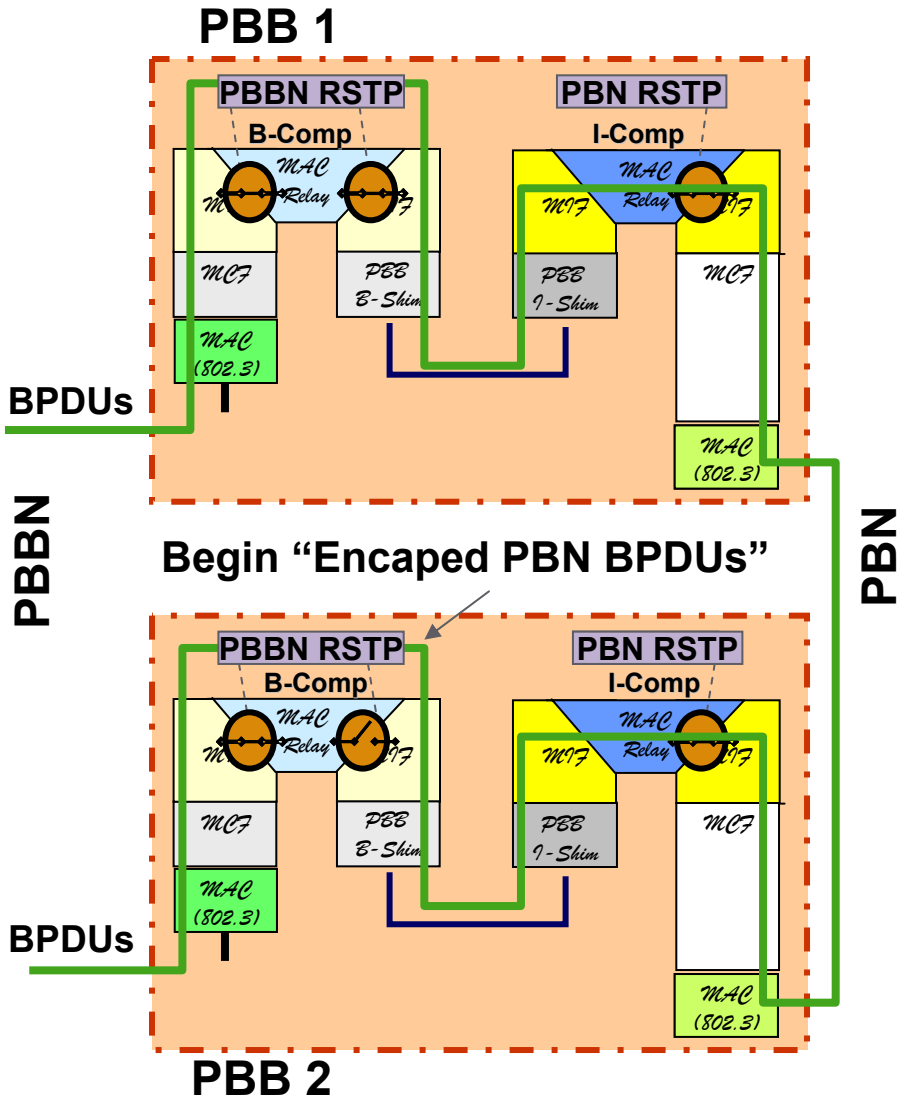
1. All PBN S-VLANs are provisioned through the I-Component of both PBB 1 and PBB 2
2. I-Components of PBBs participate in PBN RSTP protocol
3. B-Components of PBBs DO NOT participate in PBN RSTP protocol
4. PBN BPDUs from inward I-Comp port are forwarded between PBB 1 and PBB 2 I-Comps via a "Encapsuled PBN BPDUs"
5. The encapsuled PBN BPDUs are is passed over the PBBN
6. Each PBN has a dedicated BPDU S-VLAN which only connect the PBBs attaching to the PBN
7. PBN RSTP resolves loops within PBN network elements

Spanning Tree Handling: Option 1



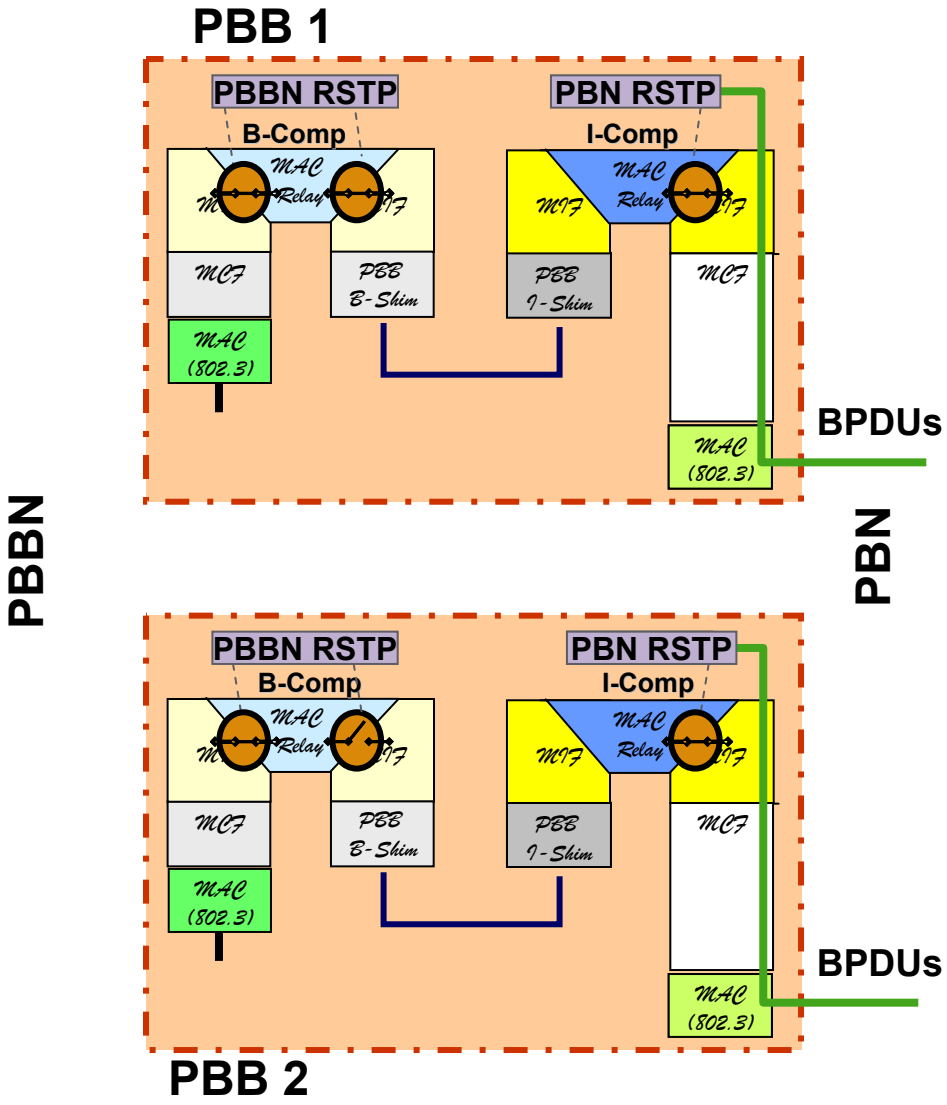
7. B-Component elements of PBB participate in PBBN xSTP protocol
8. PBB B-VLAN is routed by PBBN RSTP to the B-Component of PBB 1 and PBB 2 respectively
9. PBBN RSTP control frames not forwarded to corresponding PBB I-Component (i.e., blocked by B-shim)
10. PBBN RSTP resolves loop within PBBN network elements treating the B-Shim as an end station port

Spanning Tree Handling: Option 2



1. All PBN S-VLANs are provisioned through the I-Component of both PBB 1 and PBB 2
2. I-Components of PBBs participate in PBN RSTP protocol
3. B-Components of PBBs participate in PBN RSTP protocol
4. PBN BPDUs from inward I-Comp port are forwarded between PBB 1 and PBB 2 I-Comps via a "Encapsuled PBN BPDUs"
5. The encapsuled PBN BPDUs are is passed over the PBBN
6. Each PBN has a dedicated BPDU S-VLAN which only connect the PBBs attaching to the PBN
7. PBN RSTP resolves loops within PBN network elements

Spanning Tree Handling: Option 2

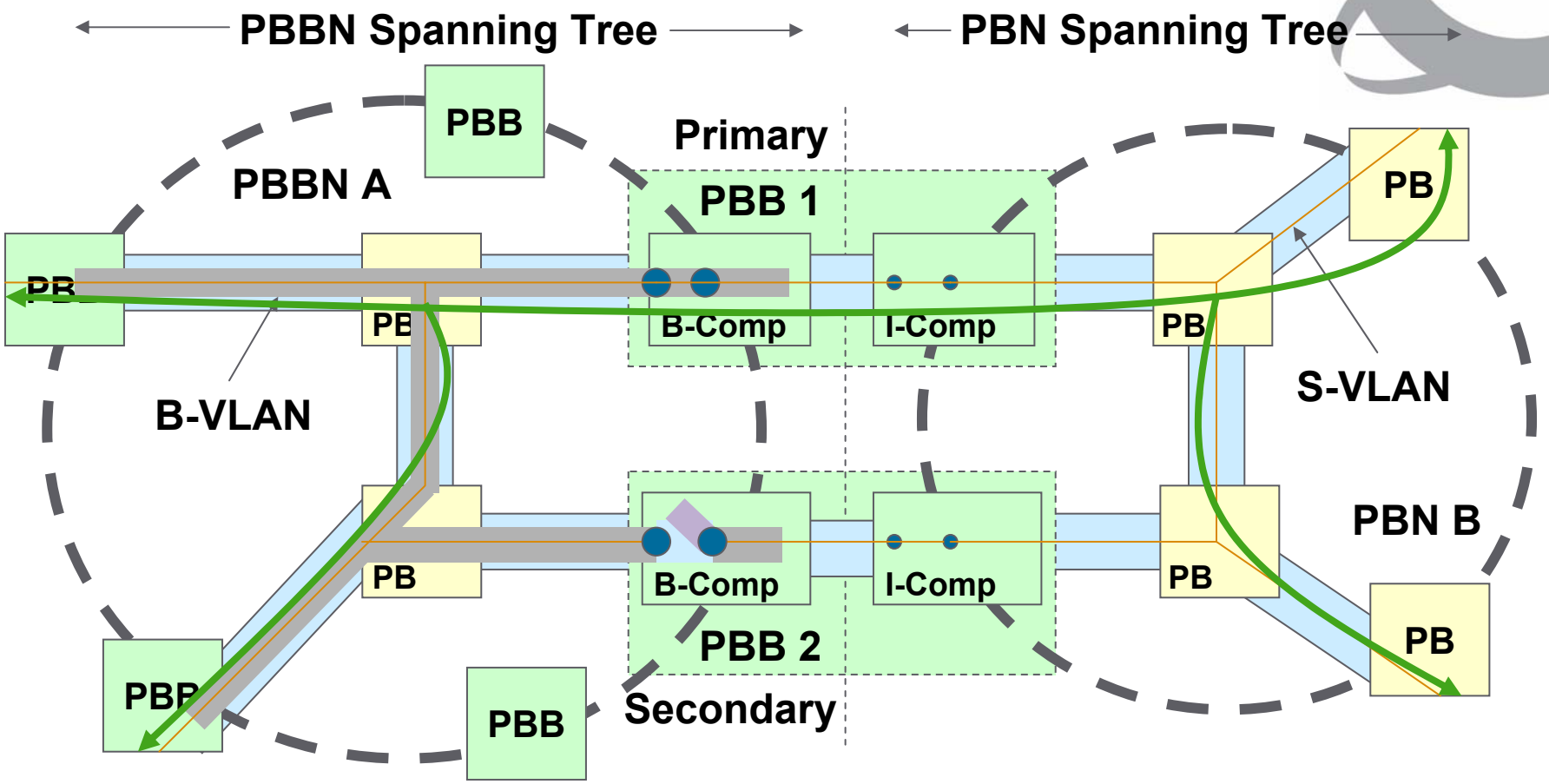


7. I-Component elements of PBB participate in PBN RSTP protocol
8. PBN S-VLAN is routed by RSTP to the I-Component of PBB 1 and PBB 2 respectively
9. PBN RSTP control frames not forwarded to corresponding PBB B-Component (i.e., blocked by I-shim)
10. PBN RSTP resolves loops within PBN network elements treating the I-Shim as an end station port

Encapsuled BPDUs

- Encapsuled BPDUs use a well known multicast which is NOT in the reserved address for bridges.
- The Encapsuled BPDU should include some management information designating the global name of the PBBN and PBN connected by the PBBs
- When sending Encapsuled BPDUs to the PBN the BPDUs either need to have a private S-VLAN between the PBBs or should be filtered at the edge of the PBN.
- When sending Encapsuled BPDUs to the PBBN the BDPUs needs to have a private S-VLAN between the PBBs or needs information which allows the PBBs to filter frames for their group.

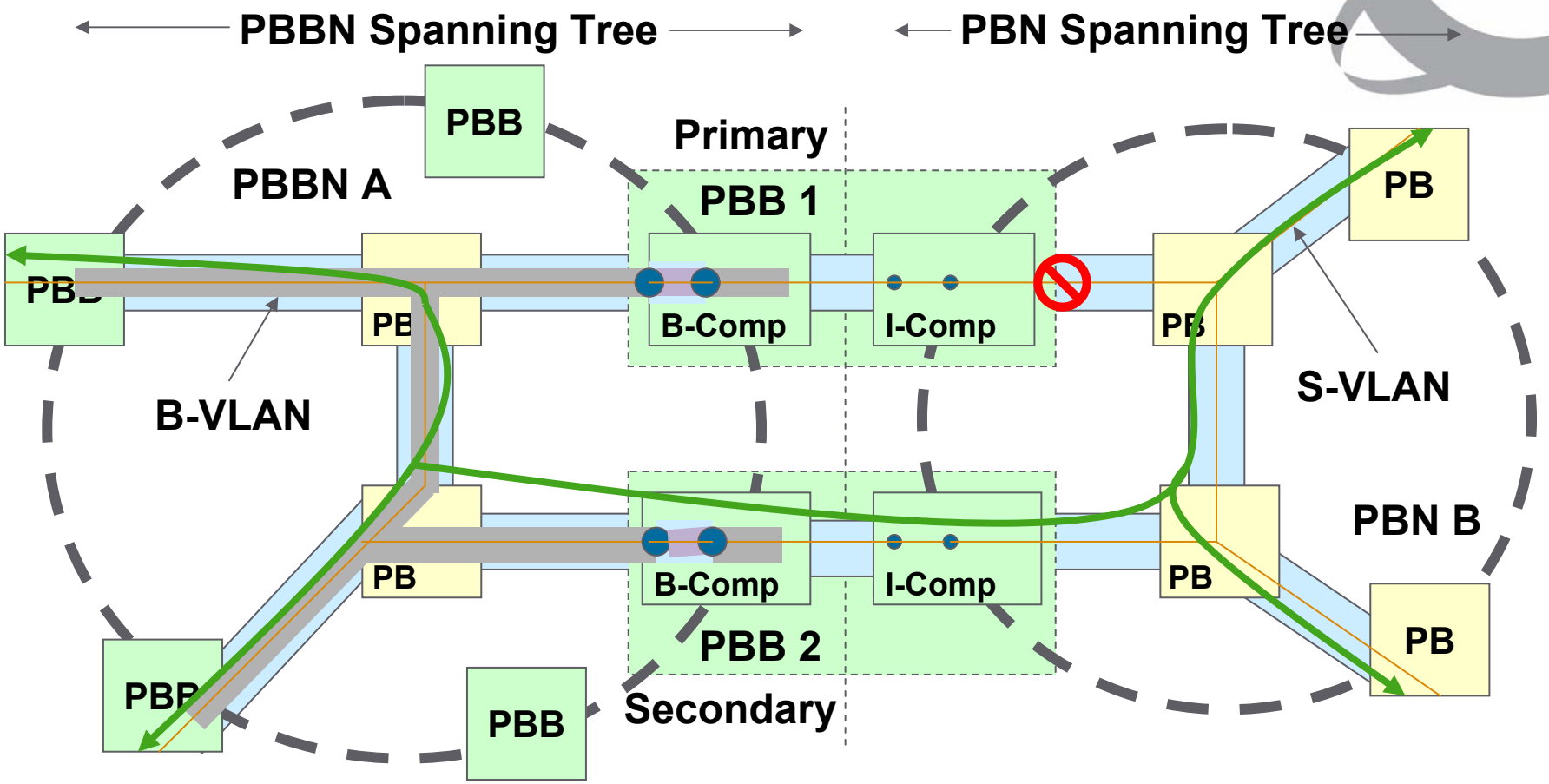
PBB Interconnect: Normal



> PBB 1 B-Comp has been selected as the designated bridge for the PBBN loop through PBN forwarding mode

> The PBBN loop is set to very high cost to prevent spanning tree forwarding through PBBN loop

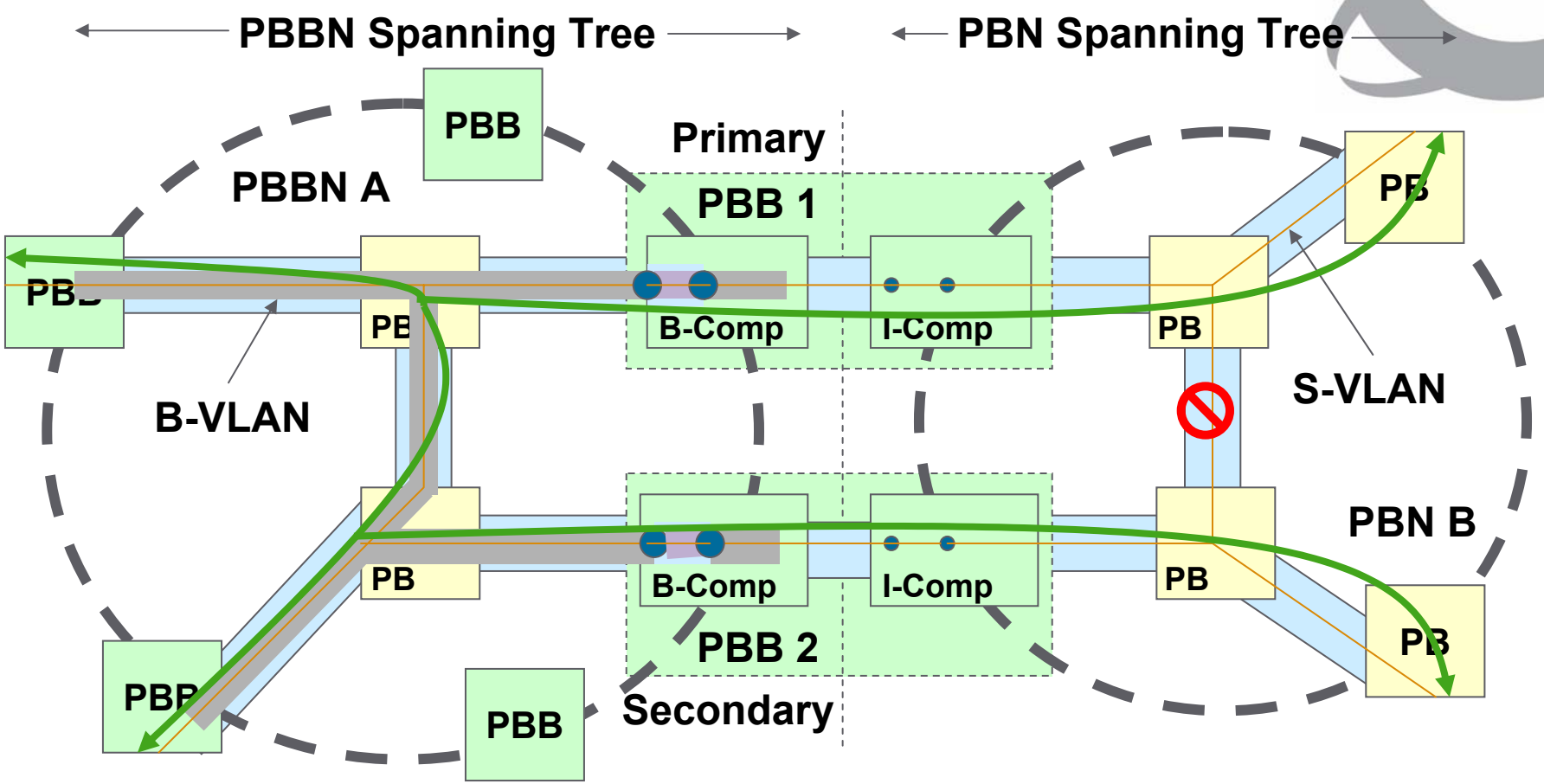
PBB Interconnect: Simple Fault



> PBB 1 B-Comp looses connectivity to PBB 2 through PBN

> PBB 2 Stops seeing the connection to PBB 1 and therefore becomes the designated bridge for the PBN loop

PBB: Segmentation Fault

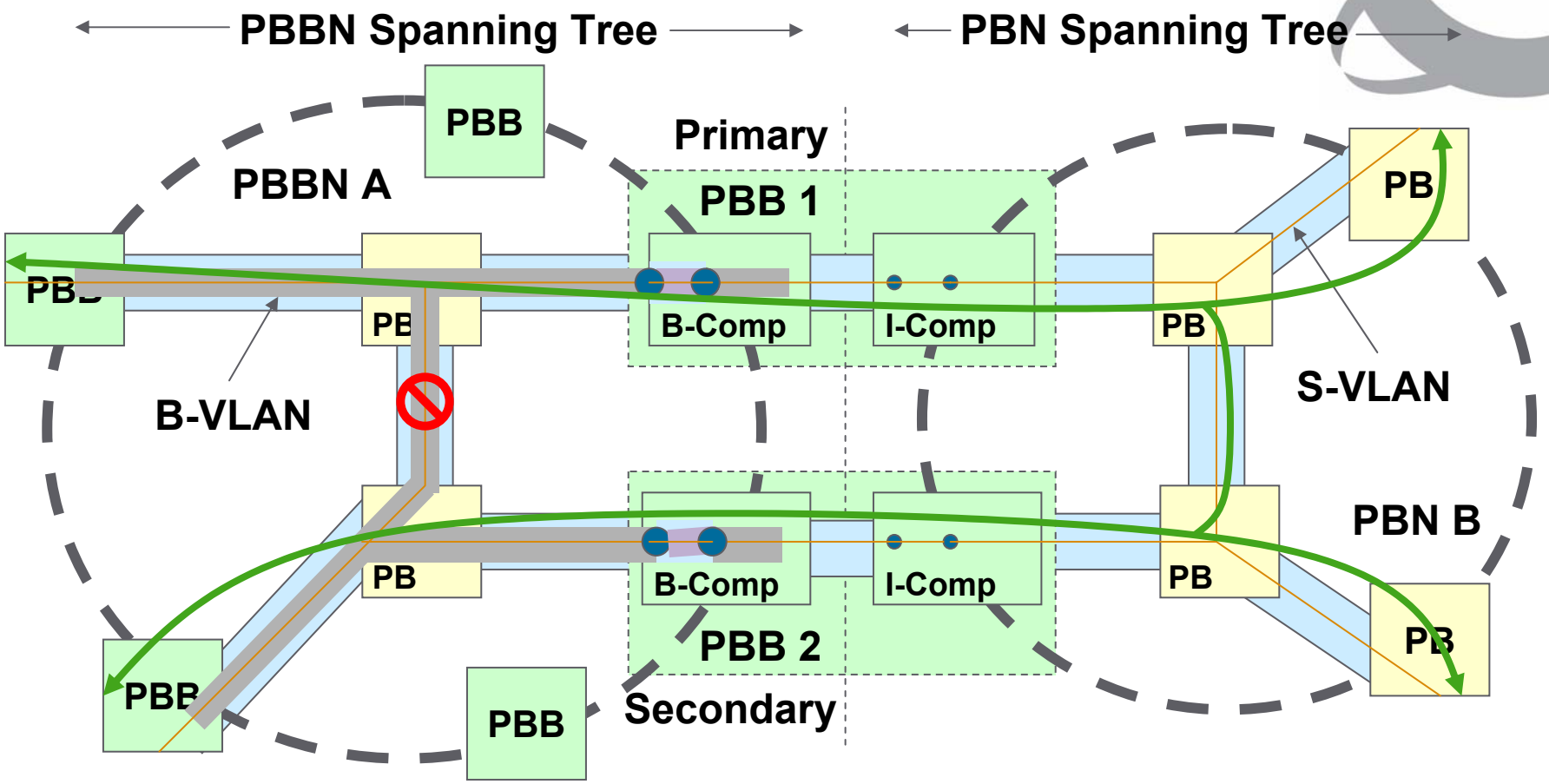


> PBB 1 B-Comp looses connectivity to PBB 2 through PBN

> PBB 2 looses connection to PBB 1 and therefore designates a port facing the PBN loop

> PBN is segment and is reconnected through PBBN

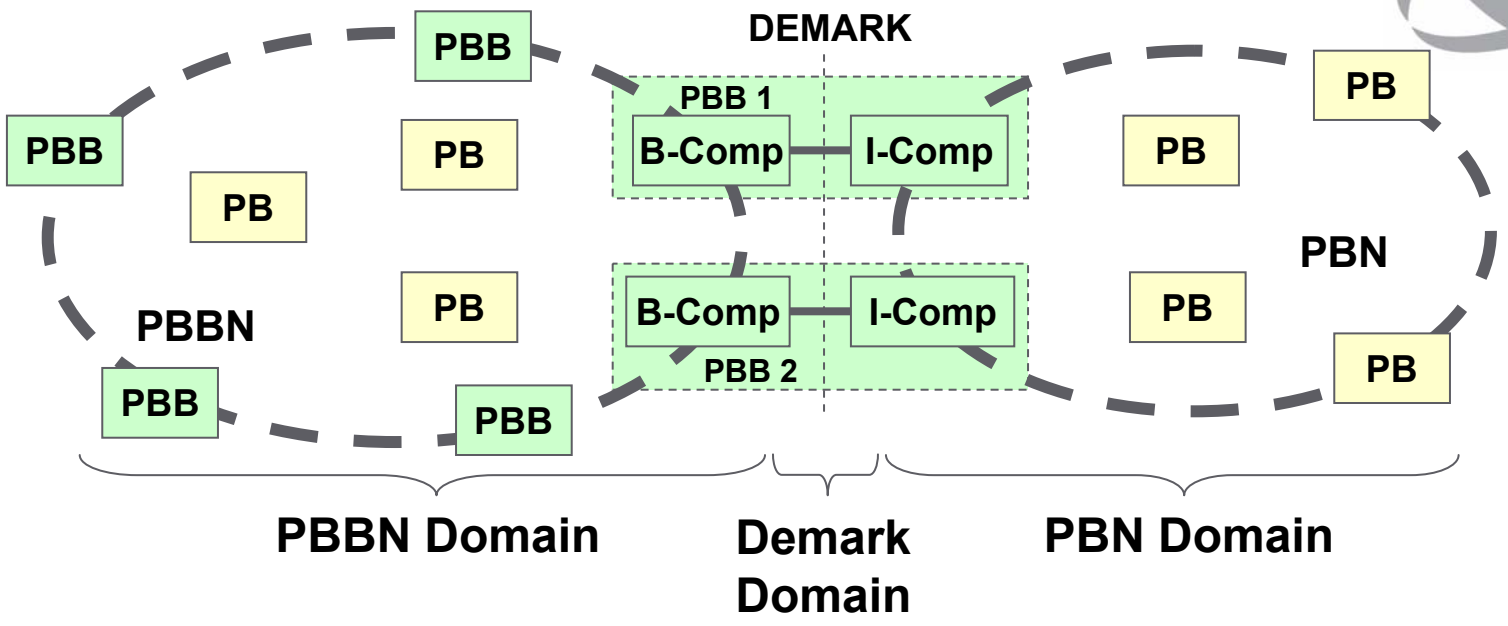
PBB: Segmentation Fault 2



> PBB 1 B-Comp looses connectivity to PBB 2 through PBBN

> PBB 2 goes to fowarding frames destine for PBB1 over PBN which is the only remaining path

802.1ag Based Fault Detection



> Maintenance Domains operating over default VLAN cover call switches in PBBN and PBN

> Add PBB TLV to CC messages to allow discovery of all PBBs and the networks they connect

> Propagate connectivity data by CC

> Propagate error data by AIS

Recommendations

- Add mechanism for redundant interconnect to 802.1ah draft
- Method added should be for Class 2 redundant interconnect
- The spanning tree extension method should be specified
- Future work should inspect 802.1ag protection switching methods

Backup Slides