Shortest Path Bridging IEEE 802.1aq Tutorial and Demo

NANOG 50 Oct 2010

Peter Ashwood-Smith
Fellow

peter.ashwoodsmith@huawei.com

Abstract

802.1aq Shortest Path Bridging is being standardized by the IEEE as an evolution of the various spanning tree protocols. 802.1aq allows for true shortest path routing, multiple equal cost paths, much larger layer 2 topologies, faster convergence, vastly improved use of the mesh topology, single point provisioning for logical membership (E-LINE/E-LAN/E-TREE etc), abstraction of attached device MAC addresses from the transit devices, head end and/or transit multicast replication , all while supporting the full suit of 802.1 OA&M.

Applications consist of STP replacement, Data Center L2 fabric control, L2 Internet Distributed Exchange point fabric control, small to medium sized Metro Ethernet control planes. L2 wireless network backhaul….

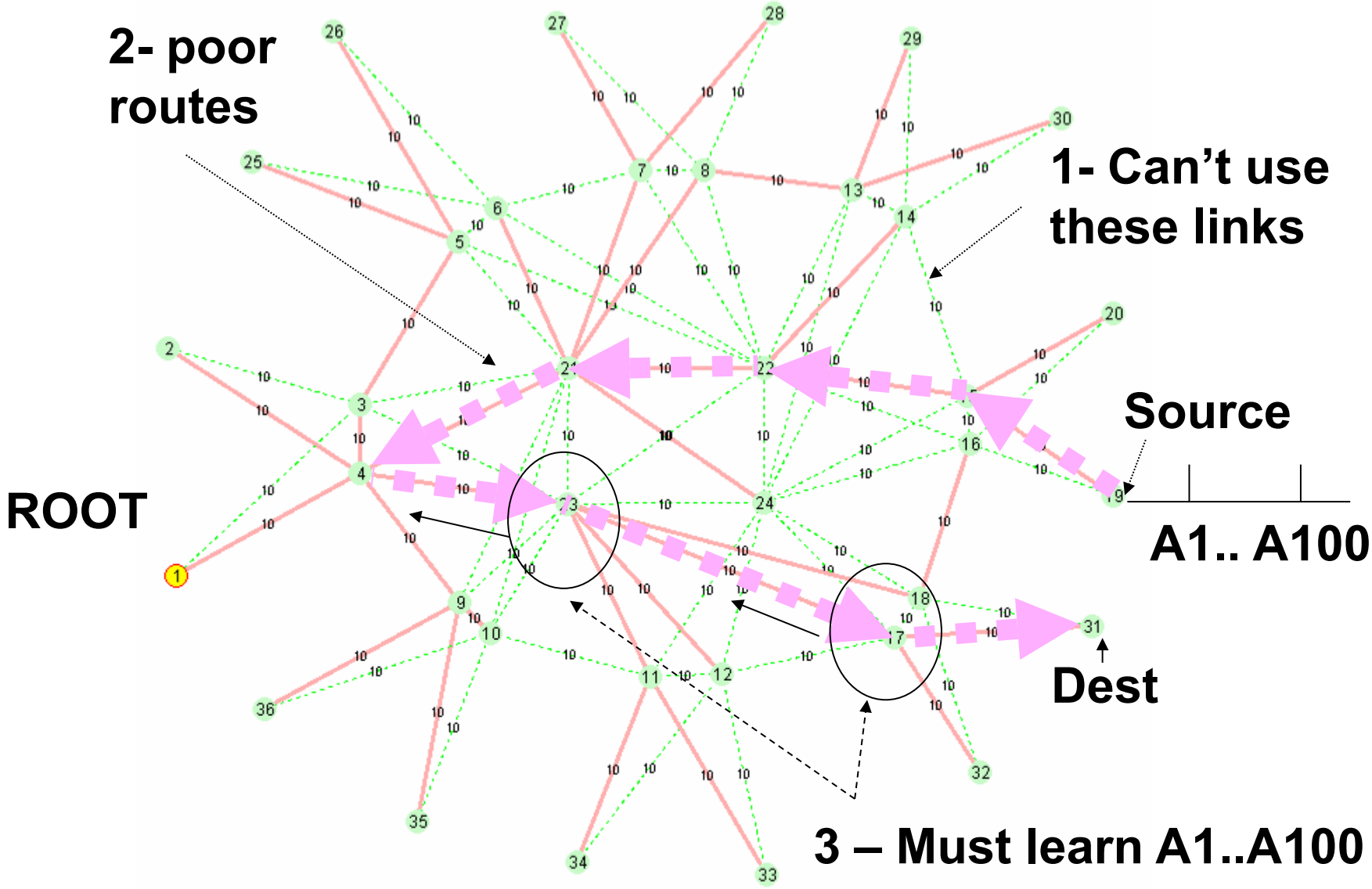
Outline

- **Challenges**
- What is 802.1aq/SPB
- Applications
- How does it work
- Example with hopefully live demo from my lab (assuming connectivity available)

Challenges

- L2 networks that scale to ~1000 bridges.
- Use of arbitrary mesh topologies.
- Use of (multiple) shortest paths.
- Efficient broadcast/multicast routing and replication points.
- Avoid address learning by tandem devices.
- Get recovery times into 100's of millisecond range for larger topologies.
- Good scaling without loops.
- Allow creation of very many logical L2 topologies (subnets) of arbitrary span.
- Maintain all L2 properties within the logical L2 topologies (transparency, ordering, symmetry, congruence, shortest path etc).
- Reuse all existing Ethernet OA&M 802.1ag/Y.1731

Example STP 36 nodes



Outline

- Challenges
- **What is 802.1aq/SPB**
- Applications
- How does it work
- Short Demo (remote switches if possible)

What is 802.1aq/SPB

- **IEEE protocol builds on 802.1 standards**
- **A new control plane for Q-in-Q and M-in-M**
  - Leverage existing inexpensive ASICs
  - Q-in-Q mode called SPBV
  - M-in-M mode called SPBM
- **Backward compatible to 802.1**
  - 802.1ag, Y.1731, Data Center Bridging suite
- **Multiple loop free shortest paths routing**
  - Excellent use of mesh connectivity
  - Currently 16, path to 1000's including hashed per hop.
- **Optimum multicast**
  - head end or tandem replication

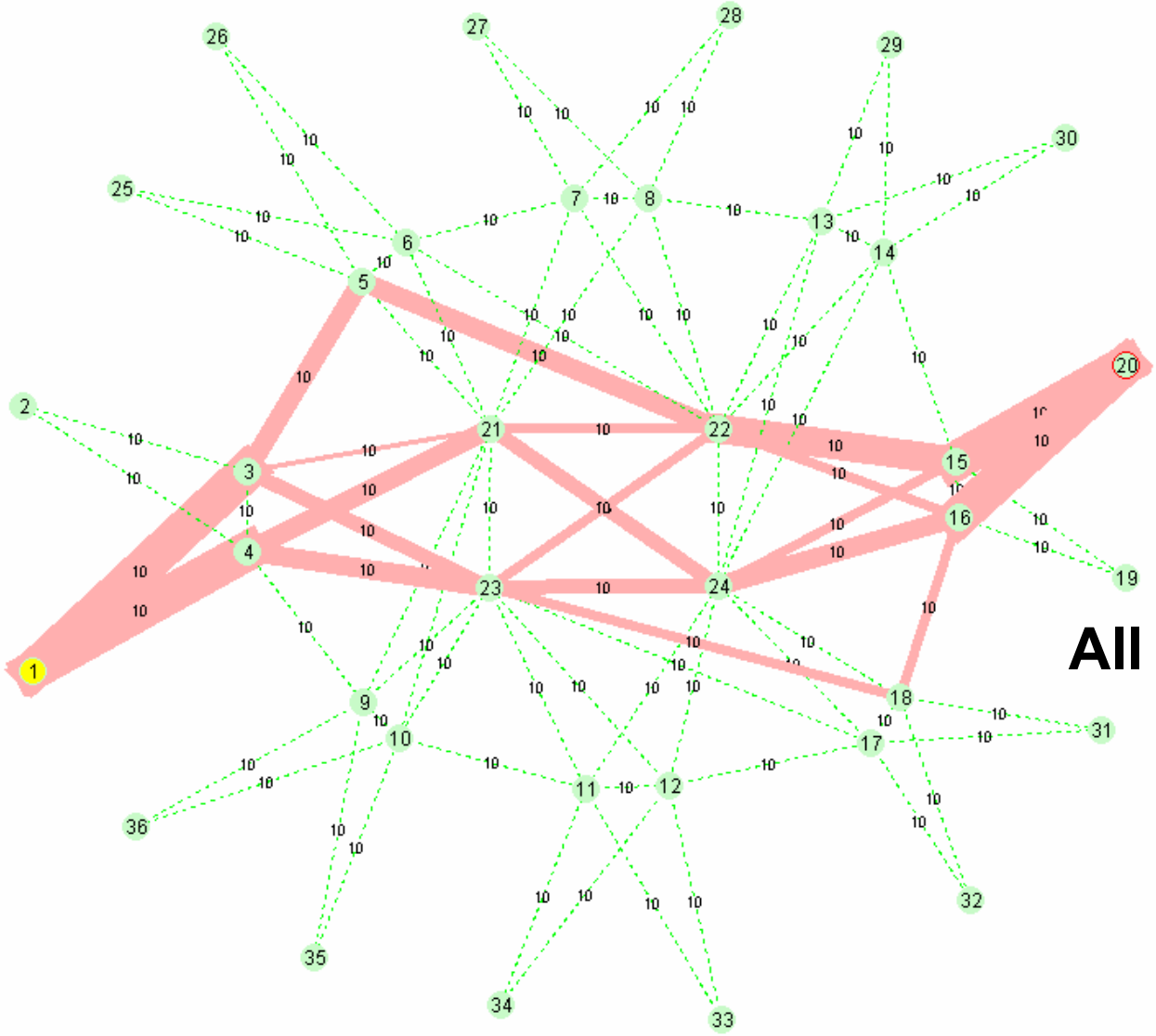
What is 802.1aq/SPB (cont'd)

- **Light weight form of traffic engineering**
  - Head end assignment of traffic to 16 shortest paths.
  - Deterministic routing - offline tools predict exact routes.
- **Scales to ~1000 or so devices**
  - Uses IS-IS already proven well beyond 1000.
  - Huge improvement over the STP scales.
- **Good convergence with minimal fuss**
  - sub second (modern processor, well designed)
  - below 100ms (use of hardware multicast for updates)
  - Includes multicast flow when replication point dies. Pre-standard seeing 300ms recovery @ ~50 nodes.
- **IS-IS**
  - Operate as independent IS-IS instance, or within IS-IS/ IP, supports Multi Topology to allow multiple instances efficiently.

What is 802.1aq/SPB (cont'd)

- **Membership advertised in same protocol as topology.**
  - Minimizes complexity, near plug-and-play
  - Support E-LINE/E-LAN/E-TREE
  - All just variations on membership attributes**.**
- **Address learning restricted to edge (M-in-M)**
  - FDB is computed and populated just like a router.
  - Unicast and Multicast handled at same time.
  - Nodal or Card/Port addressing for dual homing.
- **Computations guarantee ucast/mcast…**
  - Symmetry (same in both directions)
  - Congruence (unicast/multicast follow same route)
  - Tune-ability (currently 16 equal costs paths – opaque allows more)

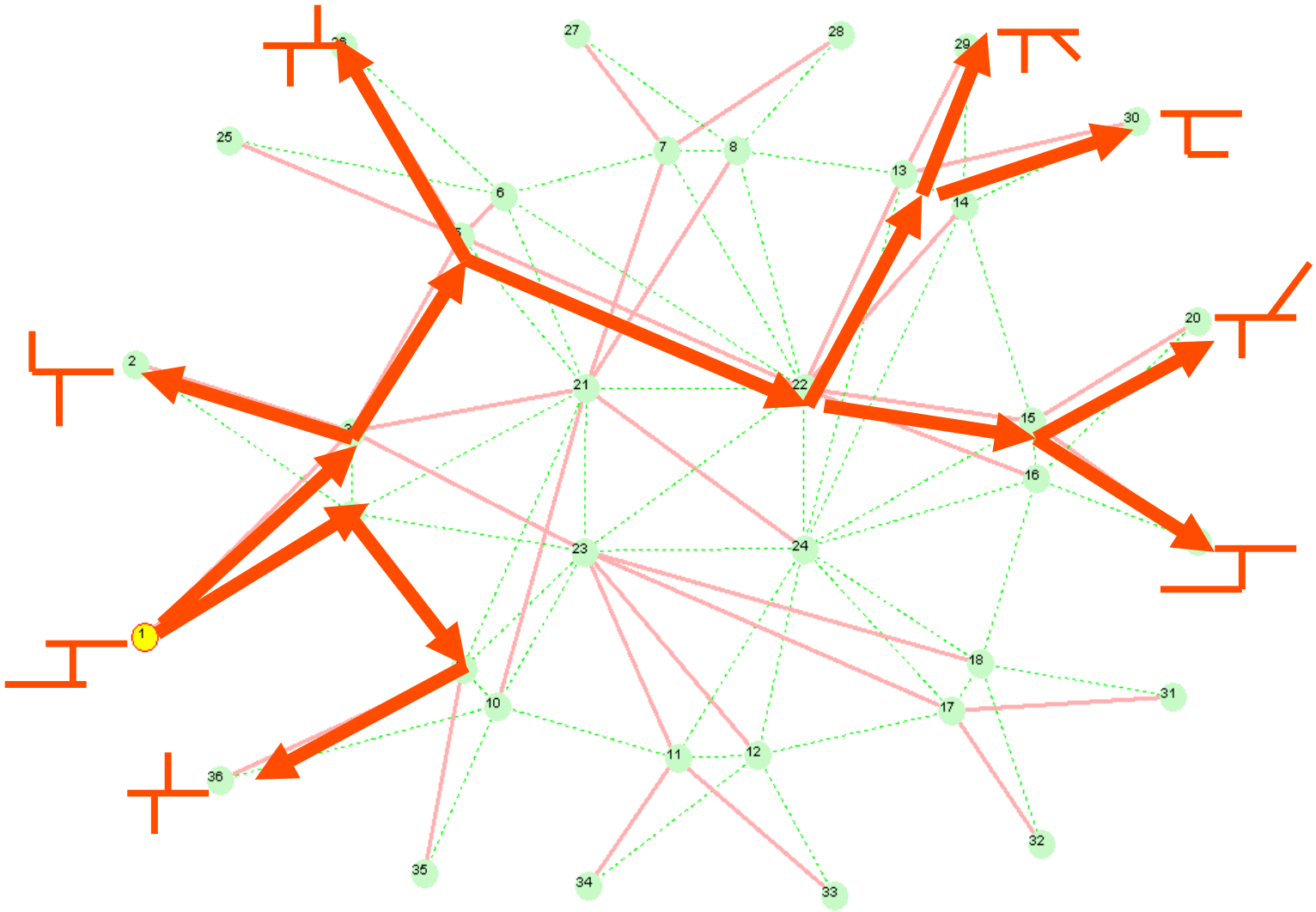
End result - Visually



All links usable

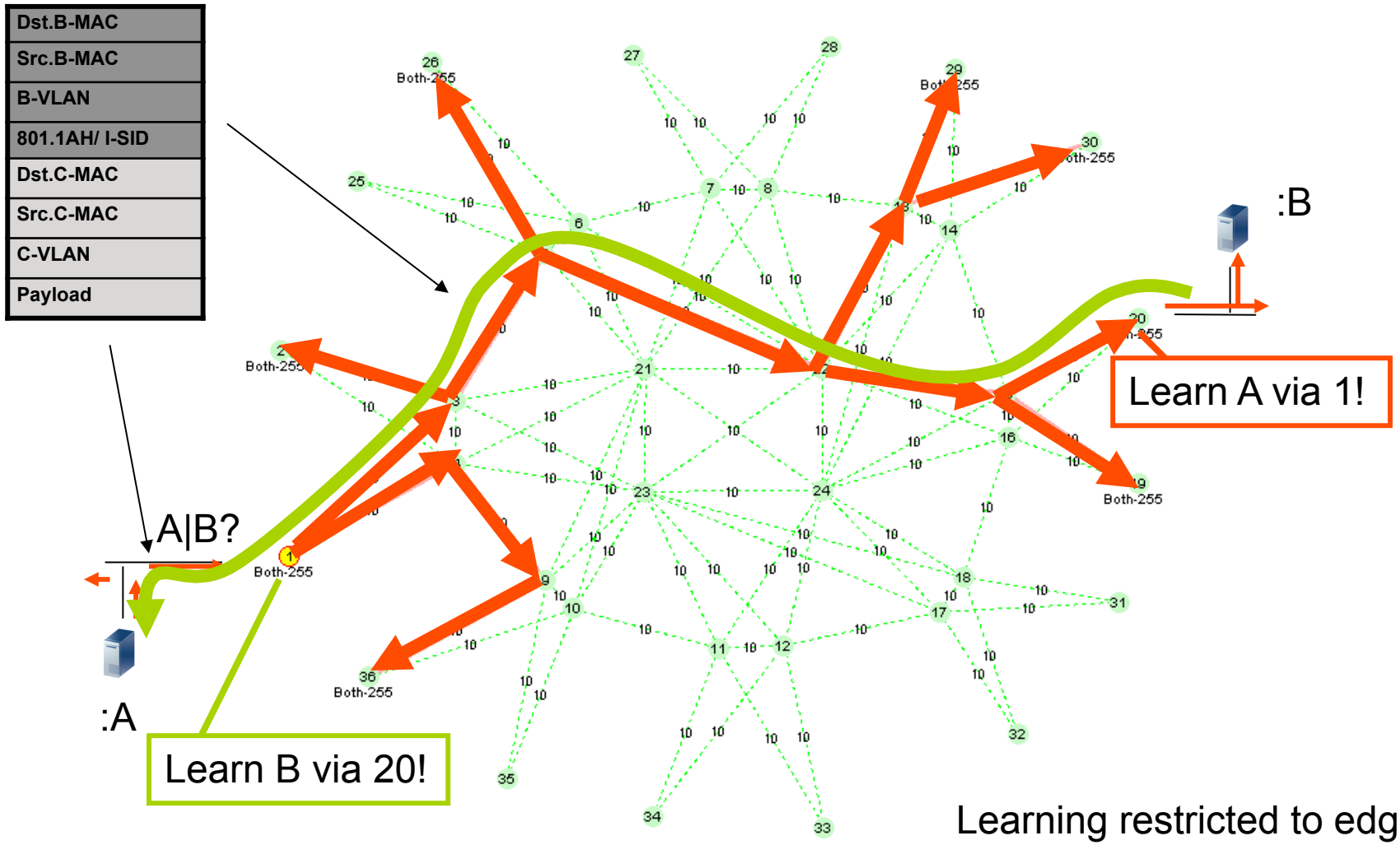
Multiple Shortest Path routing + Ethernet OA&M

SPF trees form multicast template



Shortest Path First Tree becomes template for multicast tree and is pruned automatically to proper membership.

Edge Learning - Visually



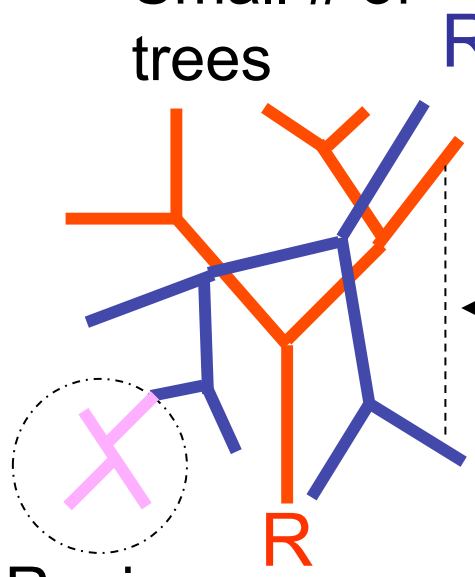
Learning restricted to edges and only where I-SID tree reaches. Mac-in-Mac encap.

Outline

- Challenges
- What is 802.1aq/SPB
- **Applications**
- How does it work
- Short Demo (remote switches if possible)

Application (M|R)STP replacement

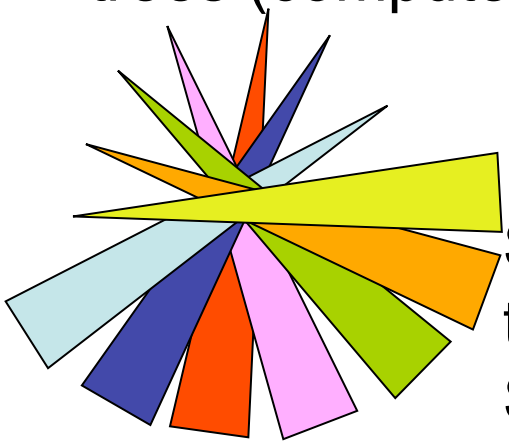
Small # of trees



Unused links

Regions for scale

Large number of source specific trees (computed).



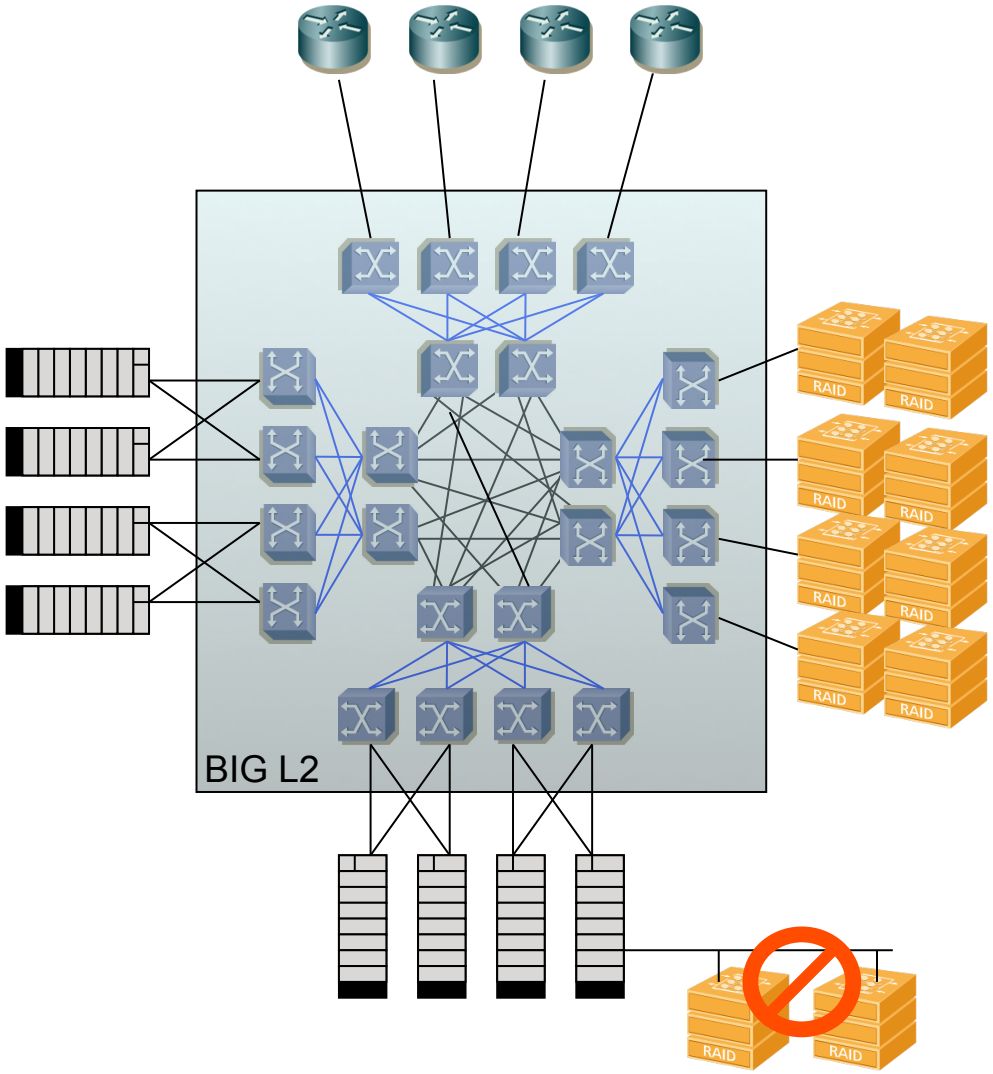
All links usable

No regions needed

Scoped exactly to C-VLAN or S-VLAN members

- Many more nodes without regions
- Low effort to get good routing
- Fast convergence – link state v.s. distance vec
- Address isolation m-in-m.

Data Center - trends

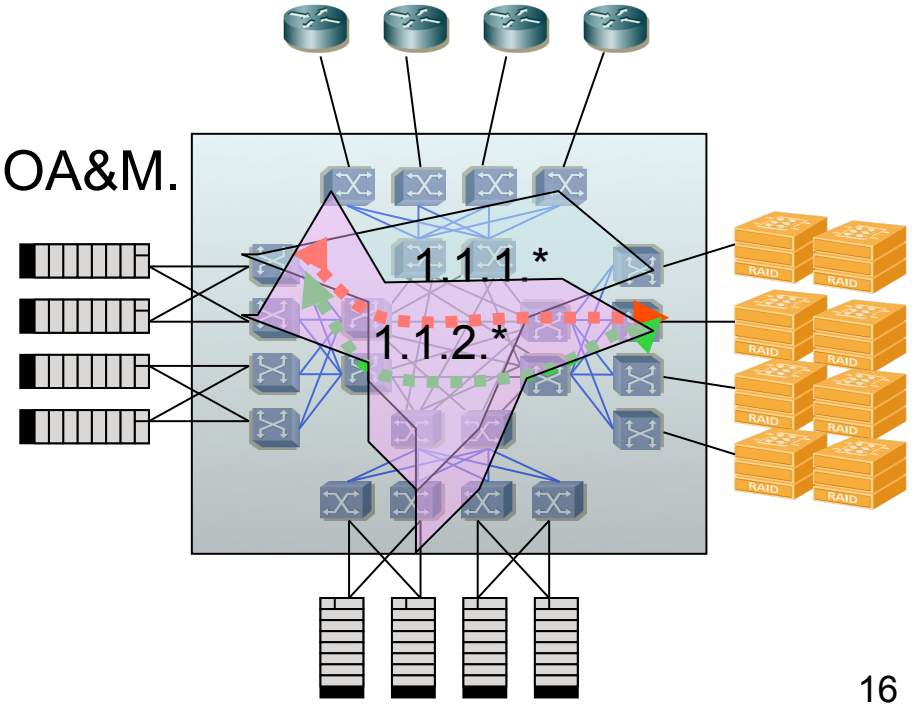


Treat DC network as
one big L2 switch by combining 100's of smaller switches in 'non blocking' topology – why?

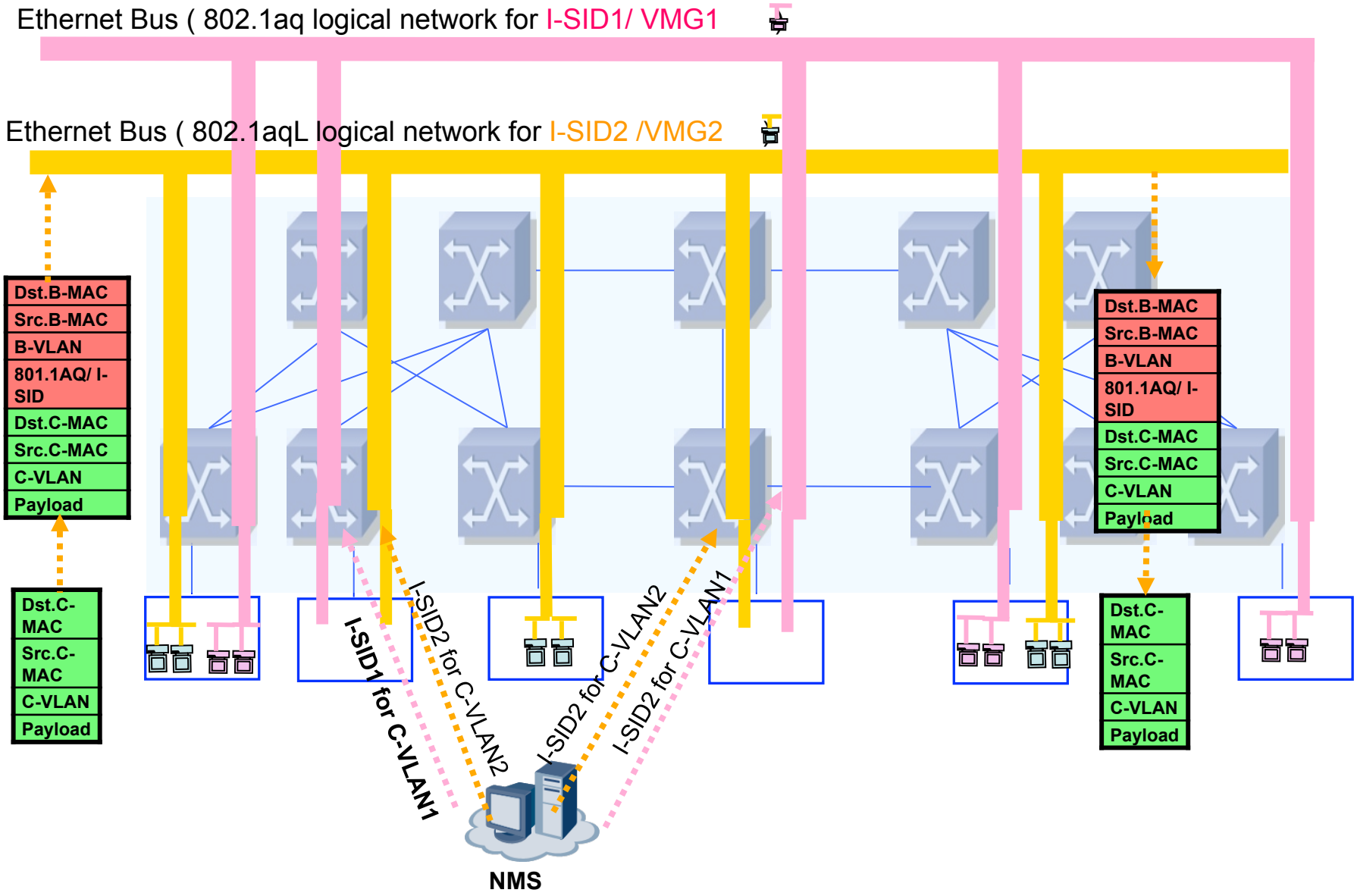
- Any server anywhere.
- Any router anywhere.
- Any appliance anywhere.
- Any VM anywhere.
  - Any IP address anywhere.
  - Any subnet anywhere.
- Any storage anywhere.
- Minimal congestion issues.
- Total flexibility for power use

Application Data Center

- Multiple shortest path routing
  - inter server traffic
- Deterministic traffic flows.
- Flexible subnet – expand/shrink anywhere.
  - Virtualization operates in subnet.
- Fully compatible with all 802.1 Data Center Bridging protocols & OA&M.
- Address isolation through m-in-m
- Fast recovery
- No loops



Application Data Center VM ‘hot’ migration (no interruption)



Application Data Center (cont'd)

- Totally compatible with Vmware server functions:
  - OA&M, motion, backup etc.
  - Apps that sit on Vmware 'just work'.

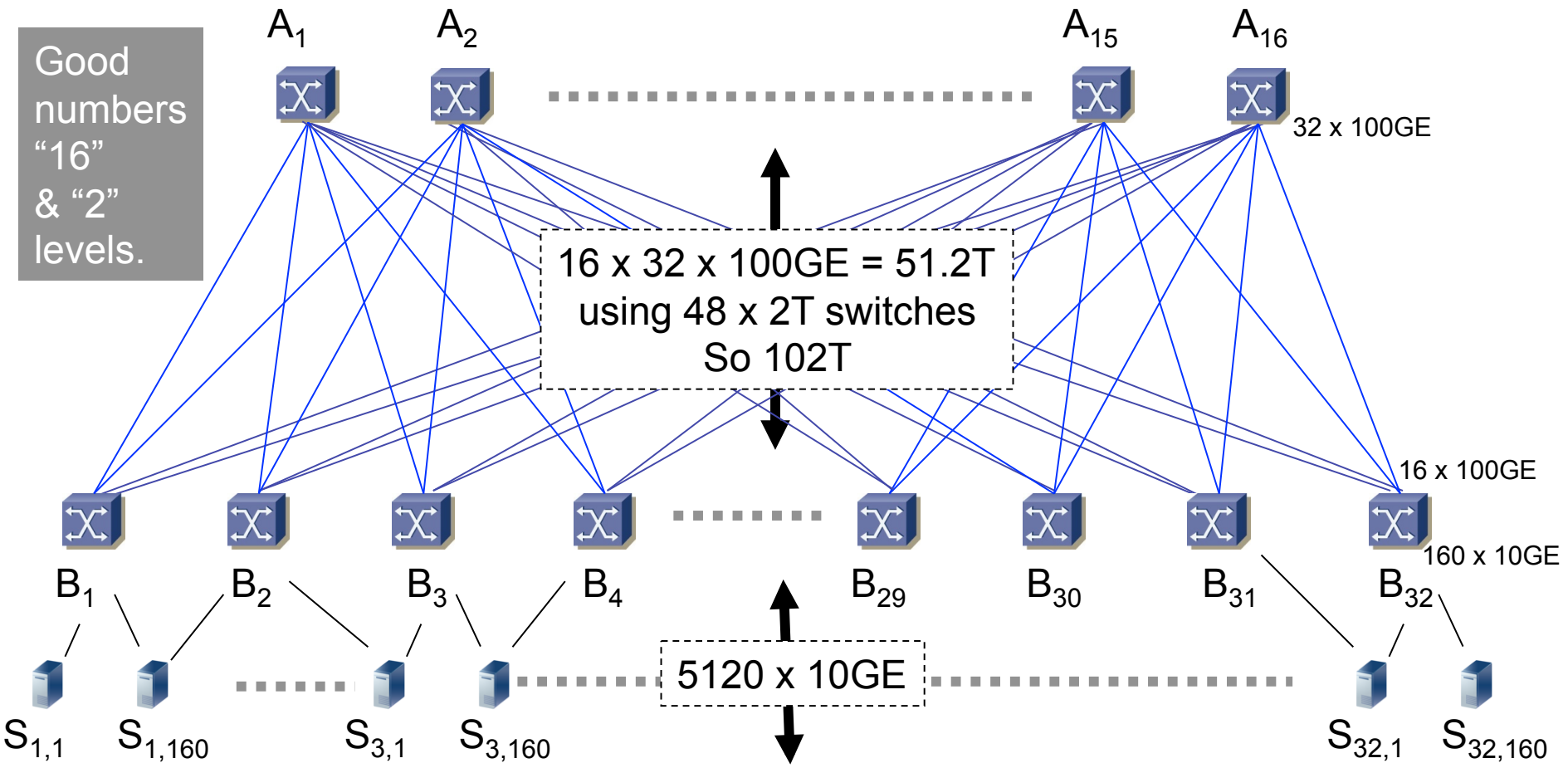
- Totally compatible with Microsoft load balancing (multicast over the L2)

- VRRP transparent (primary/stdby rtr per subnet) or proprietary variations on same protocol.

- It just makes the L2 part of the DC larger and better utilized.

- Compatible with emerging Inter DC overlay work.

Application High Performance switching cluster – assume 100GE NNI links



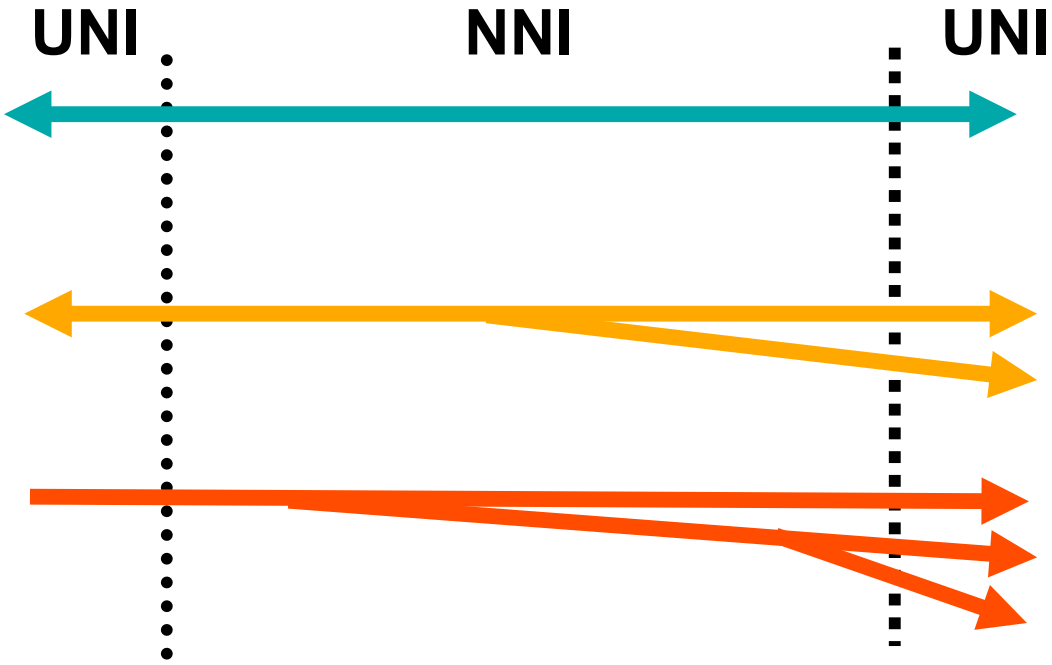
- 48 switch non blocking 2 layer L2 fabric
- 16 at "upper" layer **A_1..A_{16}**
- 32 at "lower" layer **B_1.. B_{32}**
- 16 uplinks per **B_n**, & 160 UNI links per **B_n**
- 32 downlinks per **A_n**

- (16 x 100GE per **B_n**)x32 = 512x100GE = 51.2T
- 160 x 10GE server links (UNI) per **B_n**
- (32 x 160)/2 = **2560 servers @ 2x10GE** per

100+ Terra non blocking interconnection fabric (if switches non blocking)

Application Metro/L2VPN

- Very light weight L2VPNs (2^24 data path) of:
  E-LAN, E-LINE, E-TREE flavors (a very cheap VPLS)
- Can do VPLS style head end replication
- Can do p2mp style transit replication (just one tx flag).
- Can support receive only membership (E-TREE)

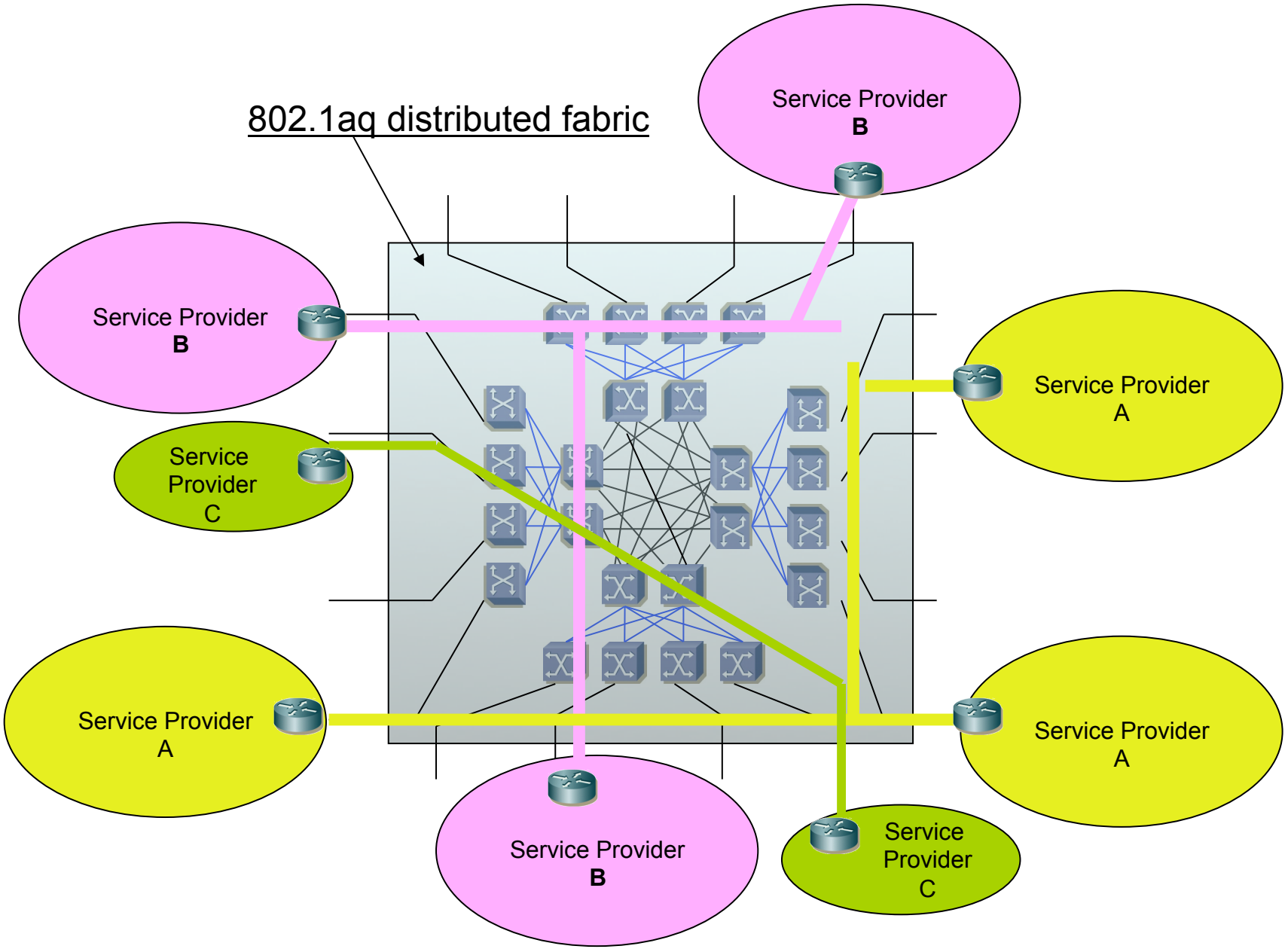


- E-LINE
= 2 tx/rx members

- E-LAN
> 2 tx/rx members

- E-TREE n rx members
1 tx member.

L2 Internet Exchange Point



Outline

- Challenges
- What is 802.1aq/SPB
- Applications
- **How does it work**

How does it work?

- **From Operators Perspective**
  - Plug NNI's together
  - Group ports/c-vlan/s-vlan at UNIs that you want to bridge (2^{24} groups='services' m-in-m mode.)
  - Assign an I-SID to each group..

- **Internally**
  - IS-IS reads box MAC, forms NNI adjacencies
  - IS-IS advertises box MACs (so no config).
  - IS-IS reads UNI port services and advertises.
  - Computations produce FIBs that bridge service members.

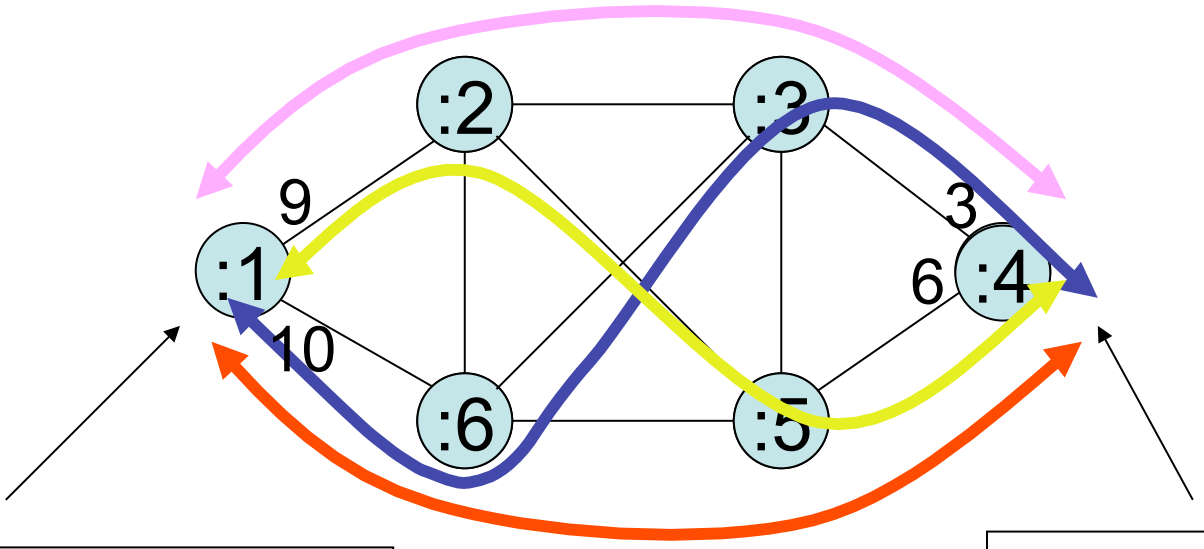
Data Path (M-in-M mode)

- C-vlan/S-vlan or untagged traffic arrives at UNI
- Its encapsulated with B-SA of bridge
- Its encapsulated with I-SID configured for group
- Its encapsulated with B-VID chosen for route
- C-DA is looked up, if found B-DA is set
- C-DA not found, B-DA is multicast that says:
  - Multicast to all other members of this I-SID group from ‘me’. Or can head-end replicate over unicast.
  - C addresses to B address association learned at UNI only.

FDB (unicast M-in-M mode)

- A unique shortest path from node to all others is computed.
- BMAC of other nodes installed in FIB pointing to appropriate out interface.
- Above is repeated for 16+ shortest paths each causes a different B-VID to be used.
- Symmetry is assured through special tie-breaking logic. 16+ different tie-breaking algorithms permit 16+ different shortest paths.

FDB visually: ucast m-in-m mode



MAC	BVID	IF
:4	1	9
:4	2	9
:4	3	10
:4	4	10

MAC	BVID	IF
:1	1	3
:1	2	6
:1	3	3
:1	4	6

FDB (mcast M-in-M mode)

If no services require tandem replication
there is no tandem FDB:

Very VPLS like .. Pretty boring….head replication over unicast paths .. Yawn..

Else (mp2mp like but without signaling)

If my node is on a unique shortest path between node **A** , which transmits for a group **I**, and node **B** which receives on group **I**, then:

merge into the FDB an entry for traffic from { **A**/Group **I** } to the interface towards **B**.

How does it work – transit multicast format (n/a for head replication)

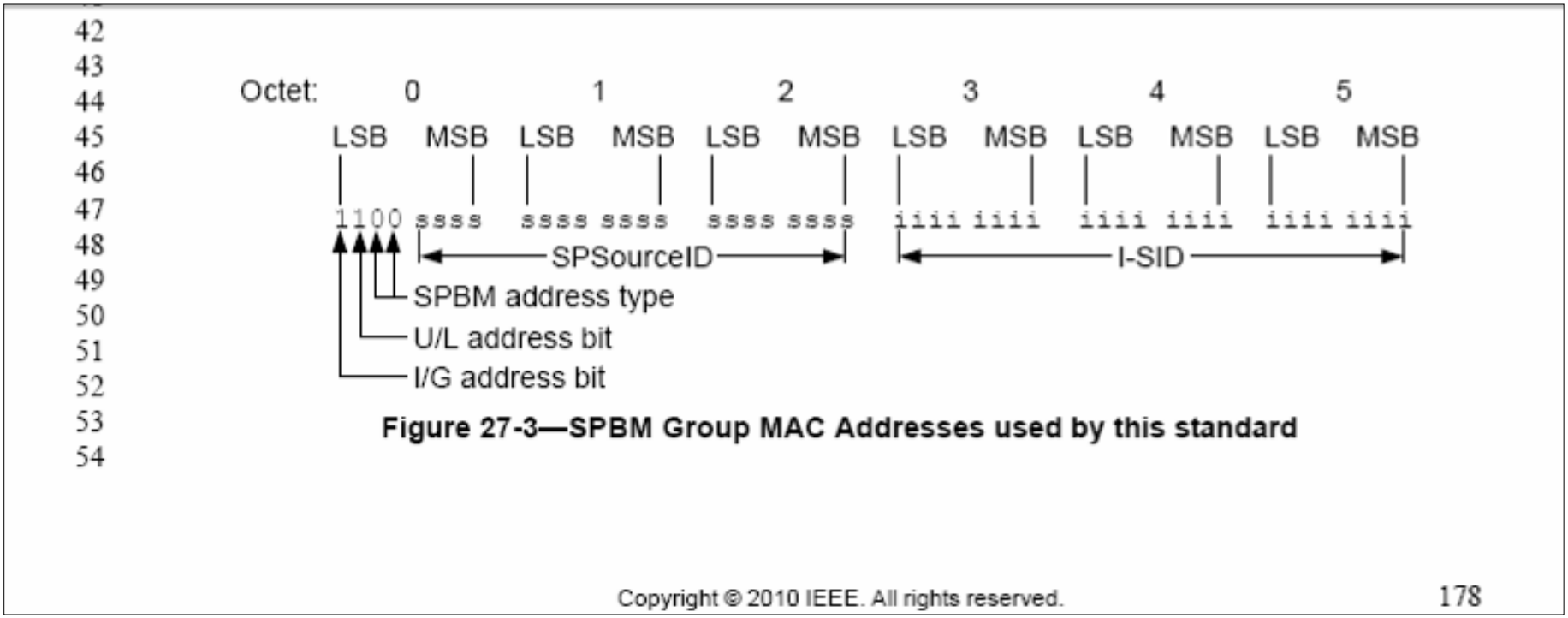


Figure 27-3—SPBM Group MAC Addresses used by this standard

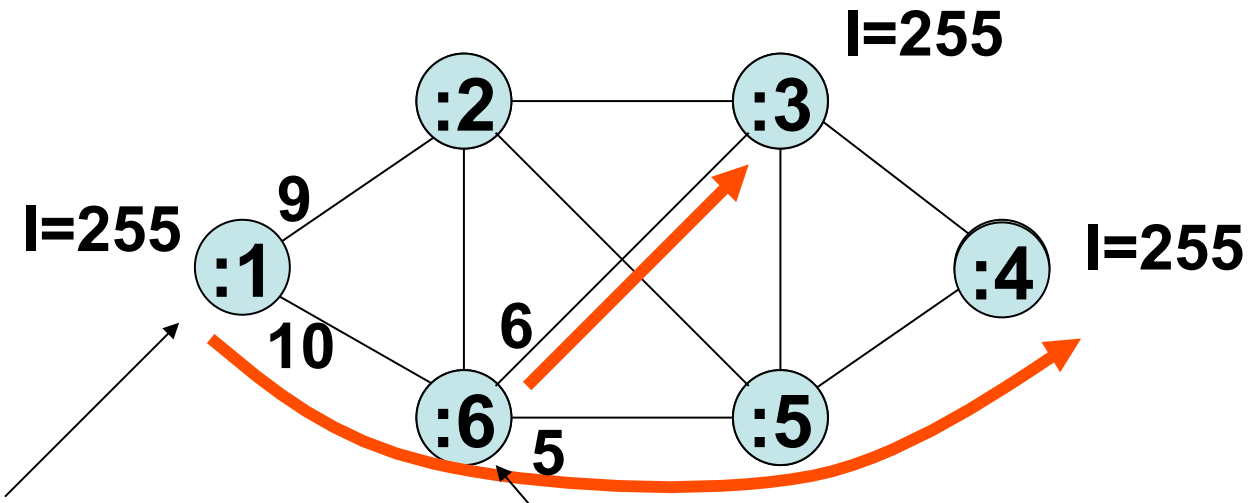
178

Example: { SOURCE: 0A-BC-DE / ISID: fe-dc-ba }

MMAC-DA: **A3-BC-DE-FE-DC-BA**

0011

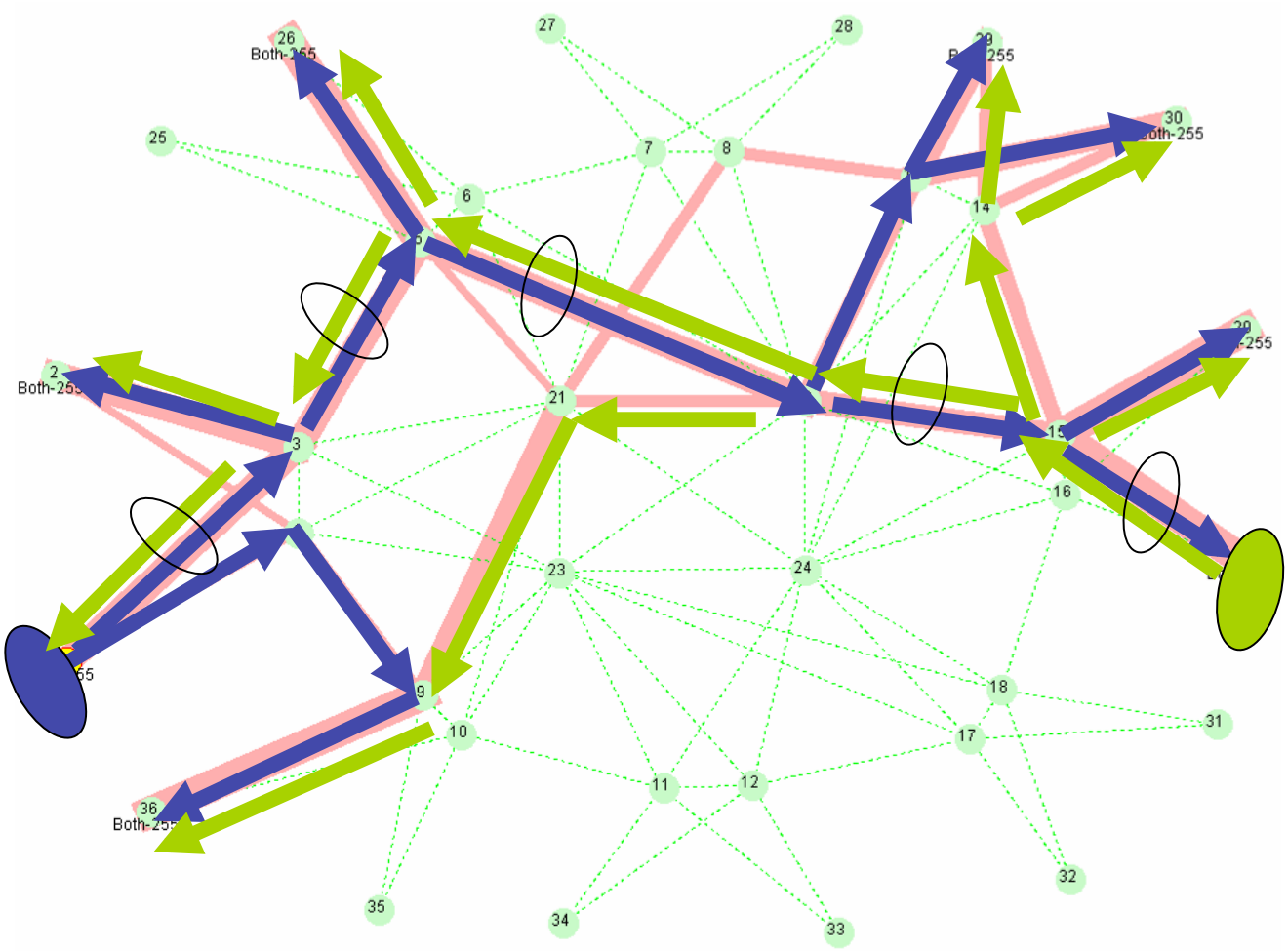
FDB visually: mcast m-in-m mode



MMAC	BVID	IF
{:1/255}	4	10

MMAC	BVID	IF
{:1/255}	4	5,6

Animation for 8 member E-LAN ‘255’



I-SID 255 has 8 members

Shown are all routes used by this I-SID in pink.

Two trees shown blue/green.

Note symmetry of trees between source/dest

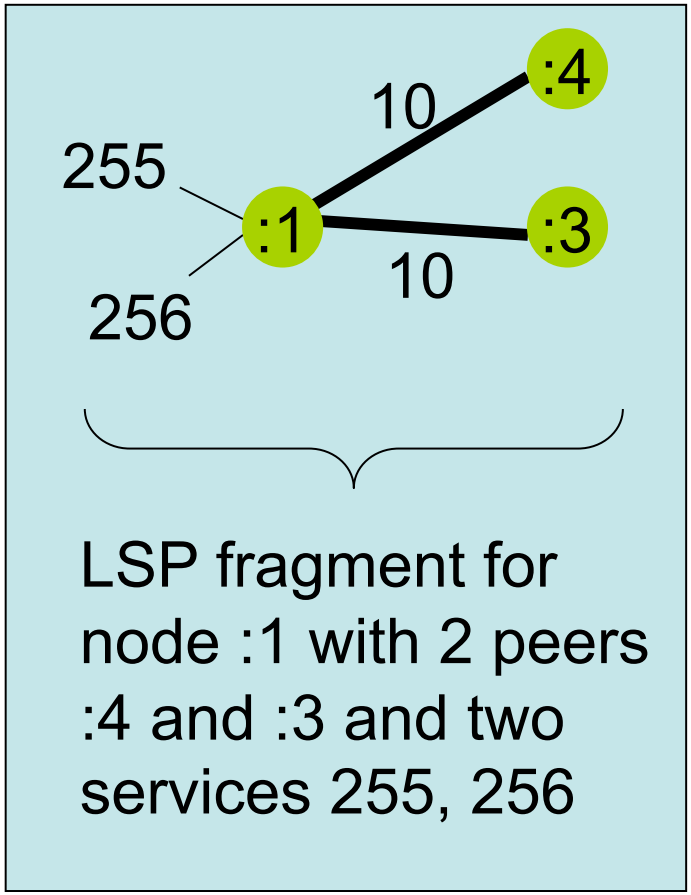
If transit multicast selected fork points in trees are replication points.

The Control Plane (m-in-m mode)

- Industry standard IS-IS Link State Protocol is basis for 802.1aq.
- Does not require any IP to operate.
- Does not preclude IPV4 or IPV6 being present in same IS-IS instance.
- SYSID carries B-MAC address
- Introduces no new PDU's to IS-IS.
- Hello TLVs augmented to pass Equal Cost Algorithm / Vid information and new NLPID.
- Update TLV's augmented to advertise SPB specific link costs.
- Update TLVs augmented to advertise ISID information.
- Update TLVs augmented to advertise nodal 'short form' name SPSOURCEID (transit mcast only).

802.1aq ISIS LSP extensions at a glance

```
LSPID                      Seq Num       Checksum ….
-------------------------------------------------

 4455.6677.0001.00-00   0x0000007a     0xc01f  …..
 SOURCE          4455.6677.0001.00  (1)
 HOST NAME       Instance_1
 NLPID           SPB (0xC1)  (2)
 AREA ADDR       22.3344                        (3)
 NBR   ID        4455.6677.0004.00   COST: 10
 NBR   ID        4455.6677.0003.00   COST: 10
 SPSOURCEID      07-00-01
 SPB   ECT-ALGORITHM 1    ECT-VID 101   (4)
 SPB   ECT-ALGORITHM 0    ECT-VID 100

 SPB       BMAC   44-55-66-77-00-01
           ECT-VID           100
           SPB   ISID           255T&R  (5)
 SPB       BMAC   44-55-66-77-00-01
           ECT-VID           101
           SPB   ISID           256T&R
```



TIE BREAKING/ECT

Symmetric Tie Breaking Requires that:

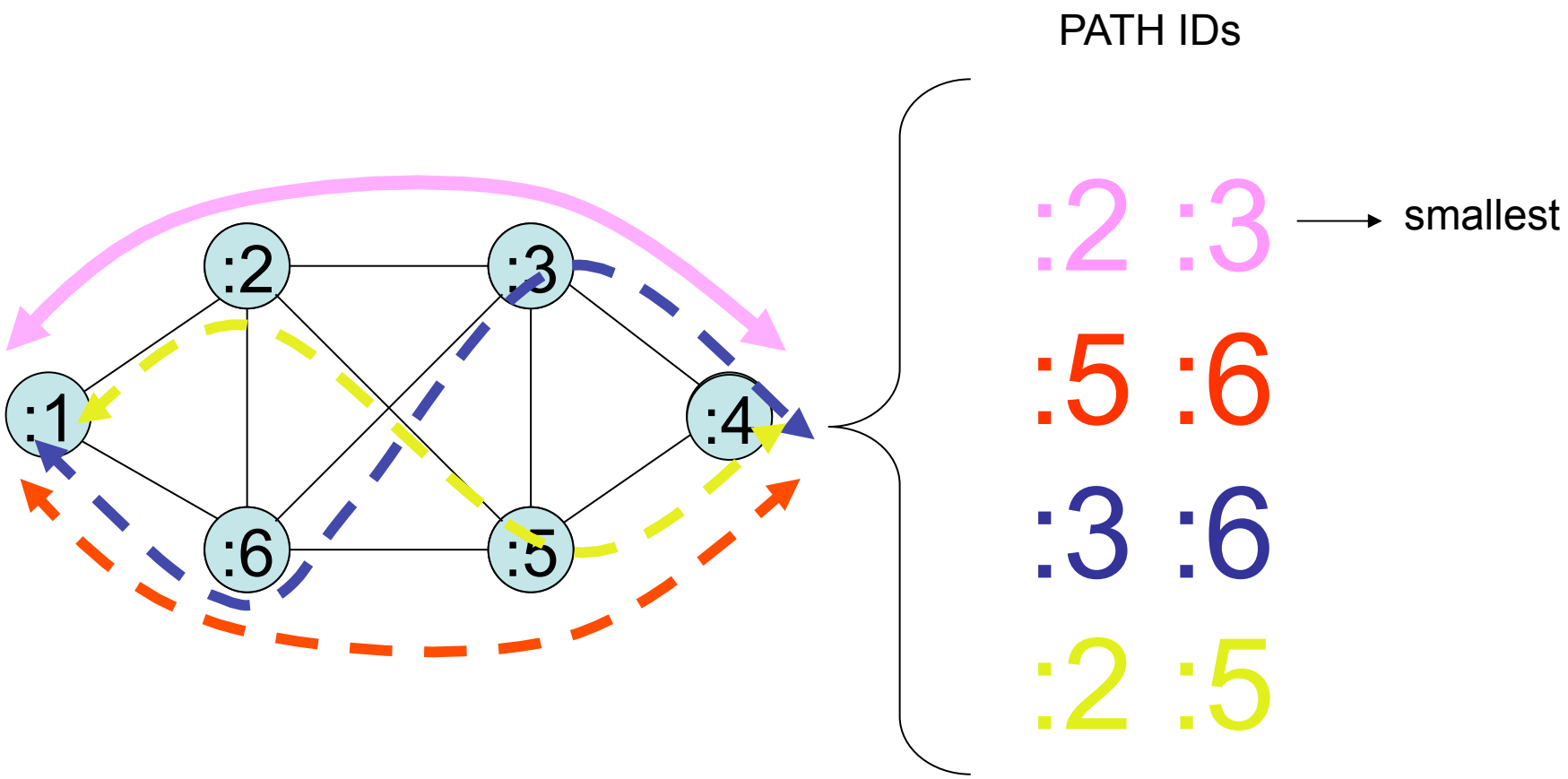
- Forward and Reverse path are the same
- Every subset of the end to end shortest path must also be a shortest path under the same tie breaking rules (transitivity): I.e.

  SPF(A,C) = SPF(A,B) || SPF(B,C) <- required for hop by hop fwding

Solution – assign to a path a ‘Path Identifier’

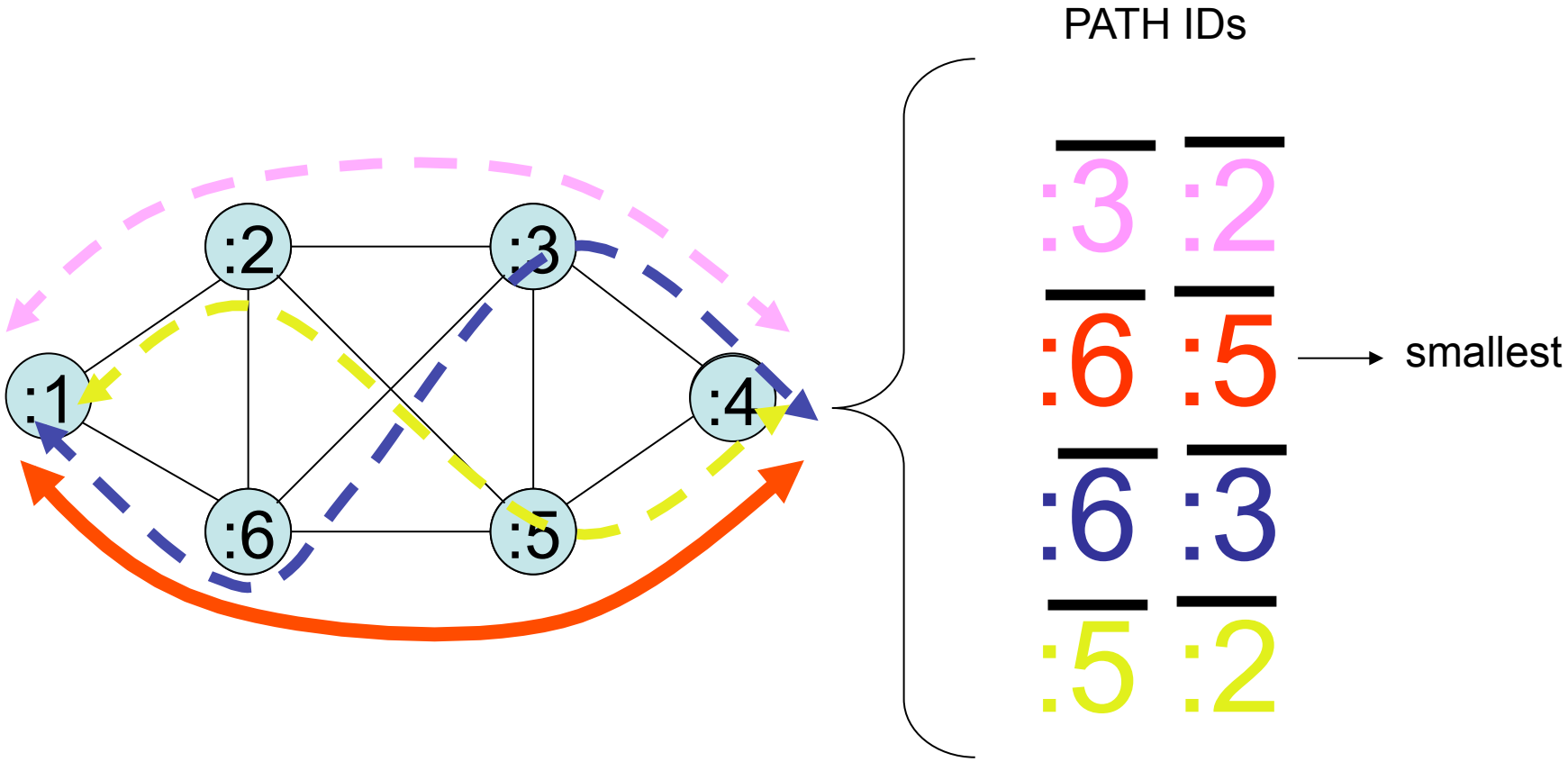
- ***Path Identifier*** := Sorted List of Nodes between SRC/DST
- Given a choice between two paths. Pick path with smallest ***Path Identifier***.

TIE BREAKING ECT – CONT'D



A path is symmetrically/deterministically chosen by picking the path with the Lowest path identifier (i.e. path with lowest Bridge Identifier on it). Transitivity proof left as exercise for the interested reader.

TIE BREAKING ECT – CONT'D



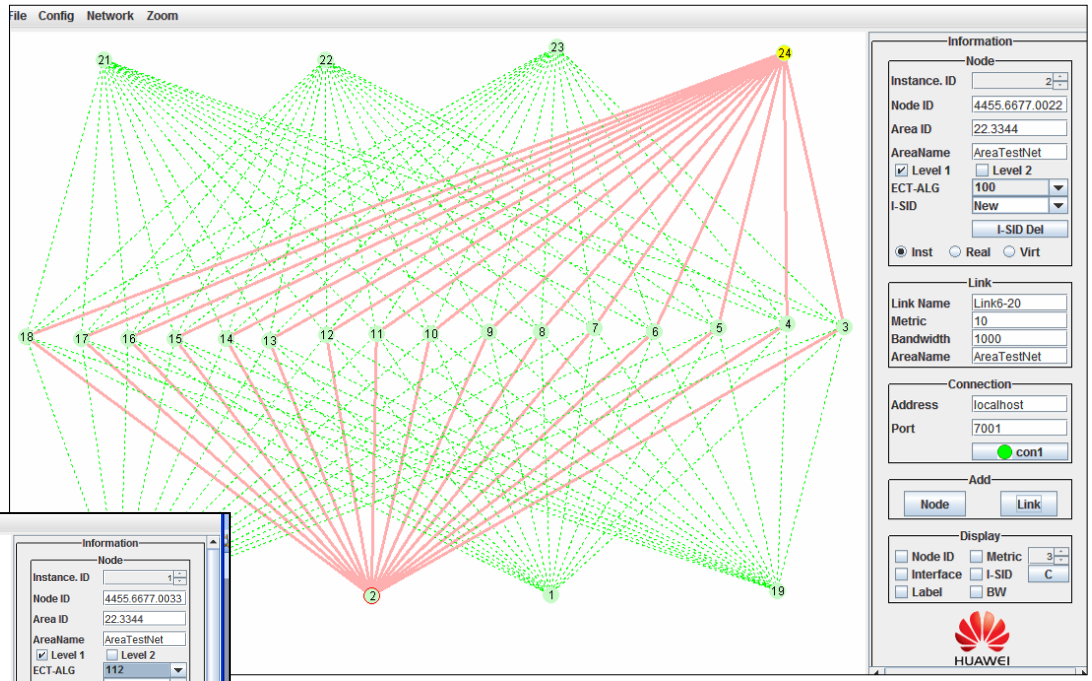
Applying a known bit mask to Bridge ID's and running the same algorithm will cause different selection. Inversion (shown) is equivalent to picking largest Path ID.

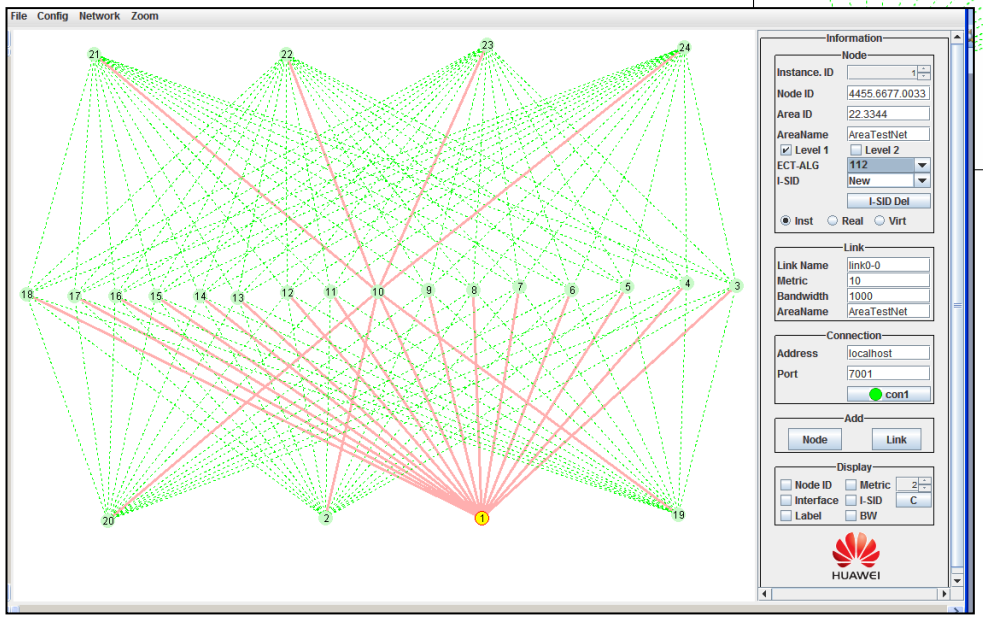
TIE BREAKING ECT – CONT'D

- In total 16 bit masks. 0x00..0x11,0x22 … 0xee,0xff are provided.
- This yields 16 possible shortest paths.
- Paths can be tuned since Bridge Priority is included in Bridge Identifier as high 2 bytes. *Bridge Identifier* = *Bridge Priority:16* || *SYSID*
- Adjusting Bridge Priority up or down allows tuning of ECT.
- Picking 16 priorities and assigning down network 'cut' yields perfect spread through cut. Eg: spine of a DC fabric.
- The 802.1aq TLV's support Opaque ECT behavior to allow further work on ECT.
- The Opaque ECT behaviors are designed to allow ECMP (hash based) as next step.

TIE BREAKING ECT – CONT'D

Can get perfect balance down spine of a two layer 16 ECT L2 Fabric. Shown Are all 16 SPF's from 2<->24





16 different SPF trees
Each use different spine as replication point.
Shown is one of the 16 SPF's from/to node 1.

Loop Suppression & Avoidance

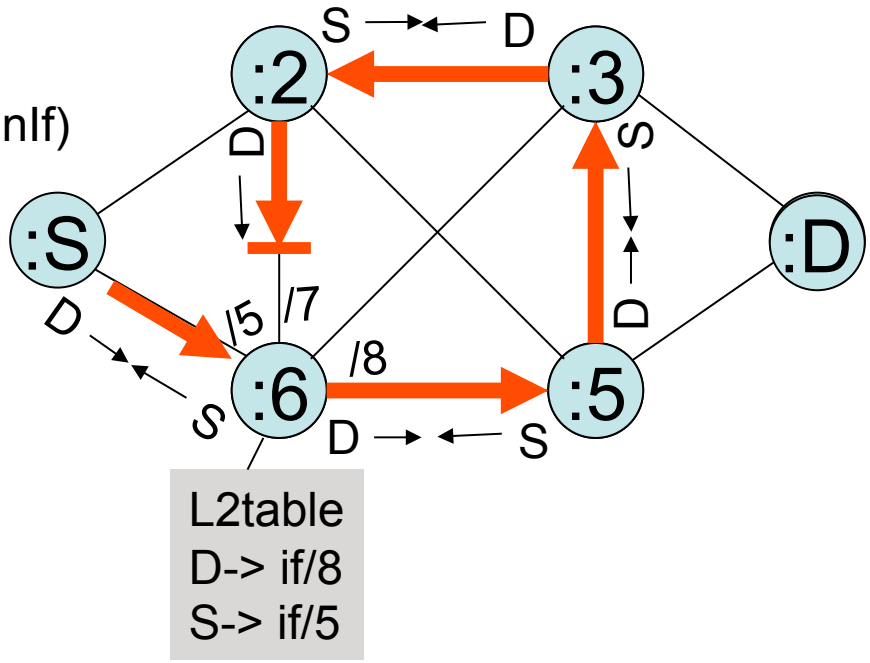
Suppression

- done on the data path using an SA check.
- prevents 99.99% loops if FDB's create one.
- no impact on convergence rates.
- exploits symmetric/congruence properties of routing.
- uses reverse learning options of most h/w to discard.

```
Forward(pkt, inIf) {
    if (l2table[pkt.sa,pkt.vid].outIf != inIf)
        discard (pkt);
    else …
}
```

On :6

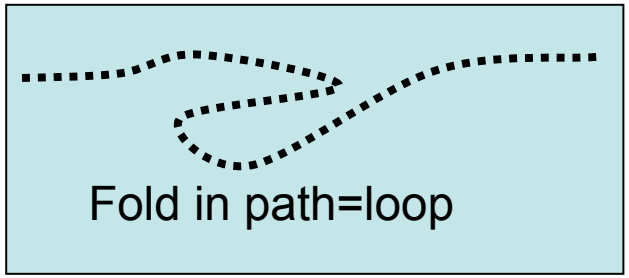
/7 != /5 so discard

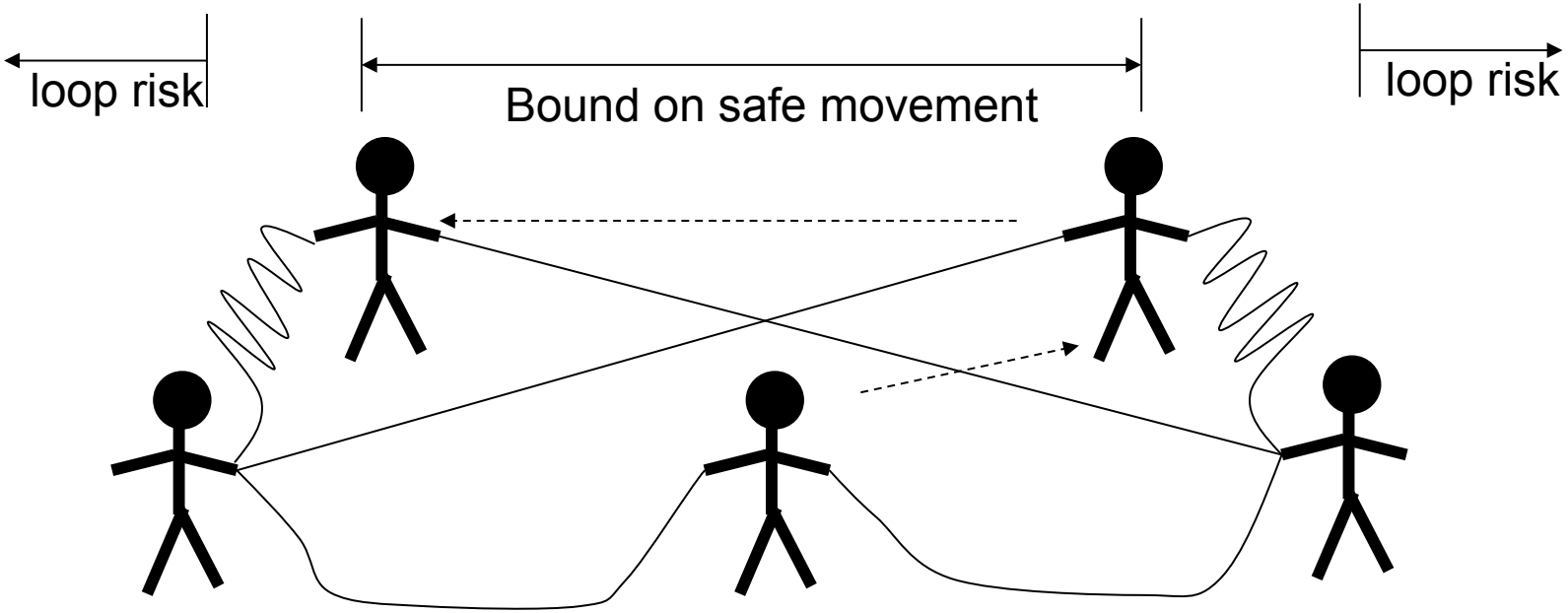


Loop Suppression & Avoidance

Avoidance

- hellos augmented with topology ‘digests’
- mismatched topology => some forwarding entries unsafe.
- blocks ‘unsafe’ entries.
- Unsafe => distance changed outside “safe zone”





802.1aq OA&M (inherited *by design*)

Service/Network Layer – 802.1ag Connectivity Fault mgmt

- Hierarchy (honors maintenance levels/abstraction)
  - Continuity Check
  - L2 traceroute
  - L2 ping

Link Layer – 802.3ah

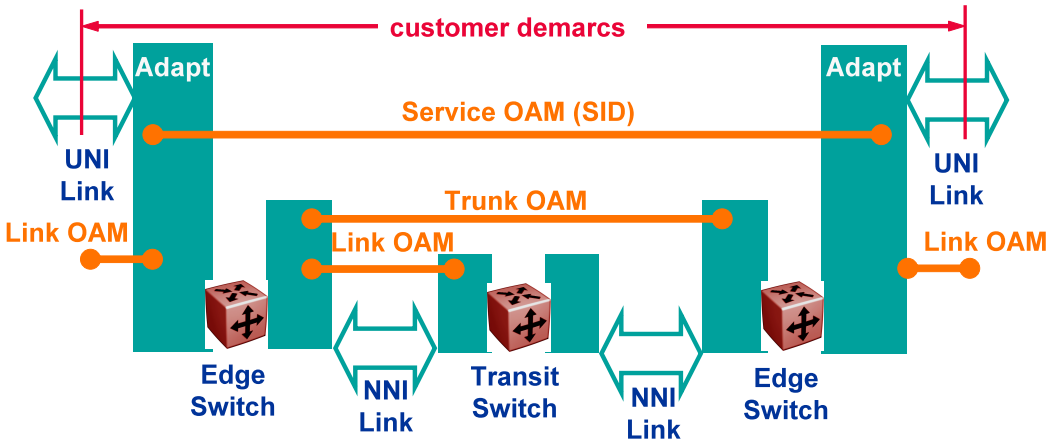
- Link Monitoring (logical/physical)
- Remote Failure Indication
- Remote Loopback

Service Layer - Y.1731

- Multicast Loopback – depends on congruency/symmetry
- Performance Measurements (Loss/Delay etc.)
- One way/two way delays – symmetry important

802.1aq OAM capabilities

1. **Continuity Check (CC)**
   a) Multicast/unidirectional heartbeat
   b) Usage: Fault detection
2. **Loopback – Connectivity Check**
   a) Unicast bi-directional request/response
   b) Usage: Fault verification
3. **Traceroute (i.e., Link trace)**
   a) Trace nodes in path to a specified target node
   b) Usage: Fault Isolation



4. **Discovery** (not specifically supported by .1ag however Y.1731 and 802.1ab support it)
   a) Service (e.g. discover all nodes supporting common service instance)
   b) Network (e.g. discover all devices common to a domain)
5. **Performance Monitoring** (MEF10 and 12 - Y.1731 for pt-pt now extending to pt-mpt and mpt-mpt)
   a) Frame Delay, Frame Loss, Frame Delay Variation (derived)
   b) Usage: Capacity planning, SLA reporting

MIBs (also all OA&M etc. inherited)

SpbSys	Basic config data
SpbSysDynamic	Basic run time data
SpbEctStatic	Desired Equal cost behaviors
SpbEctDynamic	Actual Equal cost behaviors
SpbAdjStatic	Desired Adjacency data (metric etc.)
SpbAdjDynamic	Actual Adjacencies
SpbTopNode	All other nodes in the SPB region and basic data
SpbTopEct	All other ECT behavior in the SPB region
SpbTopEdge	All other Edge data in the SPB region
SpbTopSrv	All services in the SPB region, who hosts etc.

Rapid Recovery

- ISIS augmented with multicast LSP flood to all 802.1aq nodes.
- Every 802.1aq node joins default service ISID 0xffffff.
- This E-LAN is just for control plane.
- LSPs can be advertised over this E-LAN
- Very fast distribution protocol (h/w multicast).
- On failure each end of link advertises over this ‘default’ E-LAN (in addition to normal updates).
- Reaches all 802.1aq participants at h/w multicast speed with no CPU involvement transit.
- Conceptually like having a shared LAN joining all nodes with a physical port but no DR election etc. is done, only used as unreliable very fast distribution mechanism backed up by normal IS-IS hop by hop LSPs.

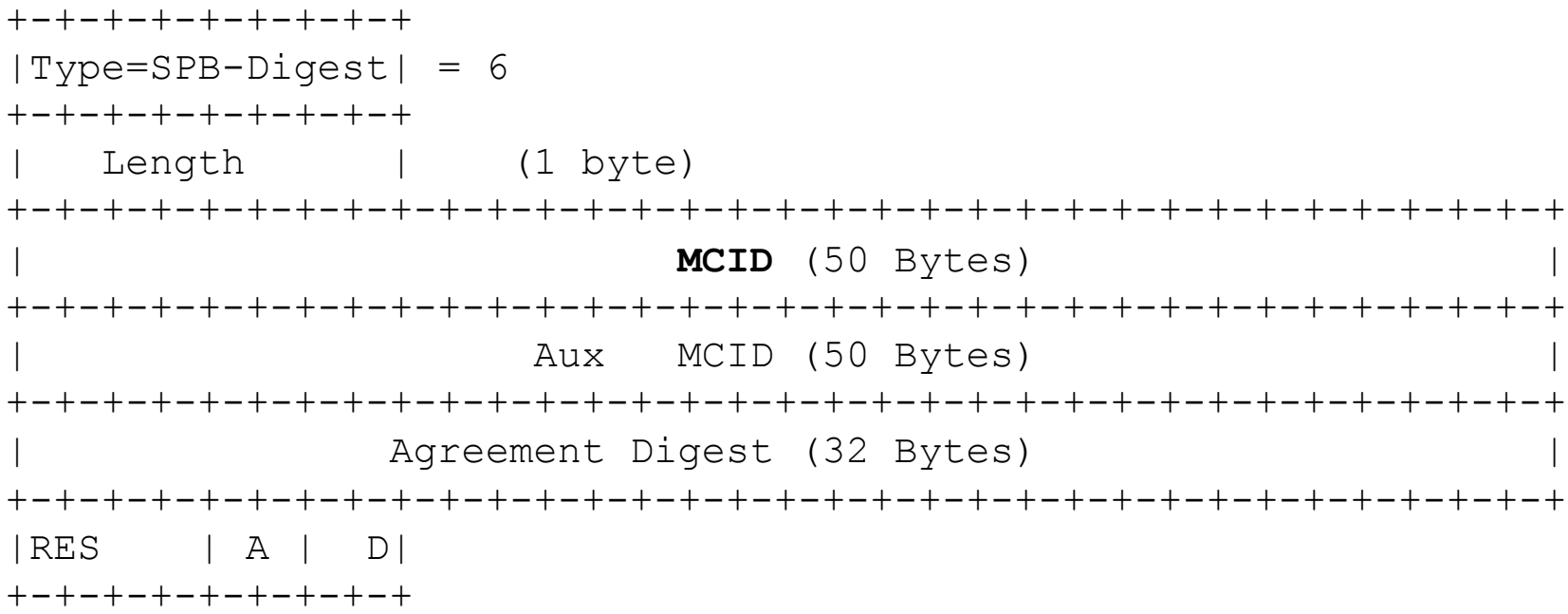
TLV Summary

PDU	TLV	SUB-TLV	TYPE	TYPE	#OCCURRENCE
IIH					
	MT-Port-Capability		143		
		SPB-B-VID		5	1
		SPB-Digest		6	1
LSP					
	MT-Capability		144		
		SPB-Inst		1	1
		SPB-I-OALG		2	>=0
		SPBM-SI		3	>=0
		SPBV-ADDR		4	>=0
	MT-Intermediate-System		222		
		SPB-Metric		12	1
		SPB-A-OALG		13	>=0

No new PDU, no changes to IS-IS state machine logic

TLV – SPB-Digest (in Hello)

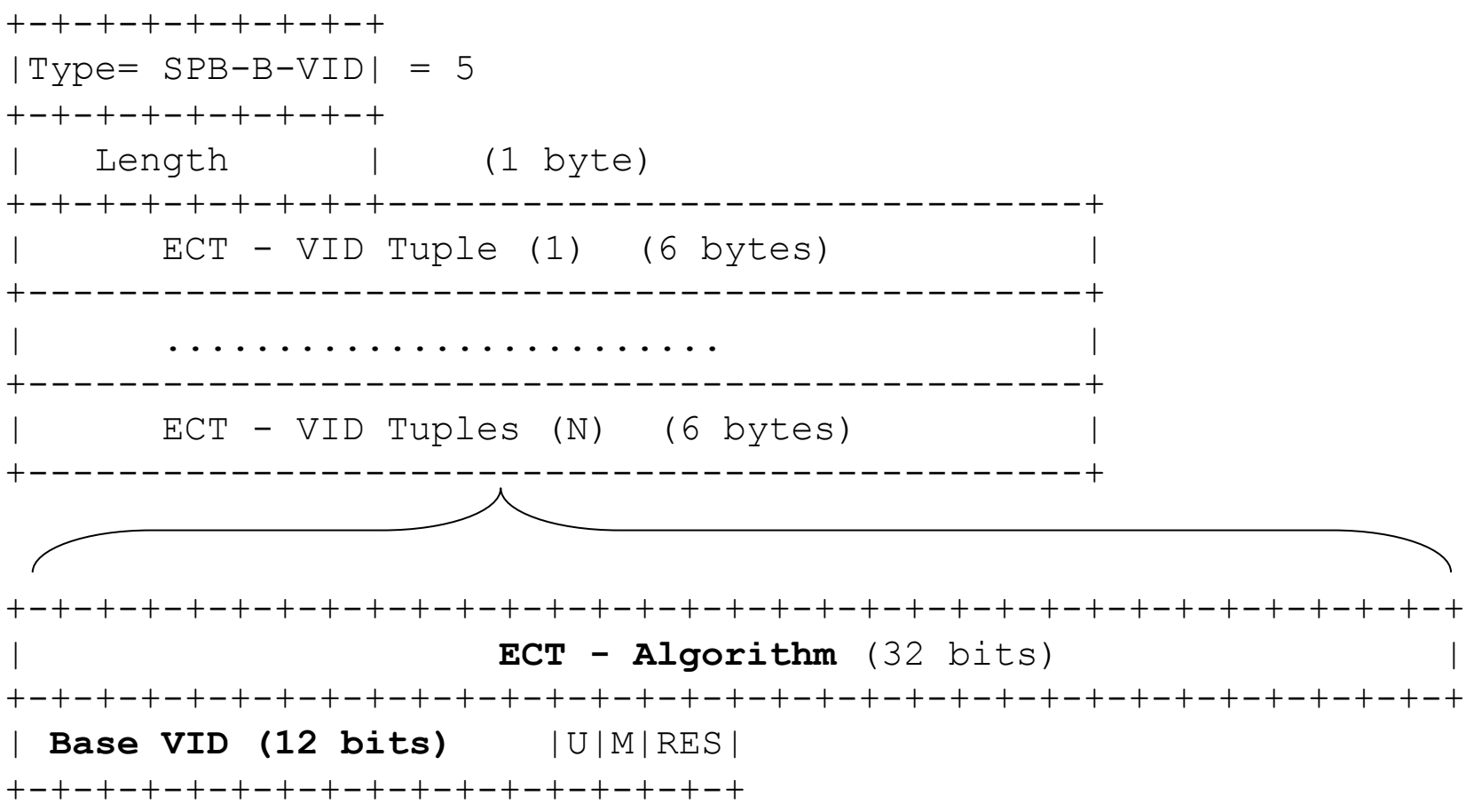
All TLV's in context of an MTID TLV



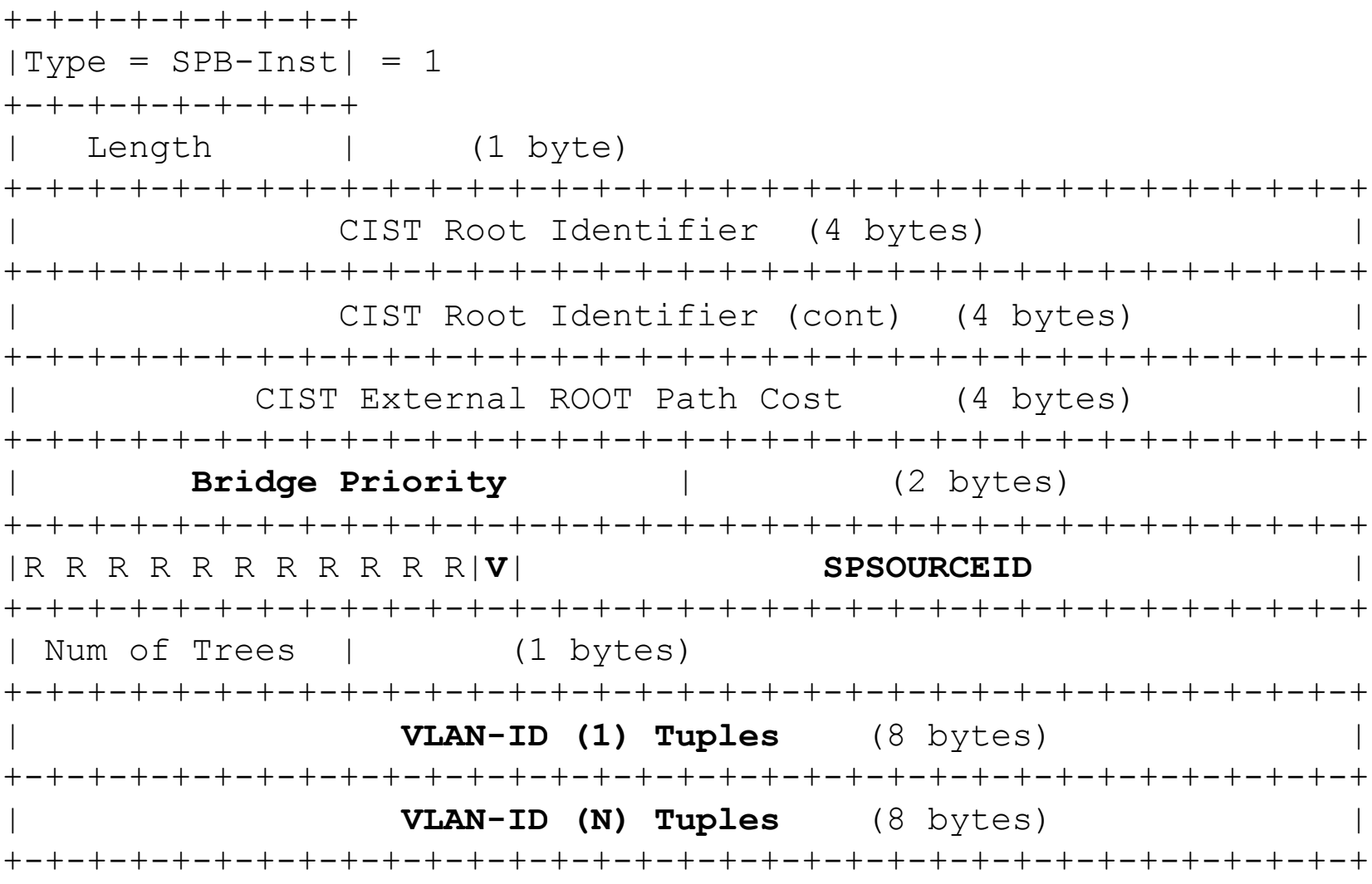
ALSO in Hello the **NLPID value 0xC1** is used and advertised.

HELLO padding unchanged ..

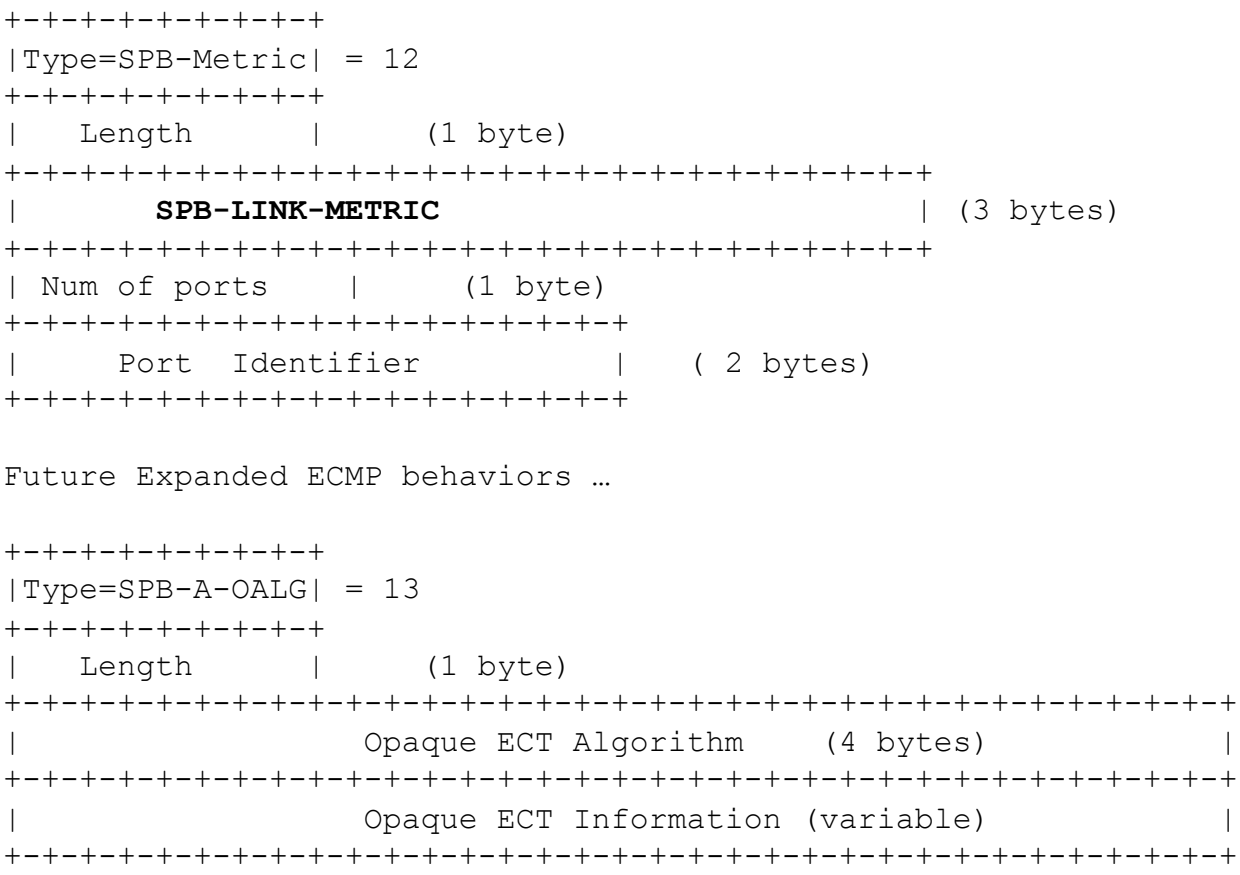
TLV – SPB-B-VID (in Hello)



TLV – SPB-Instance

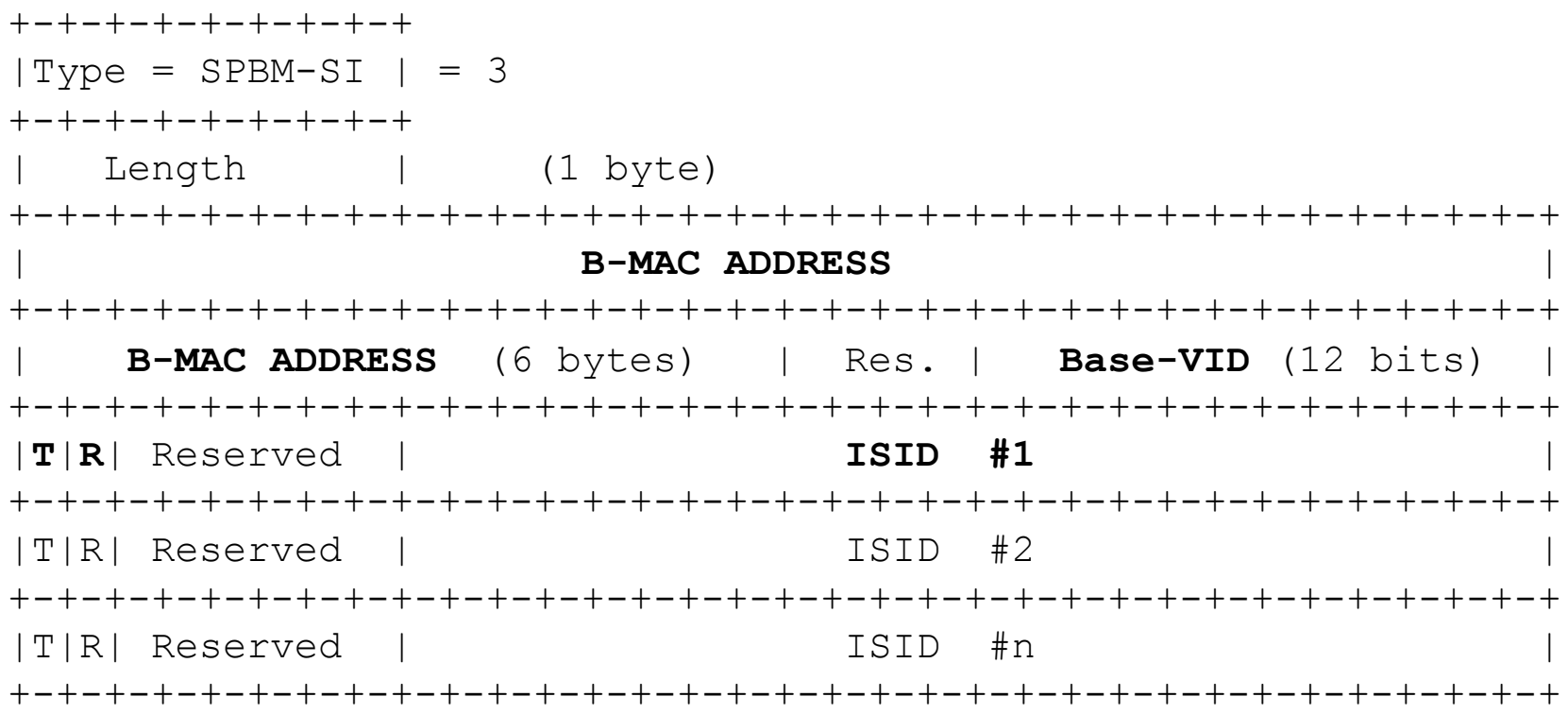


TLV – SPB-Metric



Future Expanded ECMP behaviors …

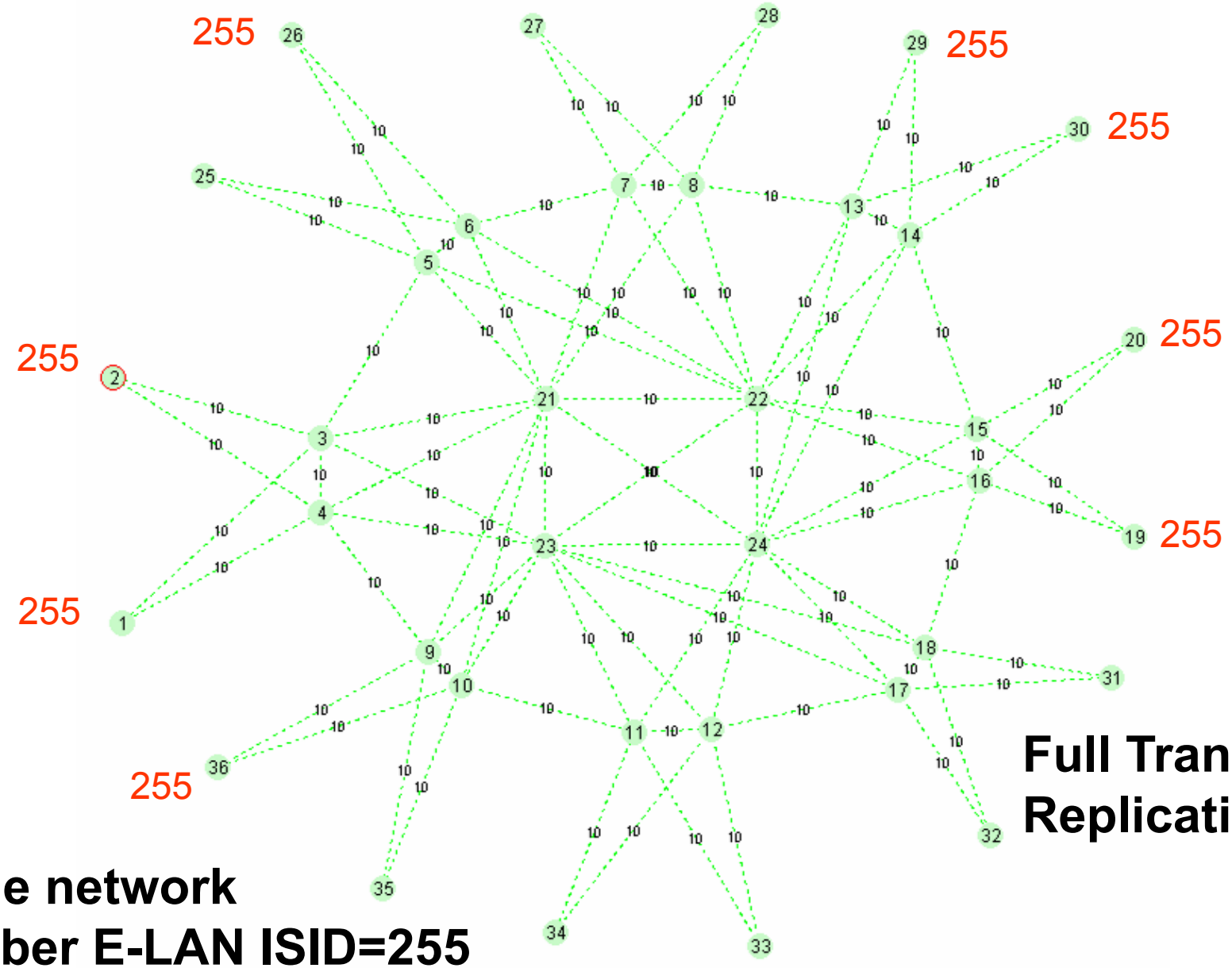
TLV – SPB-Service Instance



Outline

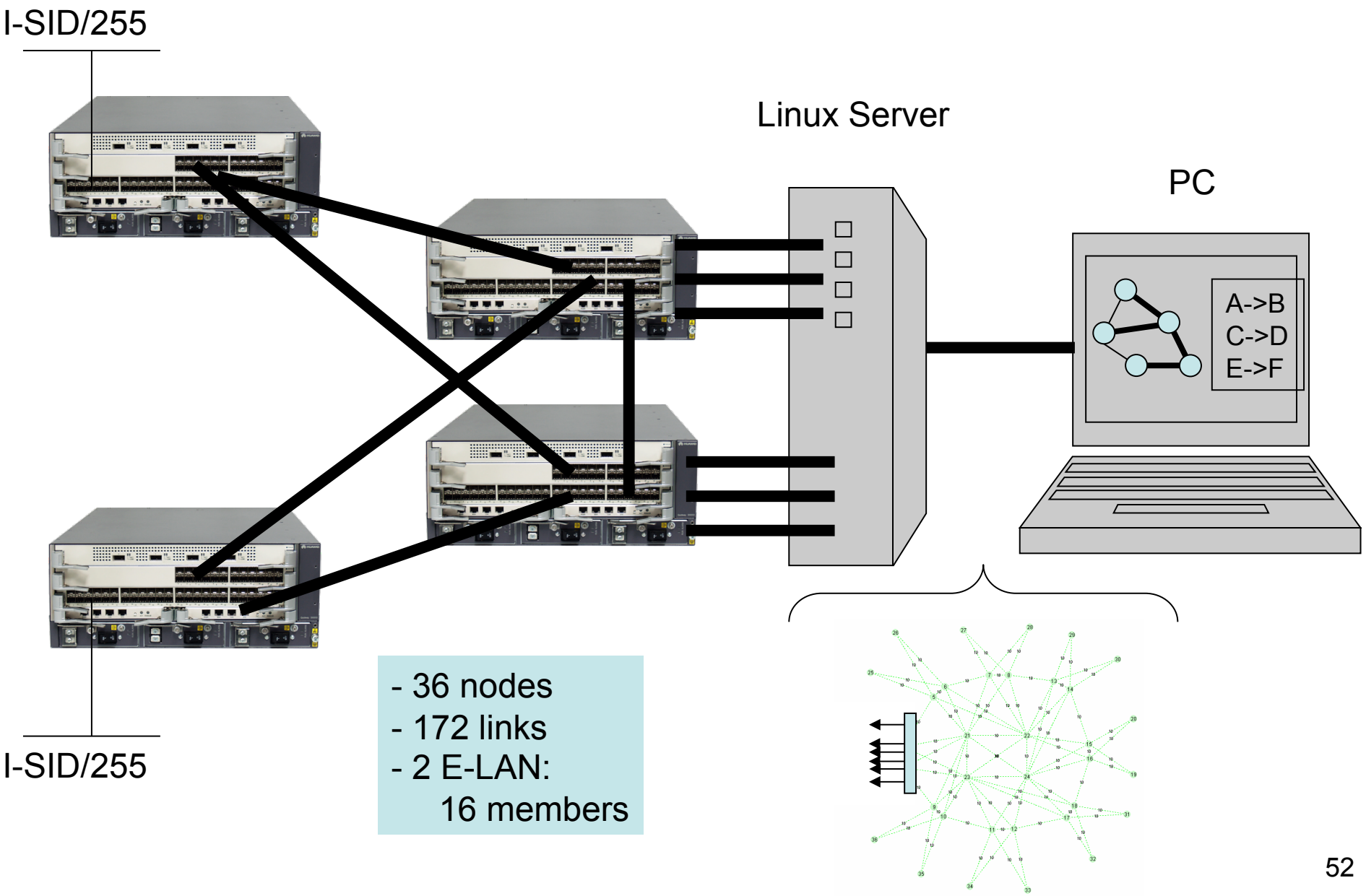
- Challenges
- What is 802.1aq/SPB
- Applications
- How does it work
- **Short Demo (remote switch if possible)**

EXAMPLE NETWORK :



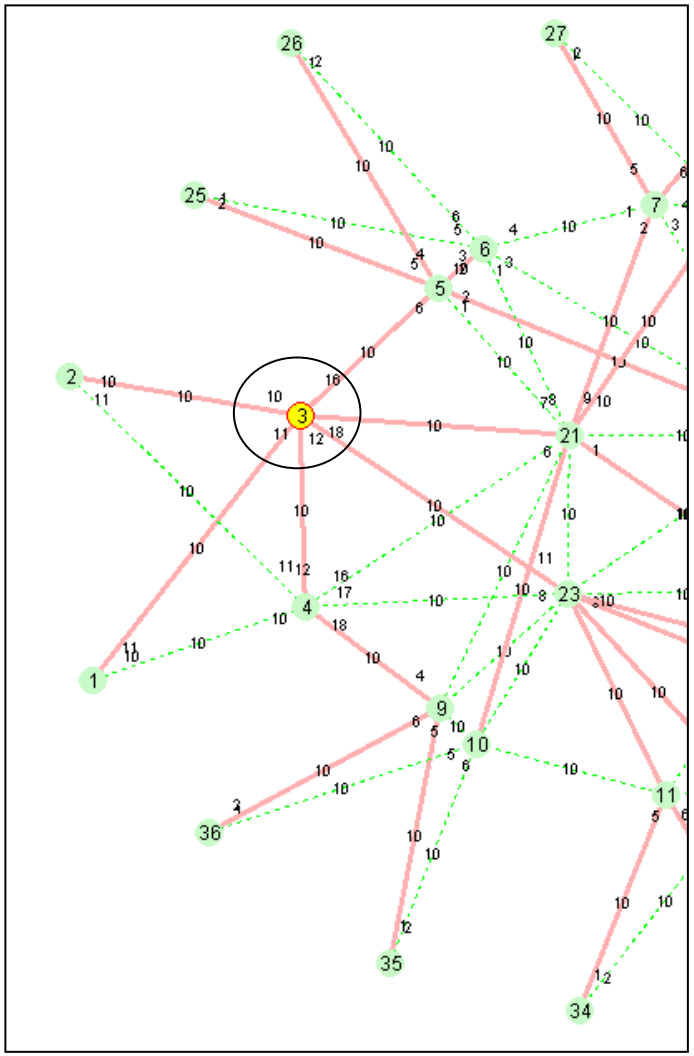
36 node network
8 member E-LAN ISID=255

DEMO BASED ON OUR S9300 SERIES TERABIT SWITCHES
S9303/1T, S9306/2T, S9312/4T



EXAMPLE – ISIS PEERS AT NODE :3

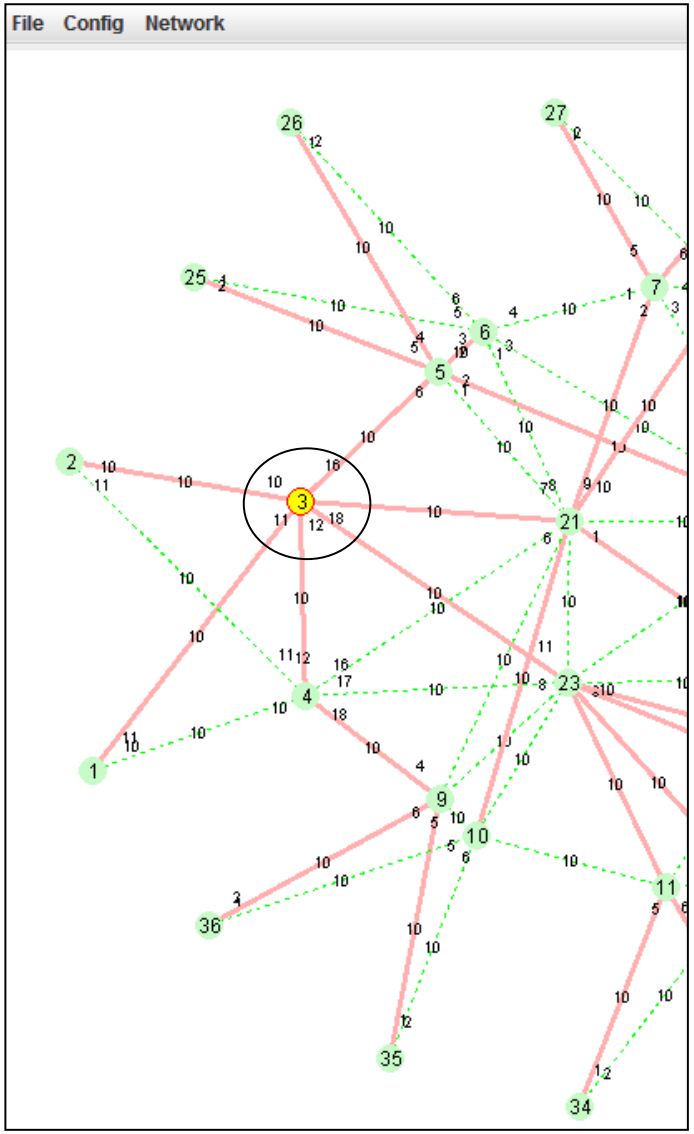
```
<ottawa-9300-3>d spb
 The current global spb information is :
 Device HMAC is 44-55-66-77-00-03
 Spsid is 07-00-03
 Ect vlan amount is 2
 Ect vlan sequence number [1] is: vlan 100 !
 Ect vlan sequence number [2] is: vlan 101 !
<ottawa-9300-3>

<ottawa-9300-3>d isis peer

                        Peer information for ISIS(1)
                        ----------------------------
  System Id     Interface          Circuit Id       State HoldTime Type
4455.6677.0001  Vlanif211          0000000002        Up    26s      L1
4455.6677.0004  Vlanif212          0000000003        Up    23s      L1
4455.6677.0005  Vlanif216          004               Up    27s      L1
4455.6677.0015  Vlanif217          005               Up    27s      L1
4455.6677.0017  Vlanif218          006               Up    25s      L1
Total Peer(s): 5

<ottawa-9300-3>
```



Logging on to node :3
We can see the basic SPB info and the ISIS peers….

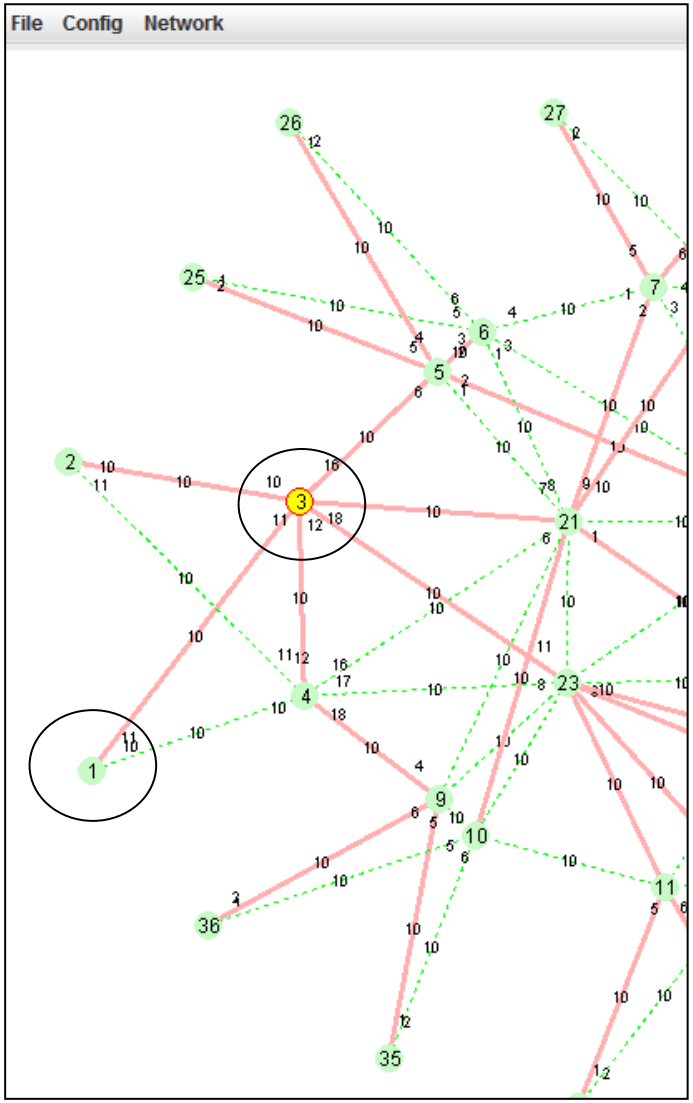
EXAMPLE – LSDB at node :3

```
                Database information for ISIS(1)
                --------------------------------
                   Level-1 Link State Database
LSPID                 Seq Num      Checksum      Holdtime      Length  ATT/P/OL
-------------------------------------------------------------------------------
4455.6677.0001.00-00  0x00000fd2   0x1cea        1044          236     0/0/0
4455.6677.0003.00-00* 0x00001448   0x3d27        683           323     0/0/0
4455.6677.0004.00-00  0x00000ff8   0xd1d9        1090          323     0/0/0
4455.6677.0005.00-00  0x00000b3b   0x9ba7        586           317     0/0/0
4455.6677.0006.00-00  0x00000b3b   0xbc31        819           293     0/0/0
4455.6677.0007.00-00  0x00000b3b   0xce10        1075          293     0/0/0
4455.6677.0008.00-00  0x00000b3e   0xebe0        288           293     0/0/0
4455.6677.0009.00-00  0x00000b3b   0x8b77        355           317     0/0/0
4455.6677.000a.00-00  0x00000b3b   0x57a         840           294     0/0/0
4455.6677.000b.00-00  0x00000b3a   0x1665        608           294     0/0/0
4455.6677.000c.00-00  0x00000b3a   0x6903        764           294     0/0/0
4455.6677.000d.00-00  0x00000b3a   0x89fd        431           294     0/0/0
4455.6677.000e.00-00  0x00000b39   0x3445        611           294     0/0/0
4455.6677.000f.00-00  0x00000b3a   0xdabe        616           294     0/0/0
4455.6677.0010.00-00  0x00000b3b   0x810e        452           294     0/0/0
4455.6677.0011.00-00  0x00000b3a   0x1b46        645           294     0/0/0
4455.6677.0012.00-00  0x00000b39   0x1b3e        447           294     0/0/0
4455.6677.0013.00-00  0x00000b3a   0x943a        419           176     0/0/0
4455.6677.0014.00-00  0x00000b3b   0xdff2        693           176     0/0/0
4455.6677.0015.00-00  0x00000b41   0xdade        1141          508     0/0/0
4455.6677.0016.00-00  0x00000b3e   0xa832        1011          464     0/0/0
4455.6677.0017.00-00  0x00000b40   0x1563        640           508     0/0/0
4455.6677.0018.00-00  0x00000b3a   0xadee        417           464     0/0/0
4455.6677.0019.00-00  0x00000b3a   0xcff         291           158     0/0/0
4455.6677.001a.00-00  0x00000b3b   0xb131        794           176     0/0/0
4455.6677.001b.00-00  0x00000b3b   0x7062        822           176     0/0/0
4455.6677.001c.00-00  0x00000b3b   0x5876        463           176     0/0/0
4455.6677.001d.00-00  0x00000b3b   0xa610        460           176     0/0/0
4455.6677.001e.00-00  0x00000b3a   0xf5c6        627           176     0/0/0
4455.6677.001f.00-00  0x00000b3b   0x8a19        825           176     0/0/0
4455.6677.0020.00-00  0x00000b3a   0x960a        584           176     0/0/0
4455.6677.0021.00-00  0x00000b3b   0x5b58        1033          176     0/0/0
4455.6677.0022.00-00  0x00000b3a   0x8927        693           176     0/0/0
4455.6677.0023.00-00  0x00000b3b   0x9352        664           158     0/0/0
4455.6677.0024.00-00  0x00000b39   0xc8ec        736           176     0/0/0

    *(In TLV)-Leaking Route, *(By LSPID)-Self LSP, +-Self LSP(Extended),
          ATT-Attached, P-Partition, OL-Overload
```



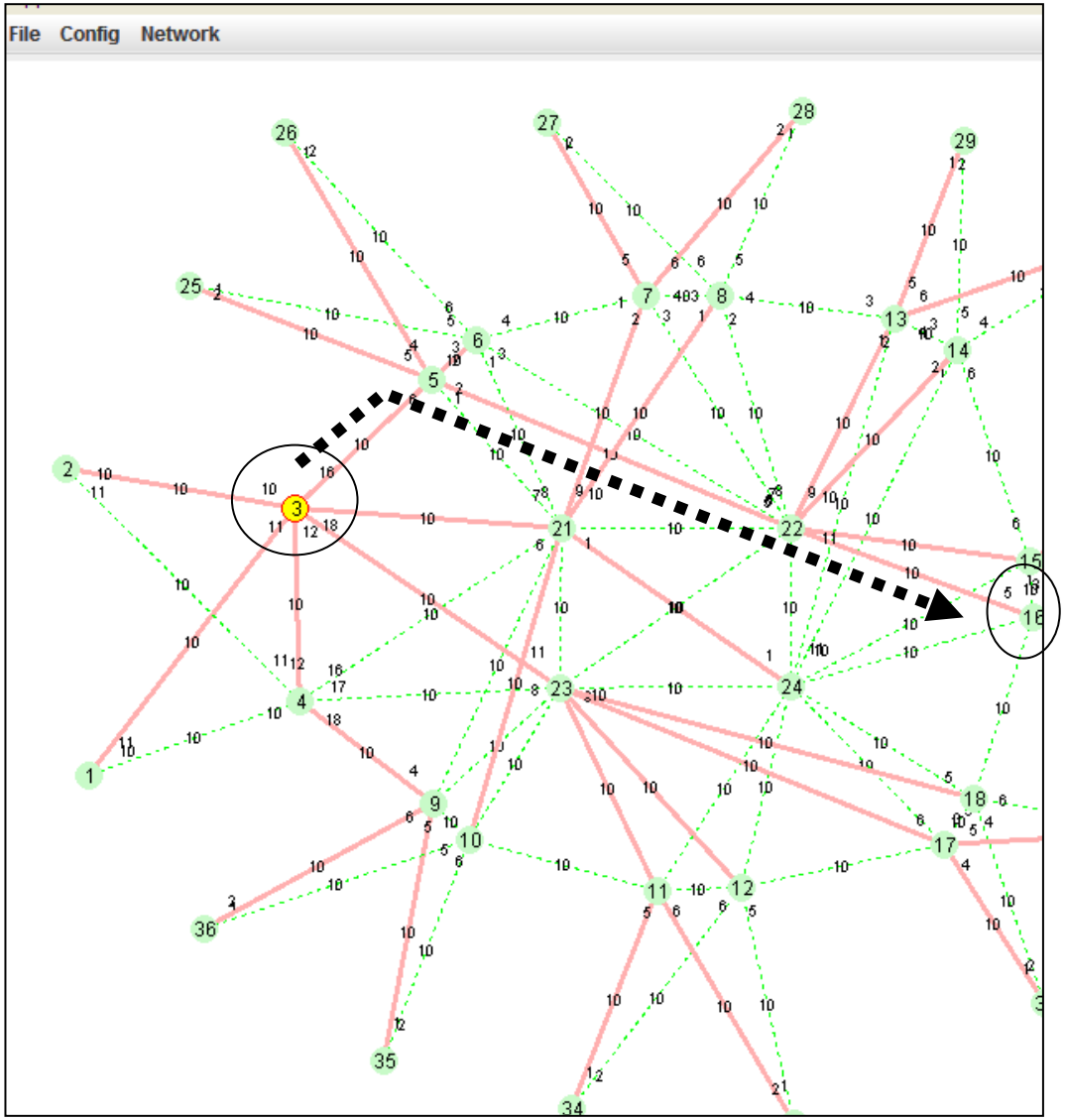
EXAMPLE LSP VERBOSE OF NODE :1 at NODE :3

```
<ottawa-9300-3>d isis lsdb 4455.6677.0001.00-00 verbose

                        Database information for ISIS(1)
                        --------------------------------

                          Level-1 Link State Database

LSPID                 Seq Num      Checksum      Holdtime
--------------------------------------------------------------------

4455.6677.0001.00-00  0x00000fd3   0x1aeb        1194
 SOURCE       4455.6677.0001.00
 NLPID        SPB(0xC1)
 AREA ADDR    22.3344
+NBR  ID      4455.6677.0003.00  COST: 10
+NBR  ID      4455.6677.0004.00  COST: 10
 SPB  ECT-ALGORITHM 0    ECT-VID 100
 SPB  ECT-ALGORITHM 1    ECT-VID 101
 SPB  ECT-ALGORITHM 2    ECT-VID 0
 …….
 SPB  ECT-ALGORITHM 15   ECT-VID 0
 SPSID           07-00-01
 SPB     BMAC  44-55-66-77-00-01
         ECT-VID         100
         SPB ISID        255T&R

<ottawa-9300-3>
```



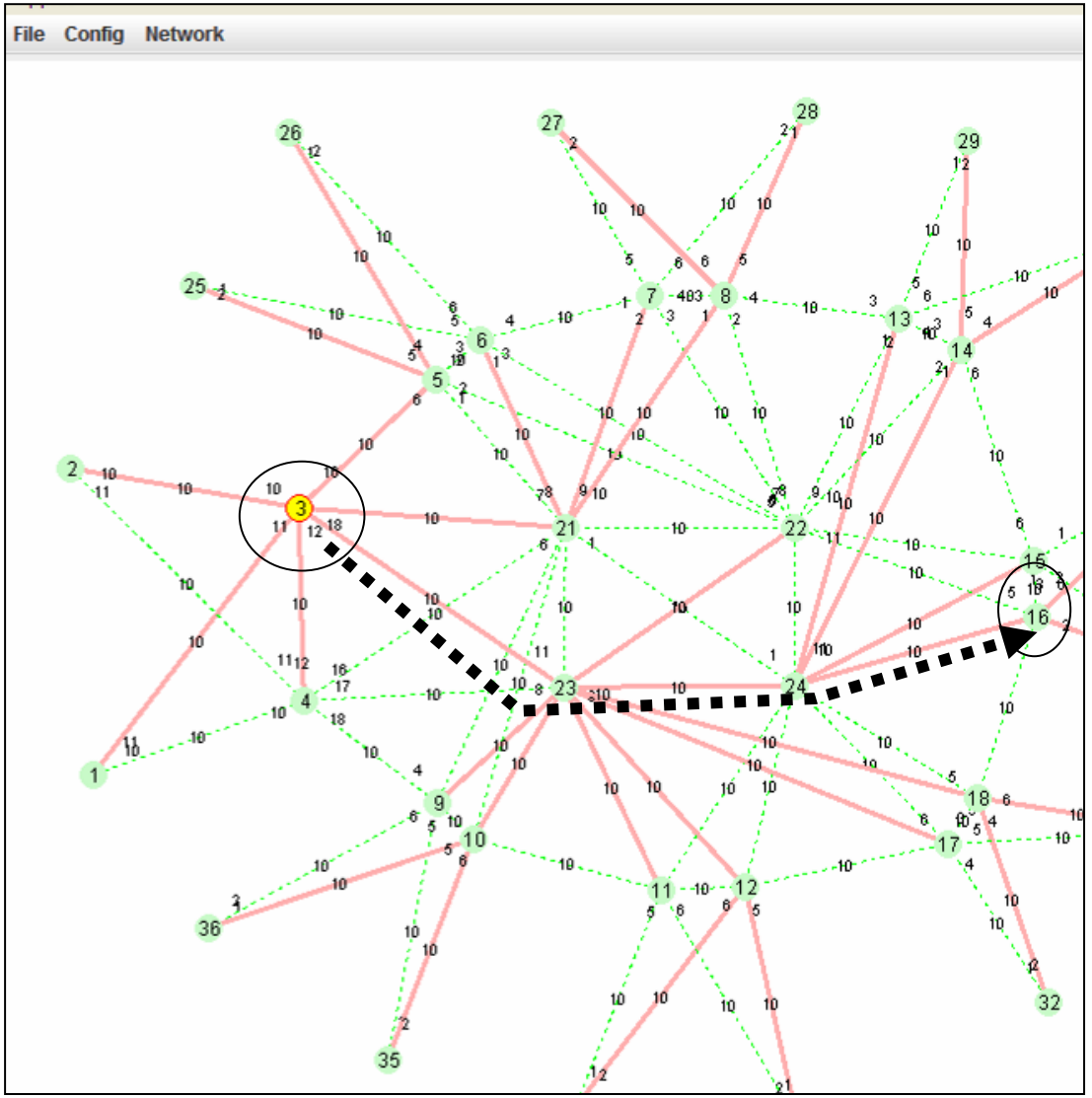
EXAMPLE – NODE :3 ROUTE TO :10 (first equal cost path)

```
<ottawa-9300-3>d spb umac
-----------------------------------------------
    BMAC              BVLAN          IF NAME
-----------------------------------------------
4455-6677-0001      100            GE2/0/11
4455-6677-0001      101            GE2/0/11
4455-6677-0004      100            GE2/0/12
4455-6677-0004      101            GE2/0/12
4455-6677-0005      100            GE2/0/16
4455-6677-0005      101            GE2/0/16
4455-6677-0006      100            GE2/0/16
4455-6677-0006      101            GE2/0/17
4455-6677-0007      100            GE2/0/17
4455-6677-0007      101            GE2/0/17
4455-6677-0008      100            GE2/0/17
4455-6677-0008      101            GE2/0/17
4455-6677-0009      100            GE2/0/12
4455-6677-0009      101            GE2/0/18
4455-6677-000a      100            GE2/0/17
4455-6677-000a      101            GE2/0/18
4455-6677-000b      100            GE2/0/18
4455-6677-000b      101            GE2/0/18
4455-6677-000c      100            GE2/0/18
4455-6677-000c      101            GE2/0/18
4455-6677-000d      100            GE2/0/16
4455-6677-000d      101            GE2/0/18
4455-6677-000e      100            GE2/0/16
4455-6677-000e      101            GE2/0/18
4455-6677-000f      100            GE2/0/16
4455-6677-000f      101            GE2/0/18
4455-6677-0010      100            GE2/0/16
4455-6677-0010      101            GE2/0/18
...
4455-6677-0024      100            GE2/0/12
4455-6677-0024      101            GE2/0/18
 Total unicast fib entries is 68
<ottawa-9300-3>
```

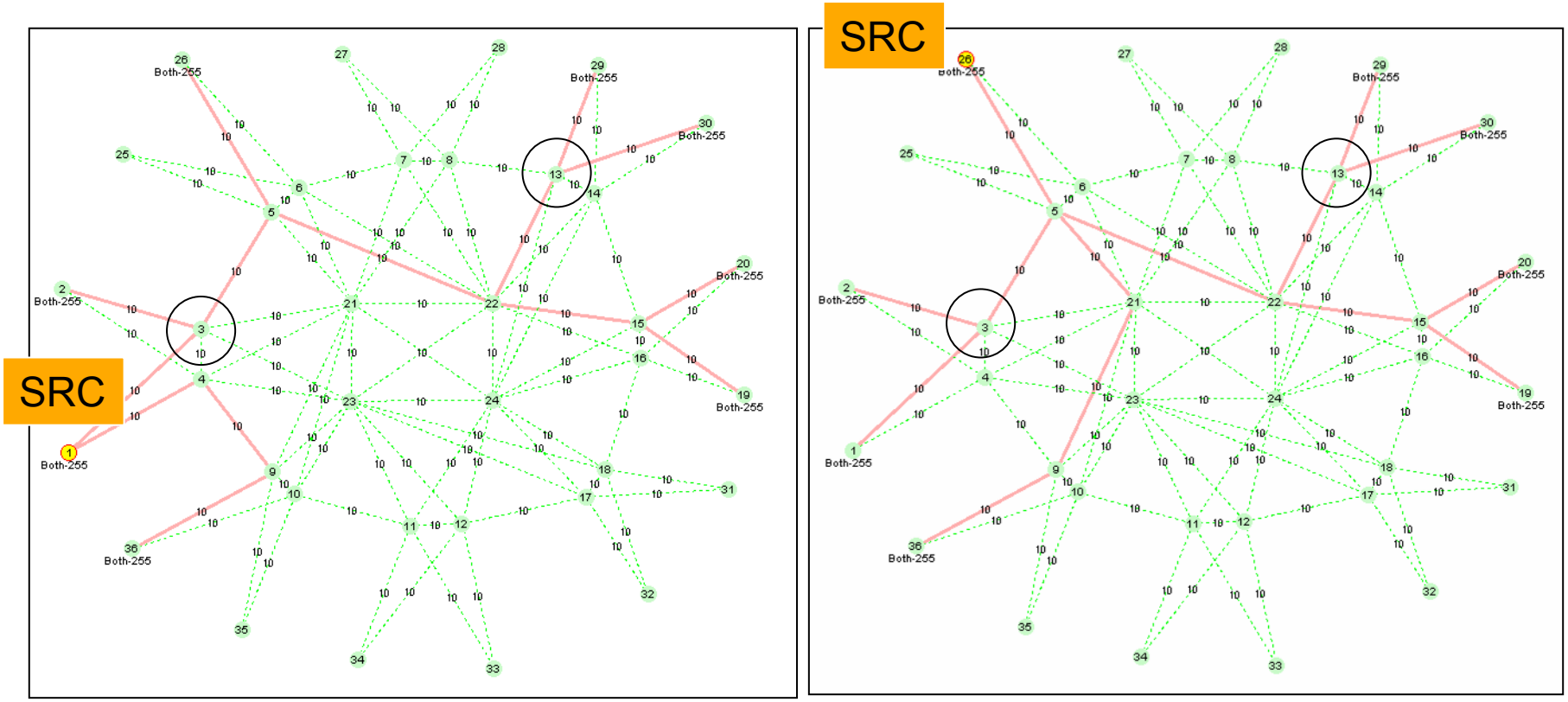


EXAMPLE – NODE :3 ROUTE TO :10 (second equal cost path)

```
<ottawa-9300-3>d spb umac
-----------------------------------------------
    BMAC            BVLAN        IF NAME
-----------------------------------------------
4455-6677-0001      100          GE2/0/11
4455-6677-0001      101          GE2/0/11
4455-6677-0004      100          GE2/0/12
4455-6677-0004      101          GE2/0/12
4455-6677-0005      100          GE2/0/16
4455-6677-0005      101          GE2/0/16
4455-6677-0006      100          GE2/0/16
4455-6677-0006      101          GE2/0/17
4455-6677-0007      100          GE2/0/17
4455-6677-0007      101          GE2/0/17
4455-6677-0008      100          GE2/0/17
4455-6677-0008      101          GE2/0/17
4455-6677-0009      100          GE2/0/12
4455-6677-0009      101          GE2/0/18
4455-6677-000a      100          GE2/0/17
4455-6677-000a      101          GE2/0/18
4455-6677-000b      100          GE2/0/18
4455-6677-000b      101          GE2/0/18
4455-6677-000c      100          GE2/0/18
4455-6677-000c      101          GE2/0/18
4455-6677-000d      100          GE2/0/16
4455-6677-000d      101          GE2/0/18
4455-6677-000e      100          GE2/0/16
4455-6677-000e      101          GE2/0/18
4455-6677-000f      100          GE2/0/16
4455-6677-000f      101          GE2/0/18
4455-6677-0010      100          GE2/0/16
4455-6677-0010      101          GE2/0/18
...
4455-6677-0024      100          GE2/0/12
4455-6677-0024      101          GE2/0/18
 Total unicast fib entries is 68
<ottawa-9300-3>
```

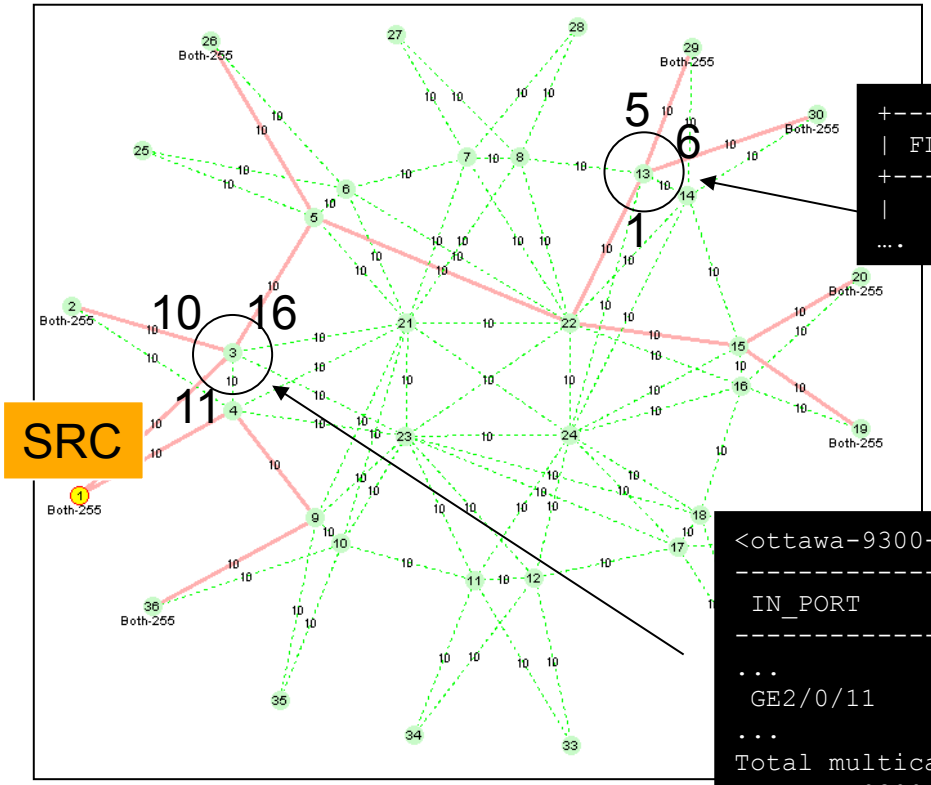


EXAMPLE: E-LAN MCAST ROUTES FROM :1 (left) and :26 (right)



Here are the multicast routes from node 1 for service 255 and also from node 26 for service 255. Note the symmetry in the route between the two multicast trees. The unicast route between :1 and :26 is also along that same path for the chosen B-VID. Since we've asked for transit replication for all members of the E-LAN we install MCAST …

EXAMPLE: E-LAN MCAST ROUTES FROM :1 (left) and :26 (right)




```
+-----+-------+-------------------+------+--------------
| FLG | IN/IF | DESTINATION ADDR  | BVID | OUT/IF(s)
+-----+-------+-------------------+------+--------------
|     | if/01 |   7300-0100-00ff  | 0100 | {if/5,if/6 }
….
```

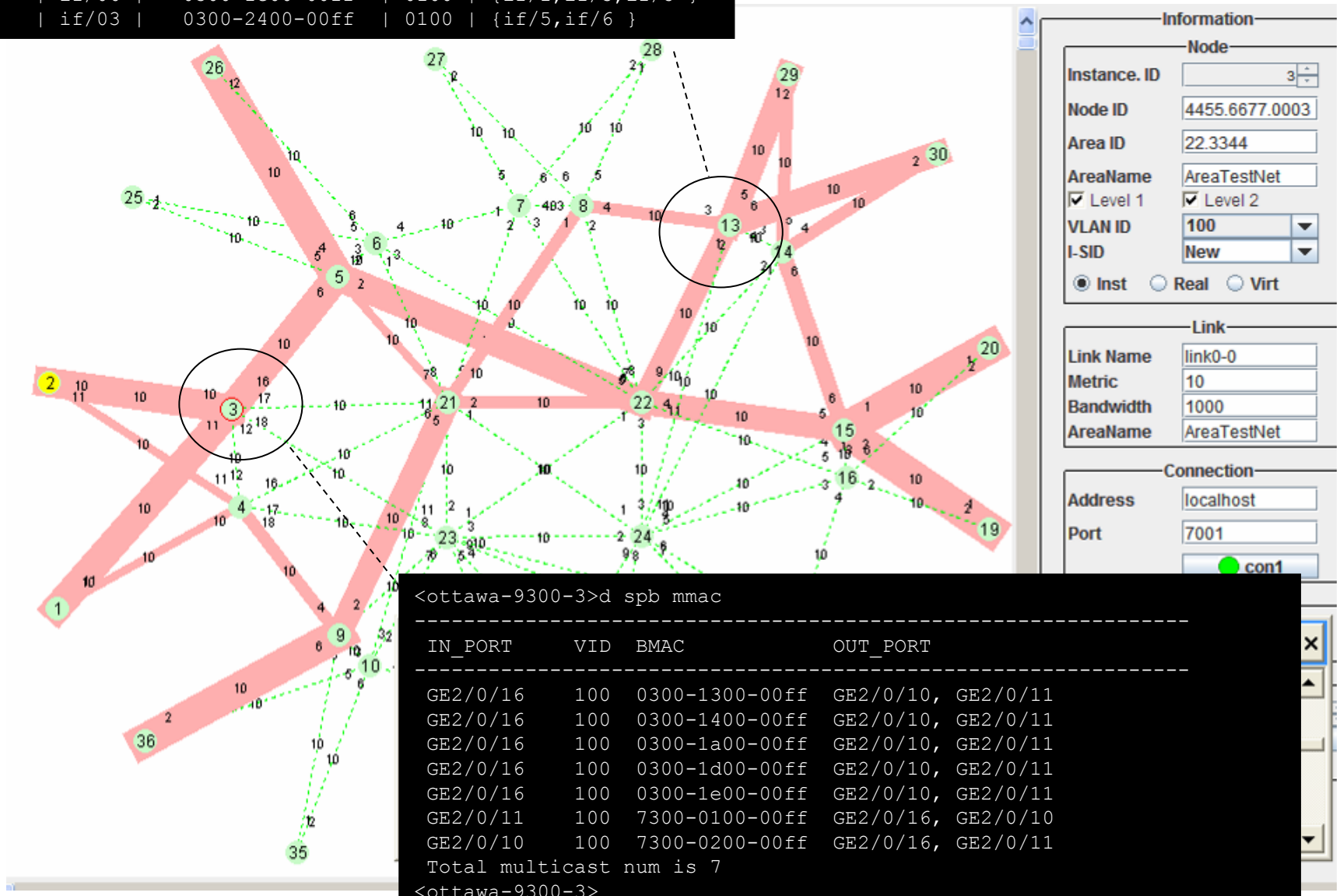
```
<ottawa-9300-3>d spb mmac
-----------------------------------------------------------------
 IN_PORT     VID  BMAC            OUT_PORT
-----------------------------------------------------------------
...
 GE2/0/11    100  7300-0100-00ff  GE2/0/16, GE2/0/10
...
Total multicast num is 7
<ottawa-9300-3>
```

MULTICAST ADDRESS IS: [SOURCE = 07-00-01 | ISID=00-00-ff]

We only get this state if we configure transmit membership in the E-LAN.
Transmit still possible without multicast state but uses serial replication at head end.
Operator chooses trade-off between state/bandwidth usage.

```
+-----+-------+-------------------+------+--------------
| FLG | IN/IF | DESTINATION ADDR  | BVID | OUT/IF(s)
+-----+-------+-------------------+------+--------------
|     | if/01 |   7300-0100-00ff  | 0100 | {if/5,if/6 }
|     | if/01 |   7300-0200-00ff  | 0100 | {if/5,if/6 }
|     | if/01 |   0300-1a00-00ff  | 0100 | {if/5,if/6 }
|     | if/05 |   0300-1d00-00ff  | 0100 | {if/1,if/3,if/6 }
|     | if/06 |   0300-1e00-00ff  | 0100 | {if/1,if/3,if/5 }
|     | if/03 |   0300-2400-00ff  | 0100 | {if/5,if/6 }
```

Here are all mFIBs on nodes :3 and :13 related to this E-LAN.




```
<ottawa-9300-3>d spb mmac
-----------------------------------------------------------------
 IN_PORT     VID  BMAC            OUT_PORT
-----------------------------------------------------------------
 GE2/0/16    100  0300-1300-00ff  GE2/0/10, GE2/0/11
 GE2/0/16    100  0300-1400-00ff  GE2/0/10, GE2/0/11
 GE2/0/16    100  0300-1a00-00ff  GE2/0/10, GE2/0/11
 GE2/0/16    100  0300-1d00-00ff  GE2/0/10, GE2/0/11
 GE2/0/16    100  0300-1e00-00ff  GE2/0/10, GE2/0/11
 GE2/0/11    100  7300-0100-00ff  GE2/0/16, GE2/0/10
 GE2/0/10    100  7300-0200-00ff  GE2/0/16, GE2/0/11
 Total multicast num is 7
<ottawa-9300-3>
```

References

"***IEEE 802.1aq***" : www.wikipedia.org:
http://en.wikipedia.org/wiki/IEEE_802.1aq

http://www.ietf.org/internet-drafts/draft-ietf-isis-ieee-aq-00.txt The IETF IS-IS draft (check for later version 01.. etc).

"***IEEE 802.1aq***" www.ieee802.org/1/802-1aq-d2-6.pdf

"***Shortest Path Bridging*** – Efficient Control of Larger Ethernet Networks" : upcoming IEEE Communications Magazine – Oct 2010

"***Provider Link State Bridging***" :
IEEE Communications Magazine V46/N9– Sept 2008
http://locuhome.com/wp-content/uploads/2009/02/
ieeecommunicationsmagazinevol46no9sep2008-carrierscaleethernet.pdf

Thank-You

HUAWEI