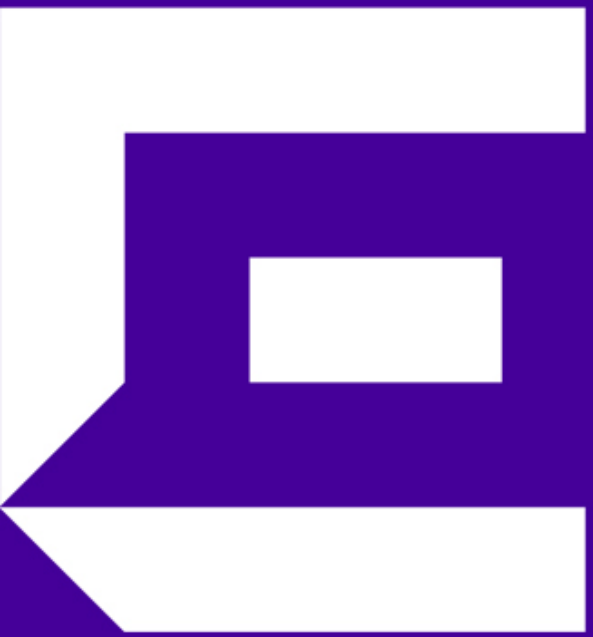
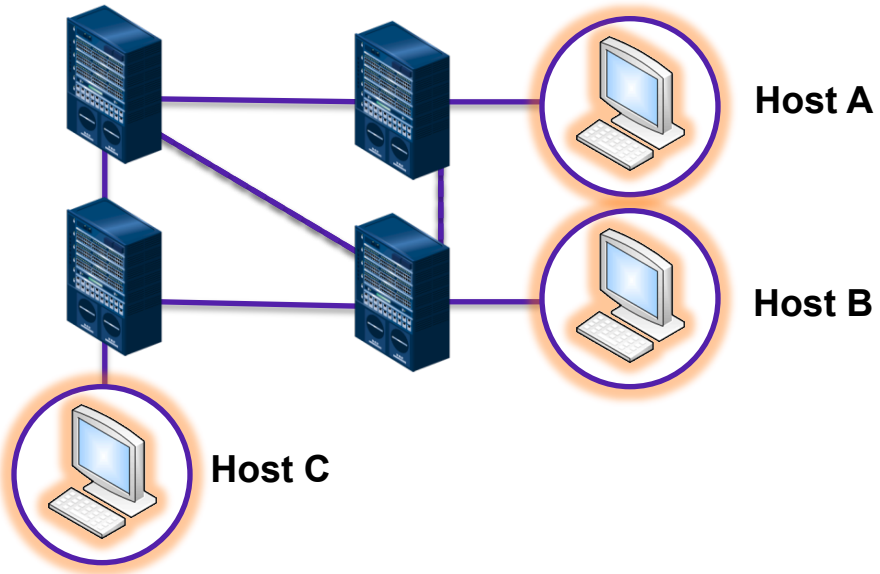


TRILL vs. SPB

Mikael Holmberg
Senior Principal Corporate Systems Engineer EMEA
mikael@extremenetworks.com

Extreme networks

Introduction – *Data Center Switching Evolving!*

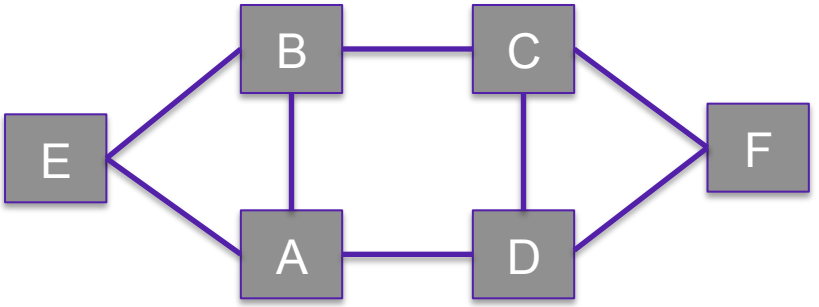


Spanning Tree Challenges

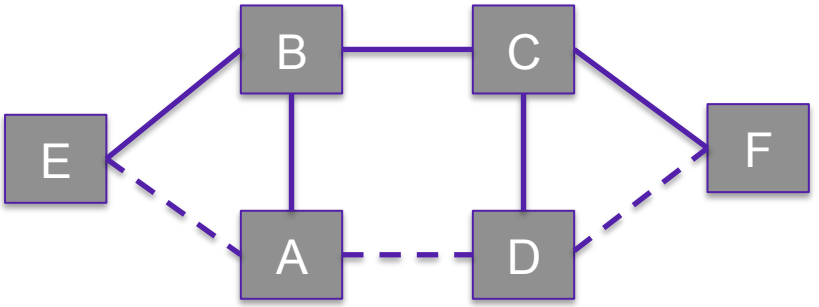
- STP introduced Blocked Ports leading to Inefficient Paths
- STP has slow convergence (in seconds) and is disruptive
- Less Aggregate Bandwidth
- MAC address tables don't scale
- Instability with Multicast Optimization
- *Could IP help?... Yes… but….*

Optimum Forwarding

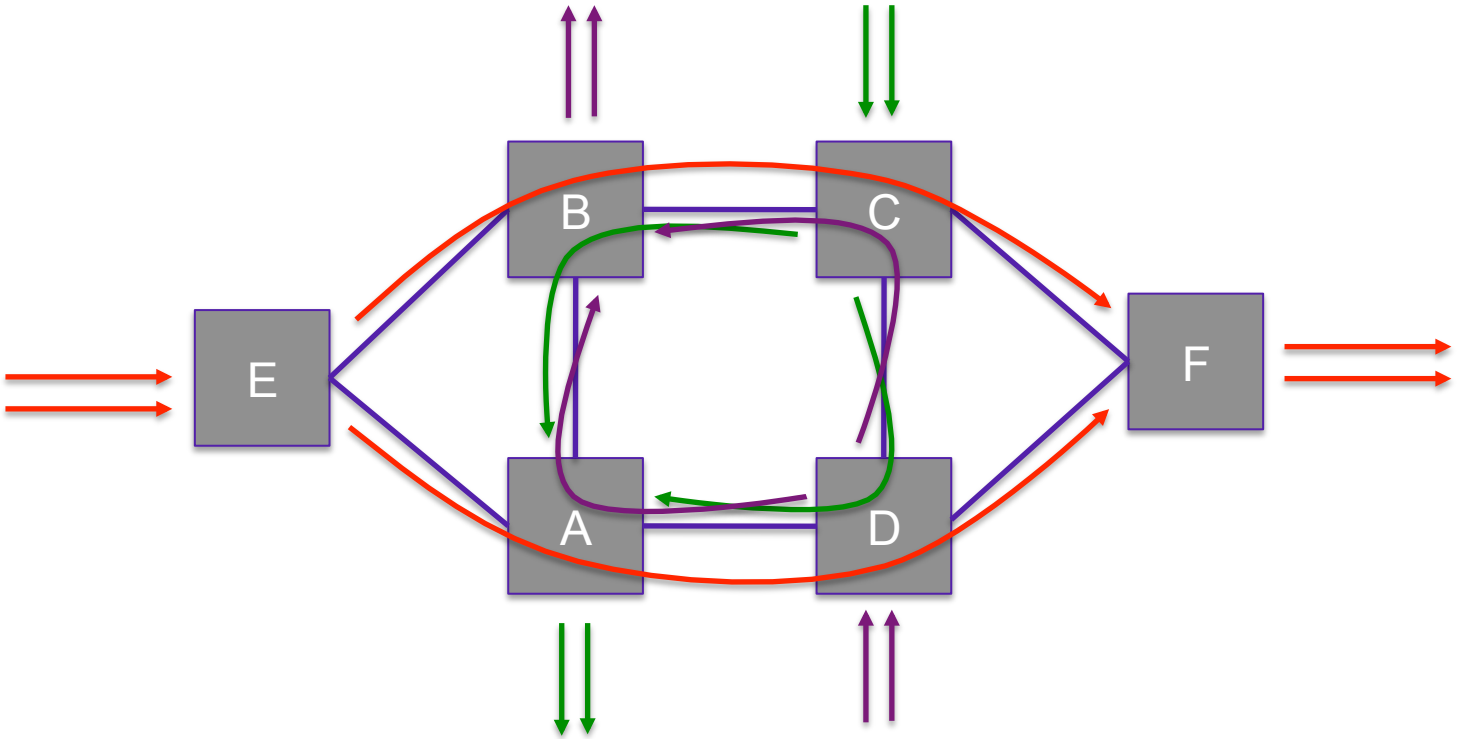
- In the following 6 Bridge Network:
  - Optimum forwarding would use all 8 links



- Loop avoidance protocols reduce available links
- Traffic limited to only one path
  - (STP, ERPS, EAPS)

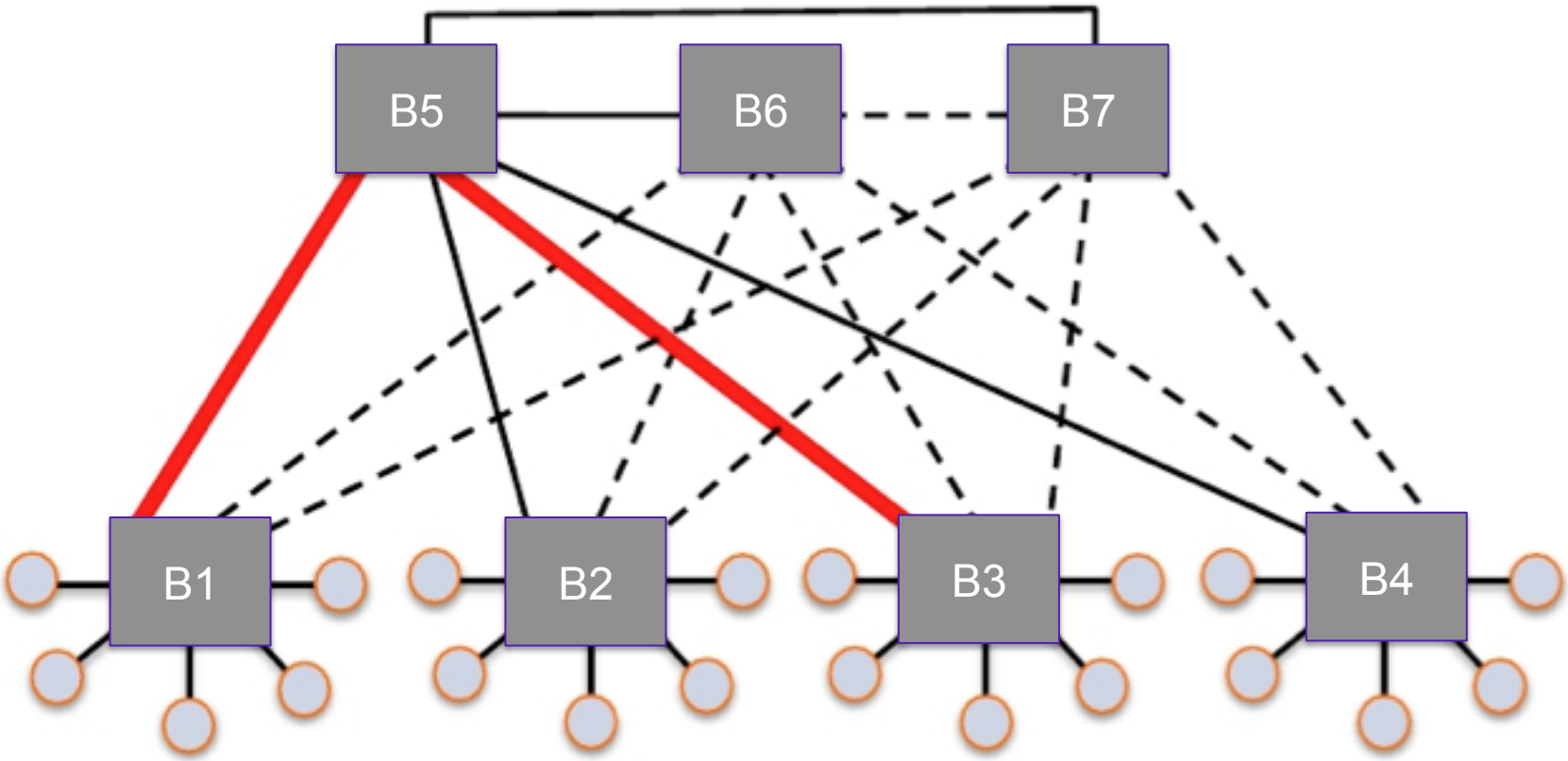


Multi-Pathing



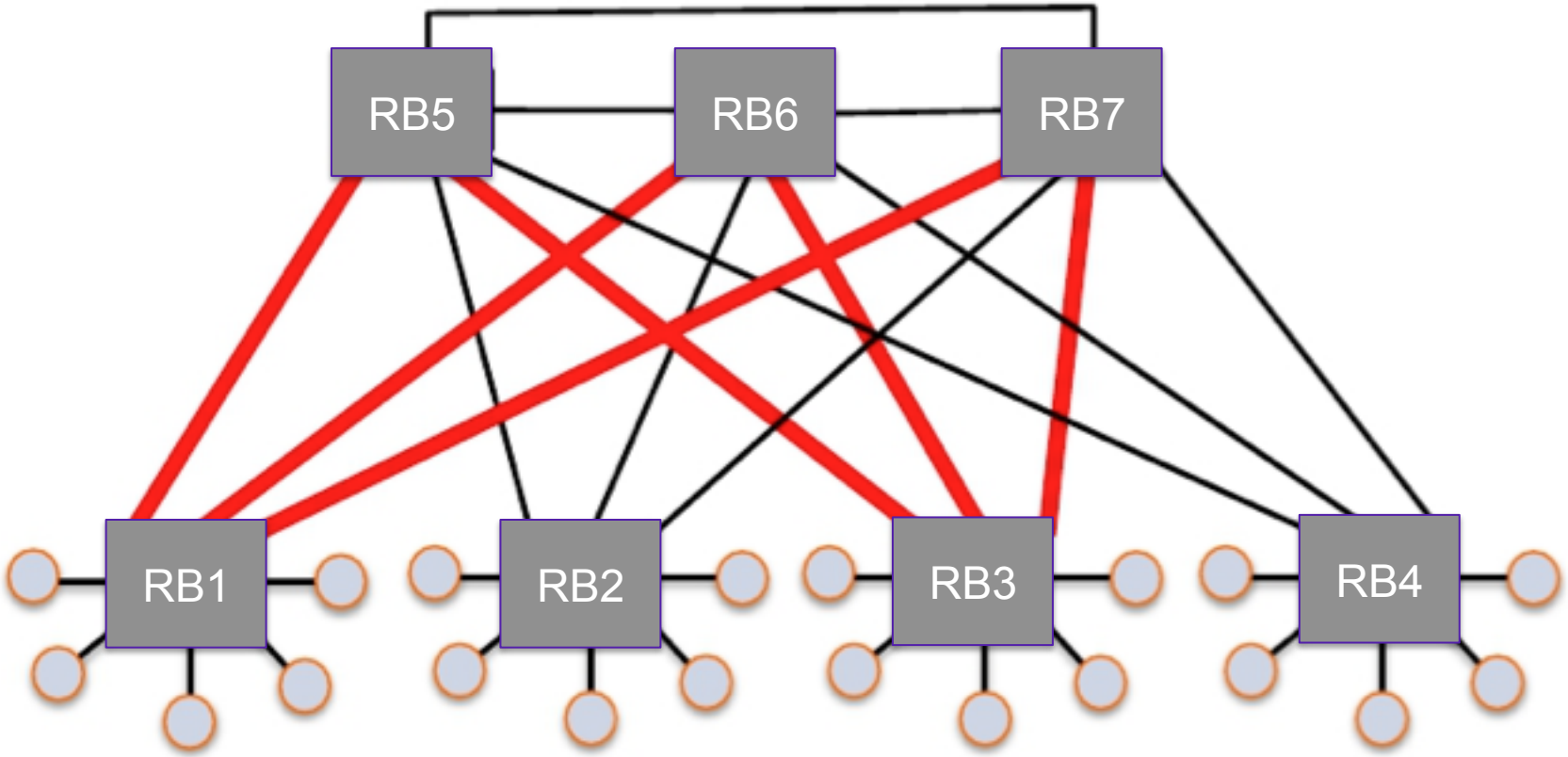
Multi-Pathing

- Bridges limit traffic to one path



Multi-Pathing

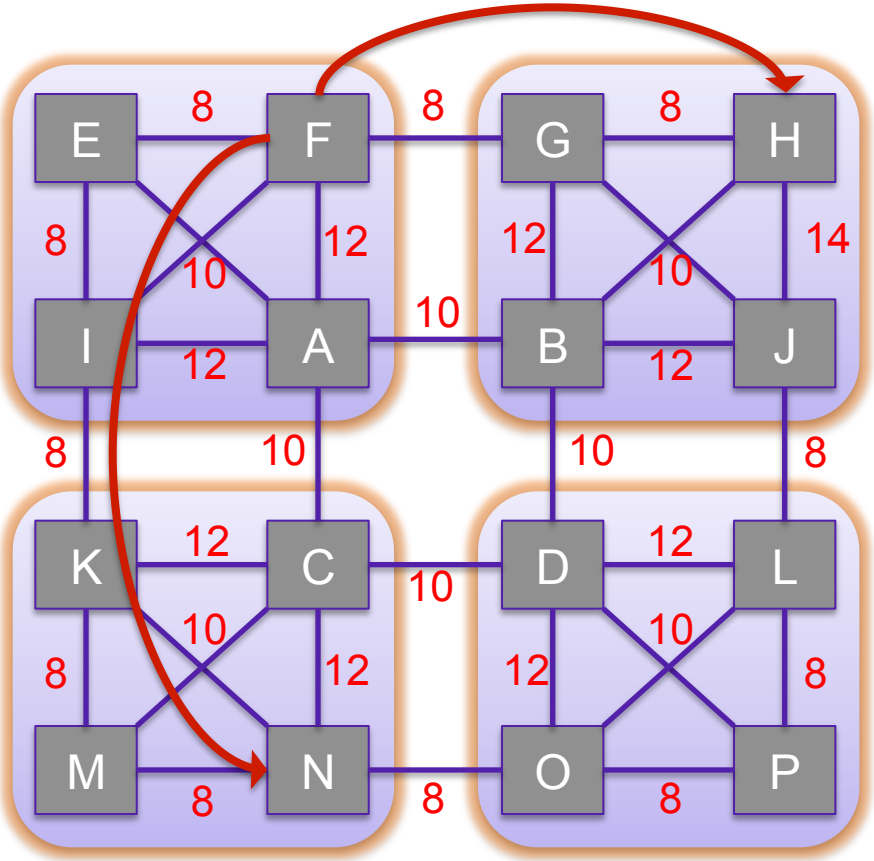
- You want something that would support multi-path for higher throughput…



Path Computation

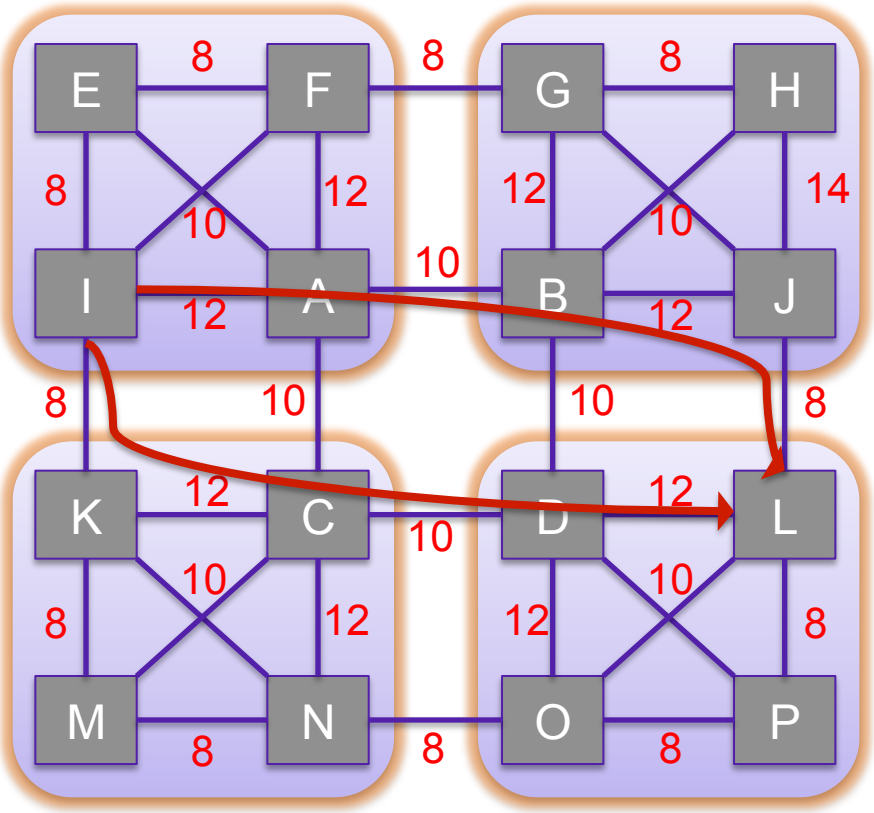
(IS-IS) UNICAST PATH CALCULATION

- TRILL uses the Dijkstra Algorithm, to calculate the best path route based on link cost to every node in the network
- Each node makes an independent decision on where to send a packet based on the packet's destination egress node
- F to H:
  - F-G-H = path cost 16
- F to N:
  - F-I-K-N = path cost 28



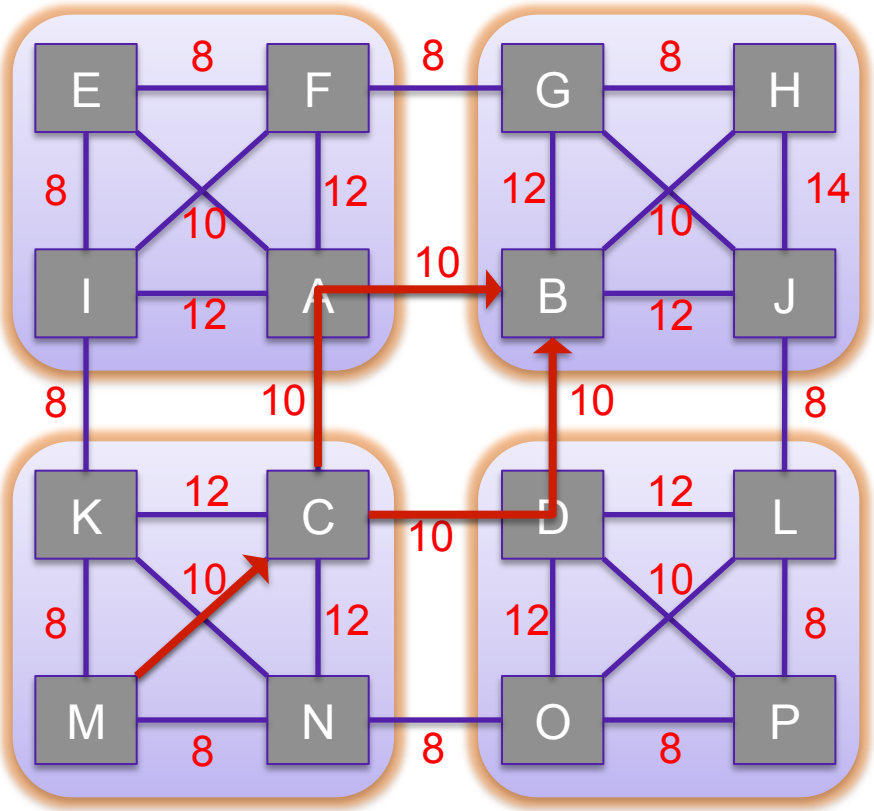
(IS-IS) MULTIPATH CALCULATION

- A link state algorithm allows multipath forwarding
- Multipath forwarding allows the ingress node to forward packets along multiple paths to reach the destination, so long as they are all considered to be the best path
- The ingress node uses a hashing algorithm to select the next hop peer.
  - The hashing algorithm operates on the encapsulated packet header so that individual flows always follow the same path
  - This can lead to bi-directional traffic flows taking different paths based on the hash
- I to L:
  - I-A-B-J-L= cost 42
  - I-K-C-C-L= cost 42



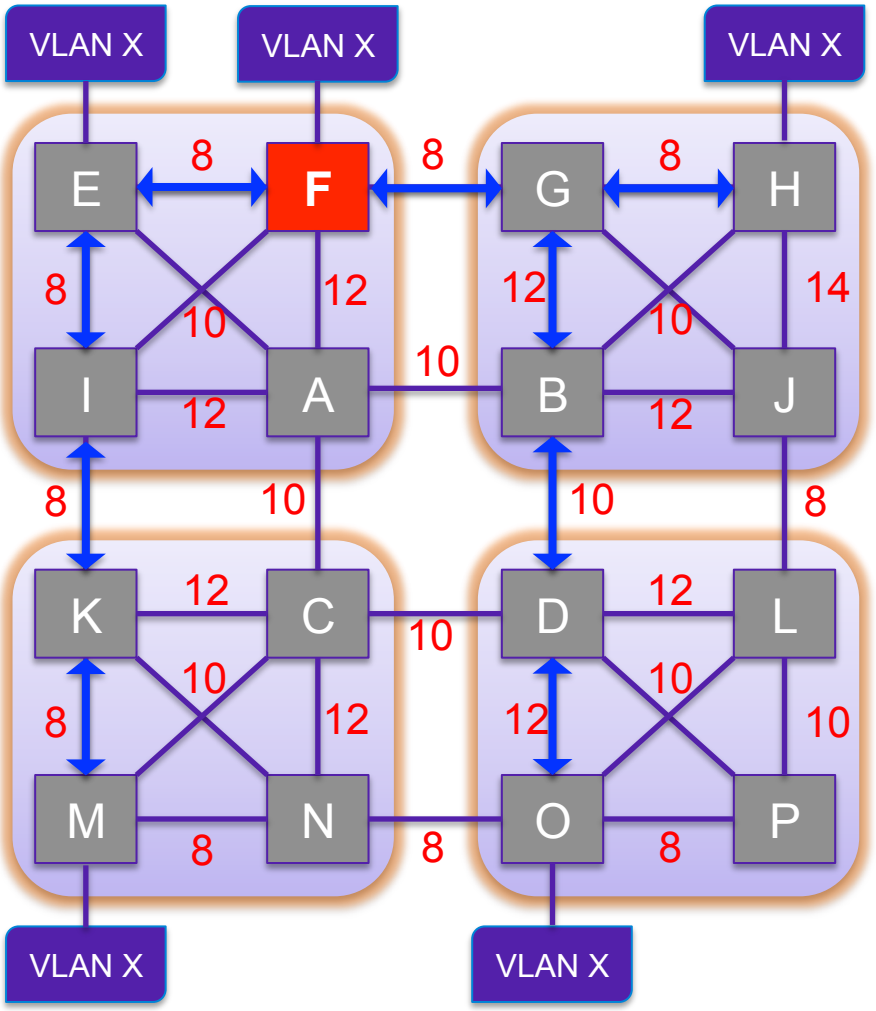
(IS-IS) PER HOP MULTIPATH CALCULATION

- Each hop along the path performs its own next hop look-up independently of the previous hops
- At each hop along the path, there may be multiple paths that were not available to the previous hops
- This provides another level of load sharing not available to Layer 2 networks
  - *This is not currently supported in Service Provider Bridging (SPB).*
- M to B:
  - Shortest path is via C
  - C to B:
    - C-A-B = path cost 20
    - C-D-B = path cost 20



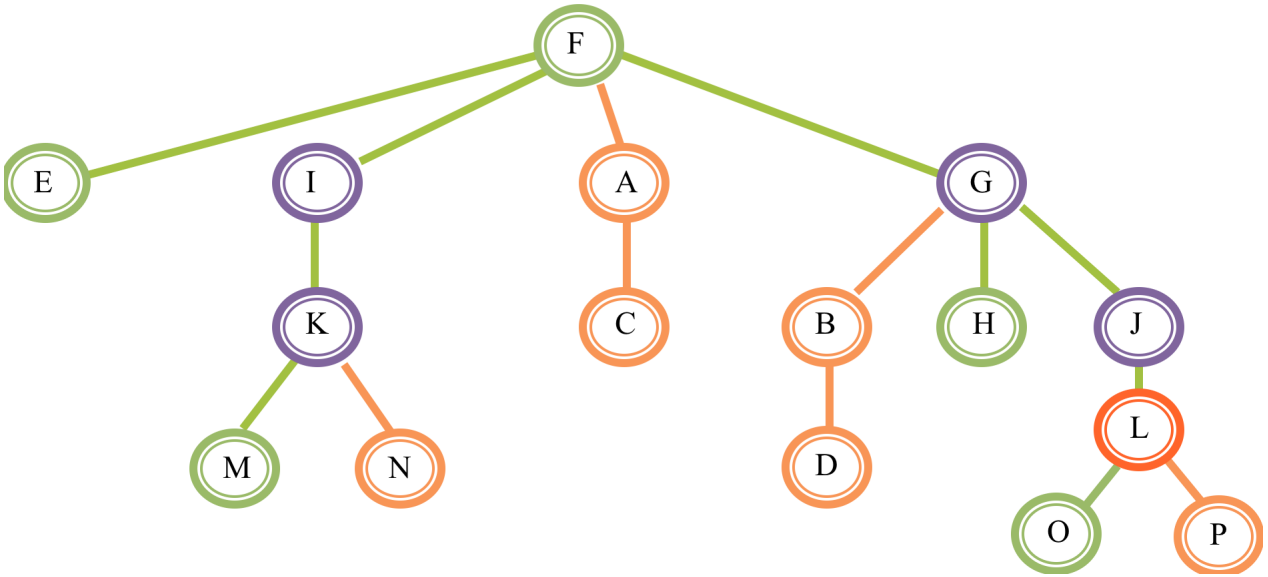
Multi-Destination Trees (TRILL)

- Broadcast, Multicast and Unknown Unicast packets are forwarded using Multicast Distribution Tress
- RBridges compute a single shared tree based on LSP database for all multi-destination traffic
- Multiple trees can be computed to load-share across multiple equal cost links
- RBridge with highest priority becomes the “TREE Root” and all distribution trees are rooted from here



Multicast Distribution Trees (TRILL)

- VLAN X attached at F,E,H,M and O
- RBridge **F** has been configured with the highest priority Root Distribution Tree
- Rbridge forwards packets with VLAN tags to only those tree adjacencies that have downstream matching Access VLANs
- RBridges K, G, and L are not required to forward traffic to some or all of the distribution tree adjacencies.
  - This effectively prunes the distribution tree and reduces packet replication and unnecessary traffic forwarding.



TRILL

Extreme networks

TRILL: Transparent Interconnection of Lots of Links

- IETF standard for L2 scalability
- Inventor of STP is inventor of TRILL – *Radia J. Perlman*
- Many RFCs:
  - RFC 5556: Problem & Applicability Statement
  - RFC 6325: Routing Bridges (RBridges): Base Protocol Specification
  - RFC 6326: TRILL use of IS-IS
  - RFC 6327: Routing Bridges Adjacency
  - RFC 6439: Routing Bridges Appointed Forwarders

WHY IS-IS …?

- The IS-IS (Intermediate System to Intermediate System) link state routing protocol was chosen for SPB over OSPF (Open Shortest Path First), the only other plausible candidate, for the following reasons:
  - IS-IS runs directly at Layer 2. Thus no IP addresses are needed, as they are for OSPF, and IS-IS can run with zero configuration.
  - IS-IS uses a TLV (type, length, value) encoding which makes it easy to define and carry new types of data.

Introduction – *Best of the Both Worlds!*

Extreme networks

L2 Switching

- Minimal Configuration
- Plug & Play
- Flat Addressing
- Slow Convergence
- Single Path
- Single Multicast Tree
- Constrained Scalability

TRILL

- Minimal Configuration
- Plug & Play
- Fast Convergence
- Multiple Paths
- Load Balancing
- Hierarchical Forwarding
- Highly Scalable

L3 Routing

- Plan & Play
- Fast Convergence
- Multiple Paths
- Load Balancing
- Hierarchical Forwarding
- Multiple Multicast Trees
- Highly Scalable

TRILL: Transparent Interconnection of Lots of Links

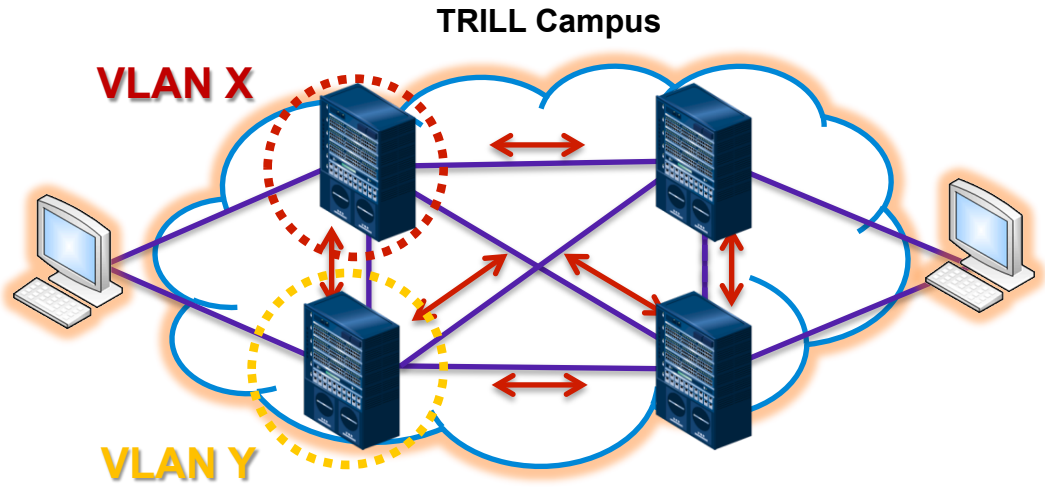
TRILL Header (8 bytes including TRILL Ethertype)

- ✓ **M (1-bit): Multi-destination bit (0 = Unicast, 1 = Multi-destination)**
- ✓ **Hop Count (6-bit): Mitigates Loop issues**
- ✓ **Nicknames (16-bit): Dynamically assigned through nickname acquisition protocol**

Dynamic Nickname Acquisition Protocol

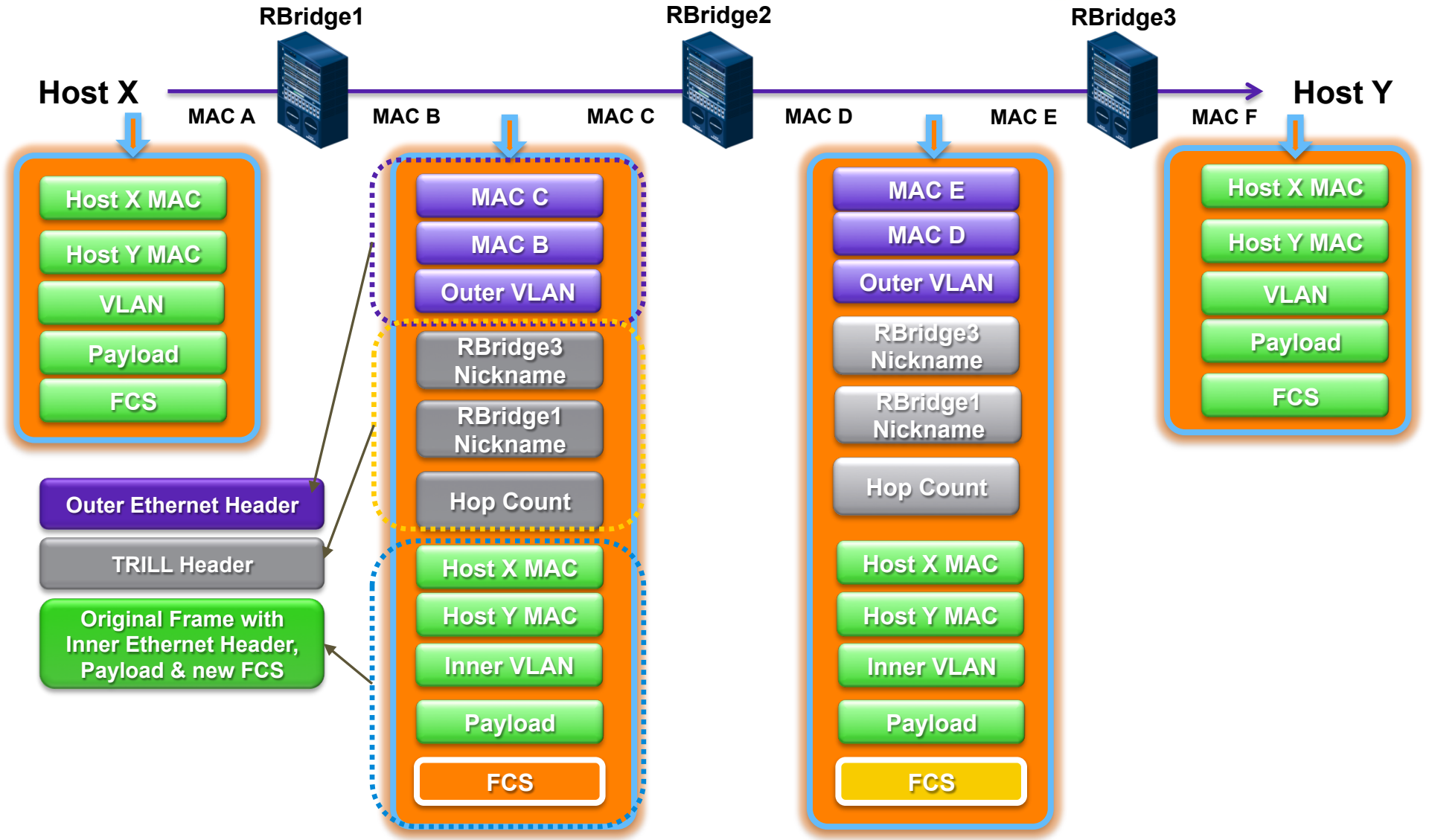
- ✓ Nicknames are manually configured or dynamically assigned
- ✓ Dynamic nicknames based on hashing parameters (System ID, time, date etc.)
- ✓ RBridge Nicknames advertised using Link State PDUs (LSP)
- ✓ Priority of the nickname is advertised in the LSP
- ✓ Nicknames are persistent across reboots

TRILL Basic Interworking



- ✓ **RBridges exchanges TRILL IS-IS Hello frames**
  - ✓ Hellos establish IS-IS connectivity on RBridge port
  - ✓ RBridges elect Designated RBridge (DRB) for each link
- ✓ **RBridges exchanges LSP to have a global link state database**
  - ✓ Includes information such as VLAN, Nicknames, link cost etc.
  - ✓ Calculates optimal paths for unicast and multi-destination traffic
- ✓ **DRB specifies the Appointed Forwarder for each VLAN**
  - ✓ Appointed Forwarders encapsulate/decapsulates TRILL data frames

TRILL Packet Encapsulation (Unicast Frames)



SPB

What is SPB? (1)

- 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 protocols
- Multiple loop free shortest paths routing
  - Excellent use of mesh connectivity
  - Currently 16 equal cost paths.
  - Optimum multicast head end or tandem replication

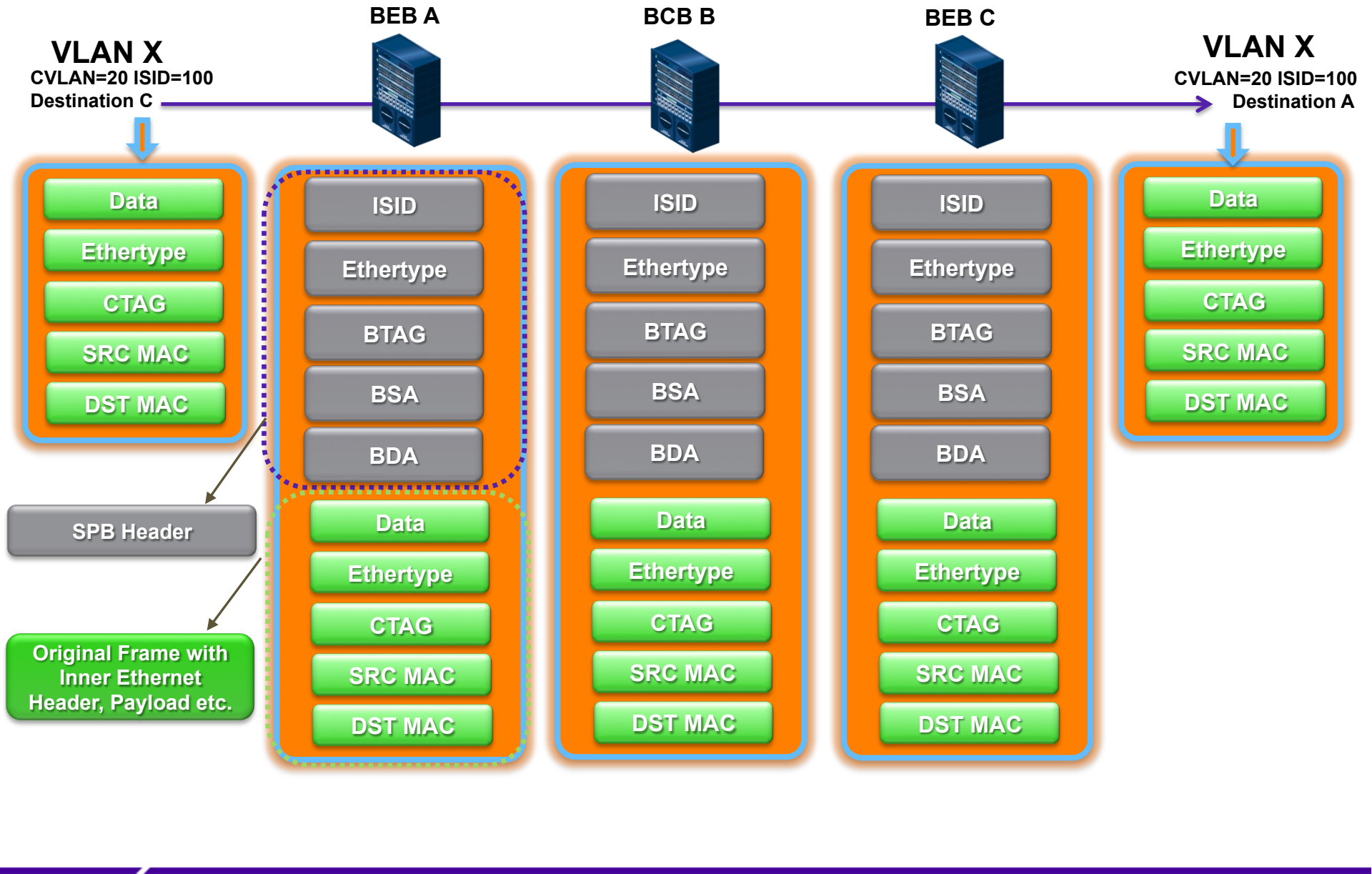
What is SPB? (2)

- 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 complexity
  - sub second (modern processor, well designed)
  - below 100ms (use of hardware multicast for updates)
  - Includes multicast flow when replication point dies.

What is SPB? (3)

- Service membership advertised in same protocol as Topology
  - Minimizes complexity, near plug-and-play
  - Support E-LINE/E-LAN/E-TREE
  - 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.
- Computations guarantee unicast/multicast:
  - Symmetry (same in both directions)
  - Congruence (unicast/multicast follow same route)
  - Tune-ability (currently 16 equal costs paths)

SPBM Packet Encapsulation



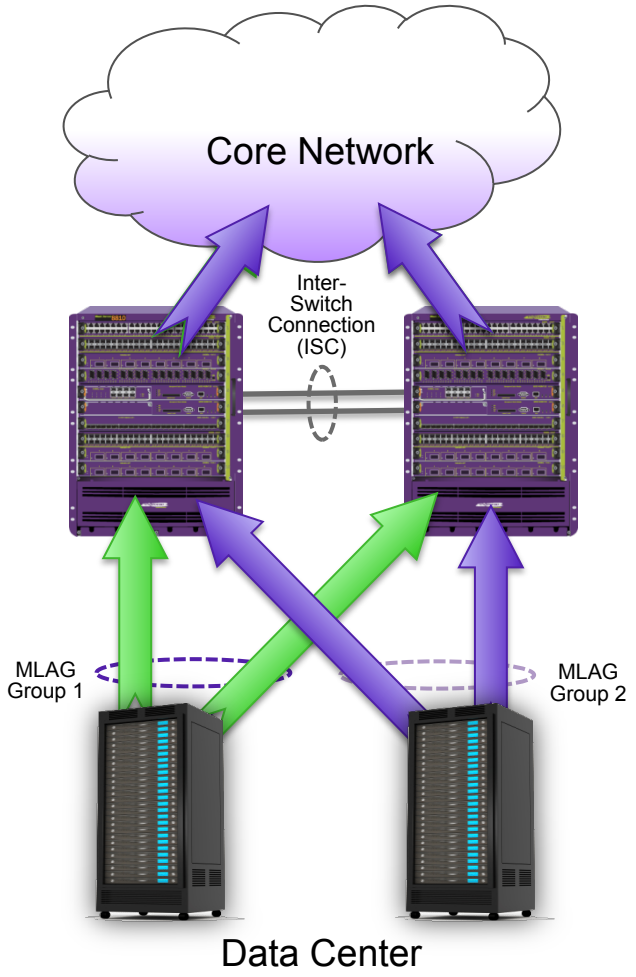
Comparison to MLAG, SPB, VPLS/MPLS

Extreme networks

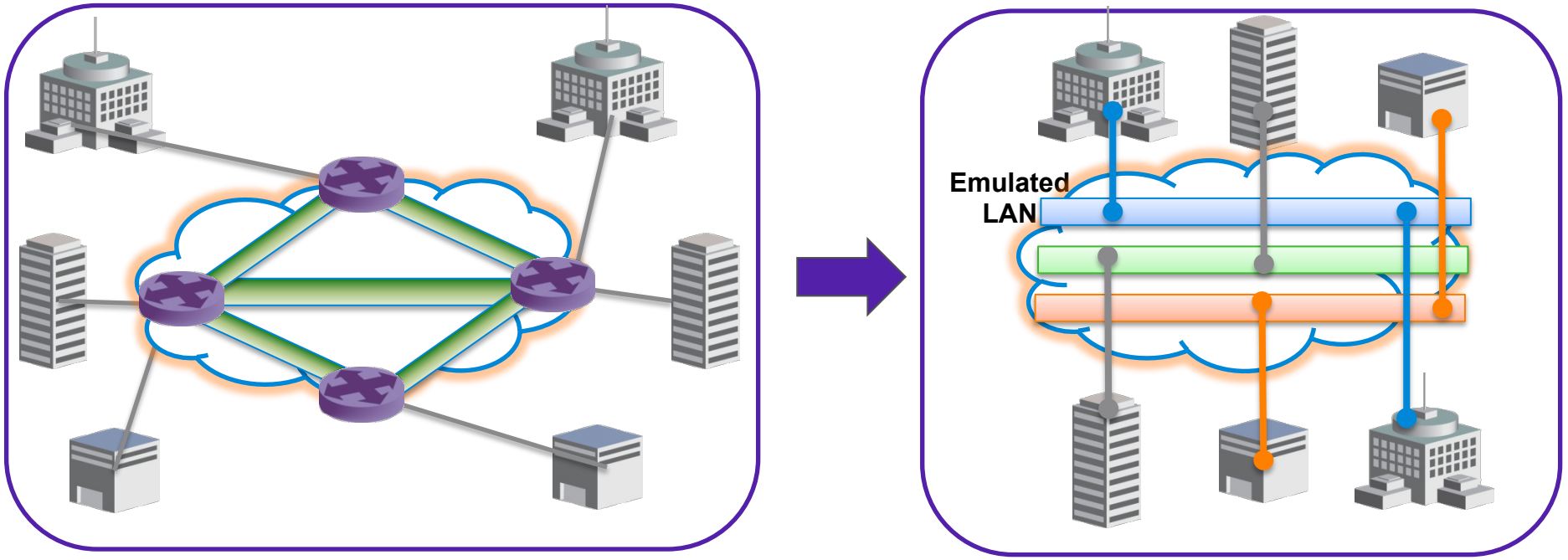
M-LAG for Active-Active Paths

Efficient Bandwidth Usage

- LAG allows combining of ports effectively increasing the bandwidth. Up to 64 ports in a LAG Group.
- M-LAG allows combining of ports on ‘two’ switches to form a single logical connection to another network device
- Active-active paths. No STP port blocking
- Fast Failover
- For both Layer-2 and Layer-3 deployments
- Interoperates across tiers
- Works with servers, switches, storage, and other network appliances



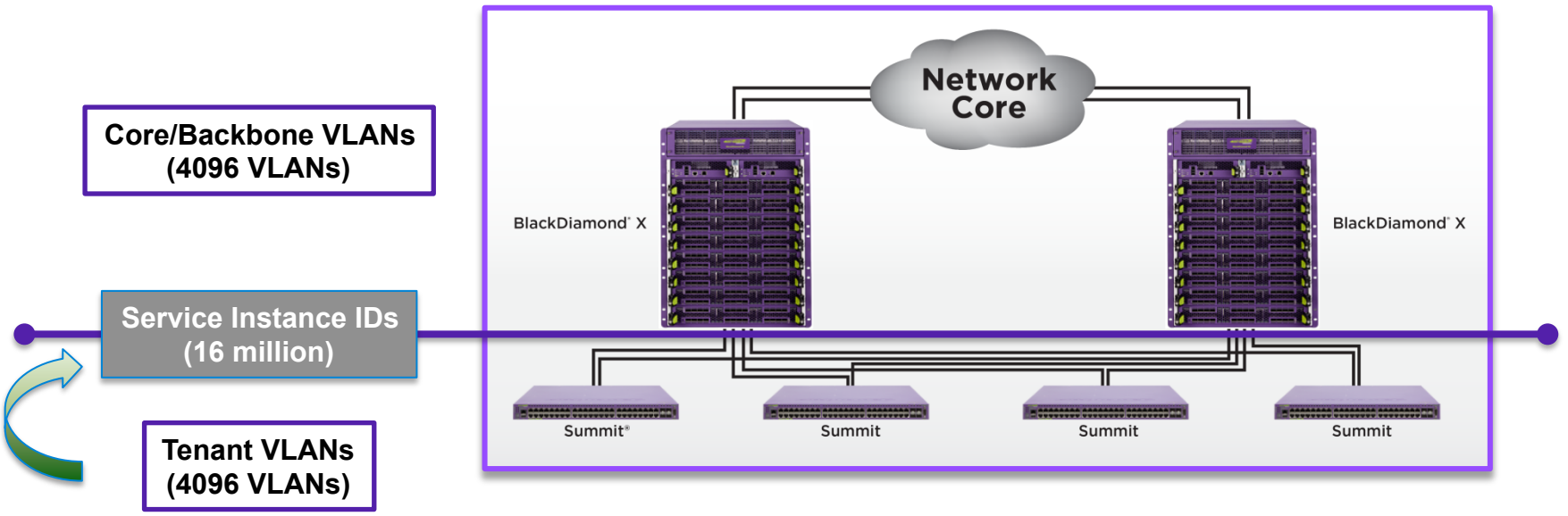
Virtual Private LAN Service (VPLS) – RFC 4761/4762



- L2 Ethernet VPN providing multi-point communication over IP/MPLS networks
- All tenants sites appear to be on the same LAN regardless of location
- VPLS provides VLAN extensions over IP/MPLS networks
- Each tenant VLAN is mapped to a virtual switch instance or VPN ID

Shortest Path Bridging (SPB) – IEEE 802.1aq

- Equal Cost Multi-Path (ECMP) solution (up to 16 trees)
- Large L2 bridging topologies (up to 16 million) based on IS-IS as link state control protocol
- Service & Infrastructure separation



IEEE Standards Evolution to scale L2 Fabrics

IEEE 802.1	IEEE 802.1Q	IEEE 802.1ad	IEEE 802.1aq
Payload	Payload	Payload	Payload
Source Address	VLAN ID (12 bit)	VLAN ID (12 bit)	VLAN ID (12 bit)
Destination Address	Source Address	SVLAN ID (12 bit)	Source Address
	Destination Address	Source Address	Destination Address
		Destination Address	Service ID (24 bit)
			B VLAN ID (12 bit)
			Source Address
			Destination Address

Comparison

	TRILL	MLAG	SPB	VPLS
Standard Body	IETF	Vendor-specific	IEEE	IETF
Technology	New	Matured	New (Variant of PBB)	Matured
Minimal Configuration	Yes	Yes	Yes B-VID needs to be configured for each ECMP	No
ECMP	Yes 16 active links with true hop-by-hop ECMP decisions	Yes 2 active links	Yes 16 active links with ingress ECMP decisions	Yes 16 ECMP LSPs can be achieved
Loop Prevention	Yes TTL and RPC	Yes	Yes RPC only	Yes
Virtualization Scale	4K networks	4K networks	Higher scale with mac-in-mac	Higher scale with VPN ID

WWW.EXTREMENETWORKS.COM