Shortest Path Bridging

Additional information

Norman Finn

Addendum

- This deck contains material that is supplemental to aq-nfinn-shortest-path.pdf.
- It is not an update to that deck.

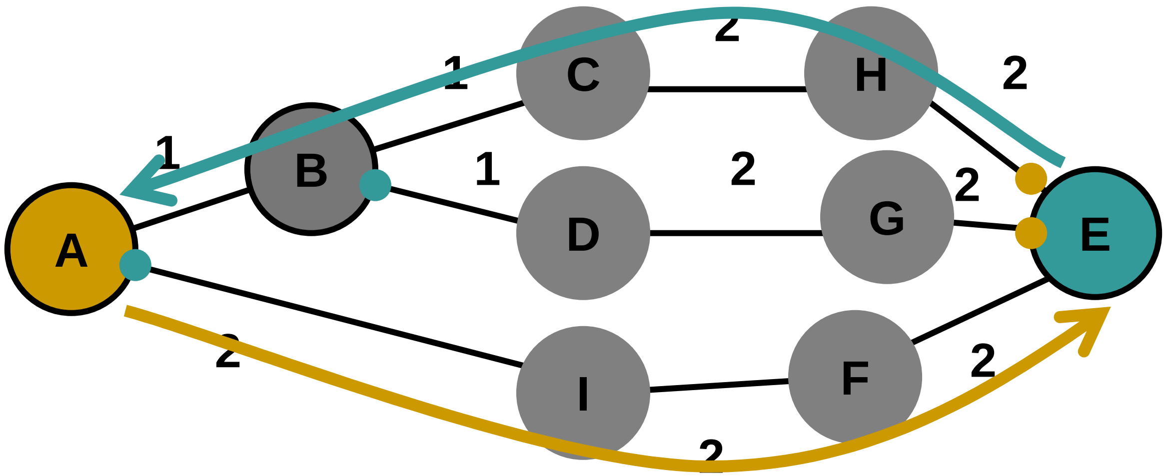
Quick Summary of Previous Work

Summary

Cisco.com

- Use link state protocols to create (at least) one spanning tree rooted at each bridge.
- Path from Bridge A to Bridge B on Bridge A's spanning tree is the same as the path on Bridge B's spanning tree.
- Frames are forwarded on a tree rooted at the ingress Bridge.
- Data plane is intact – frame forwarding is the same as in 802.1Q.

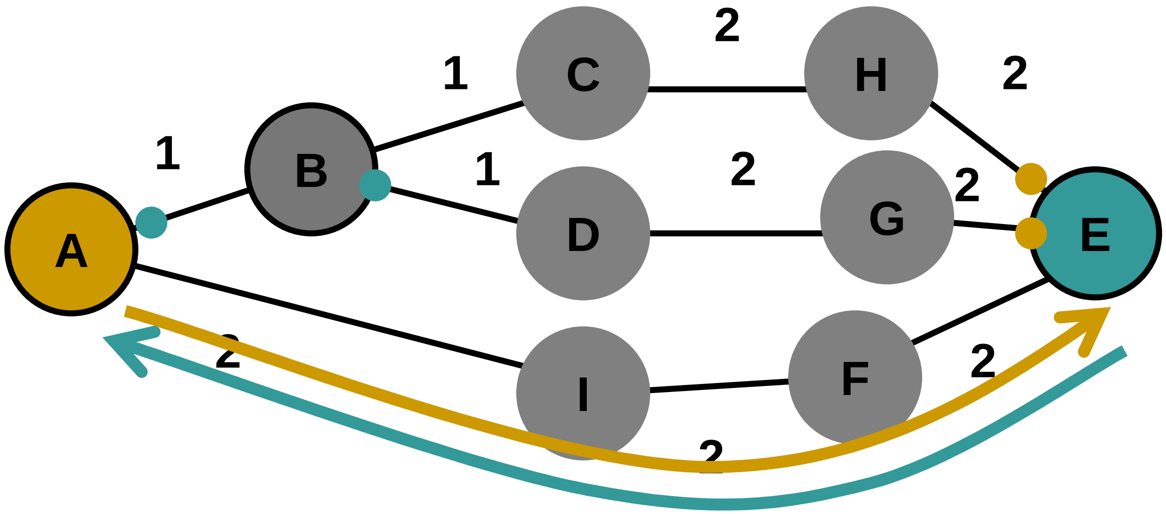
Asymmetrical Paths



- **If routing algorithms are applied naively, paths are asymmetrical, and learning doesn't work.**
  - **The path from E to A is E-H-C-B-A.**
  - **The path from A to E is A-I-F-E.**
  - **Learning doesn't work.**

Asymmetrical Paths

Cisco.com



- **It's easy to fix link-state routing algorithms to give symmetrical paths.**

  E.g. add the lowest-ID bridge encountered along the path as an additional metric, used to break the tie among equal-cost paths.

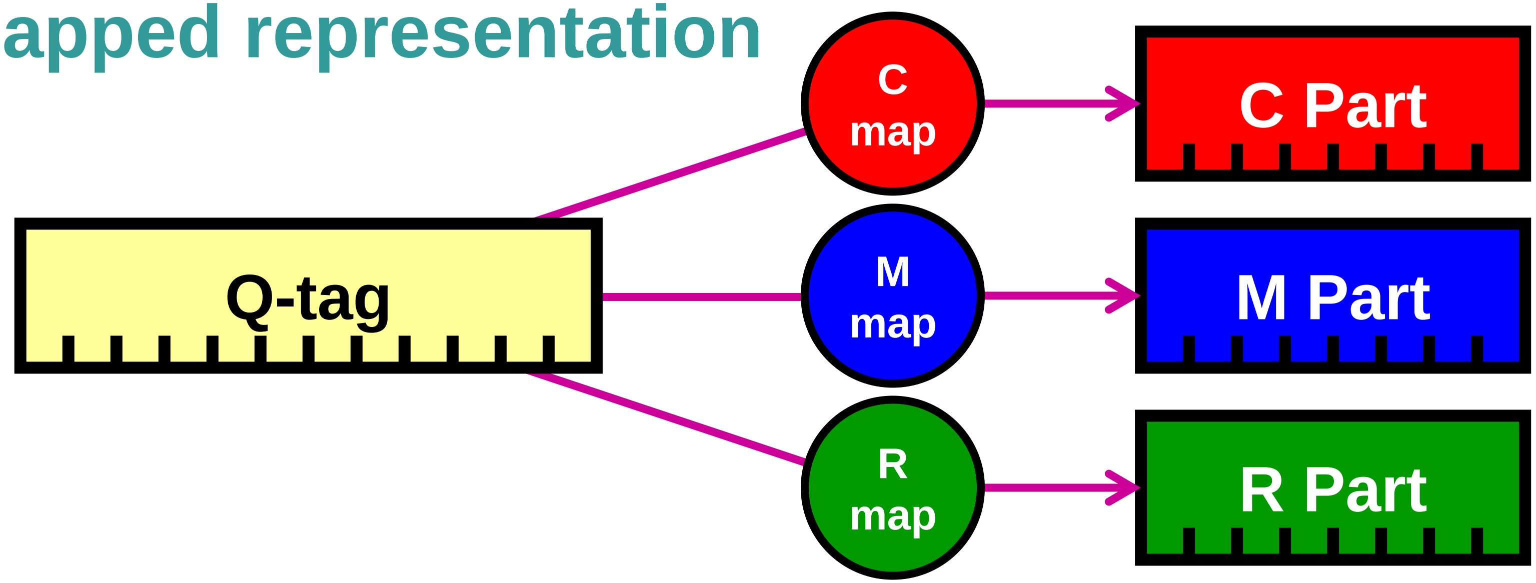
Why preserve learning?

- There is no reliable method for determining the list of hosts' MAC addresses attached to a bridge port.
- A new core with old edge bridges leads to large numbers of unicast MAC addresses to be passed in control packets.

Enterprise: Community-Multipath-Root (CMR) Tagging

Cisco.com

Mapped representation



- **C Part is the "Community of Interest".**
- **M Part is chooses among Multiple engineered topologies (if needed).**
- **R Part identifies the spanning tree Root.**

Provider Backbone
Community-Multipath-Root (CMR) Tagging

802.1ah representation

DA, SA	Type, Pri, DE	R + M part	Type Pri?DE?	C Part	Rest of the frame	CRC(s)
	802.1Q or .1ad B-tag		802.1ah I-tag			

- **Put the R and M parts in the B-tag.**
- **Put the Community part in the I-tag.**
- **DA, R, and M parts can be sufficient to do all the routing.**
- **C part can be meaningful only at the edge.**

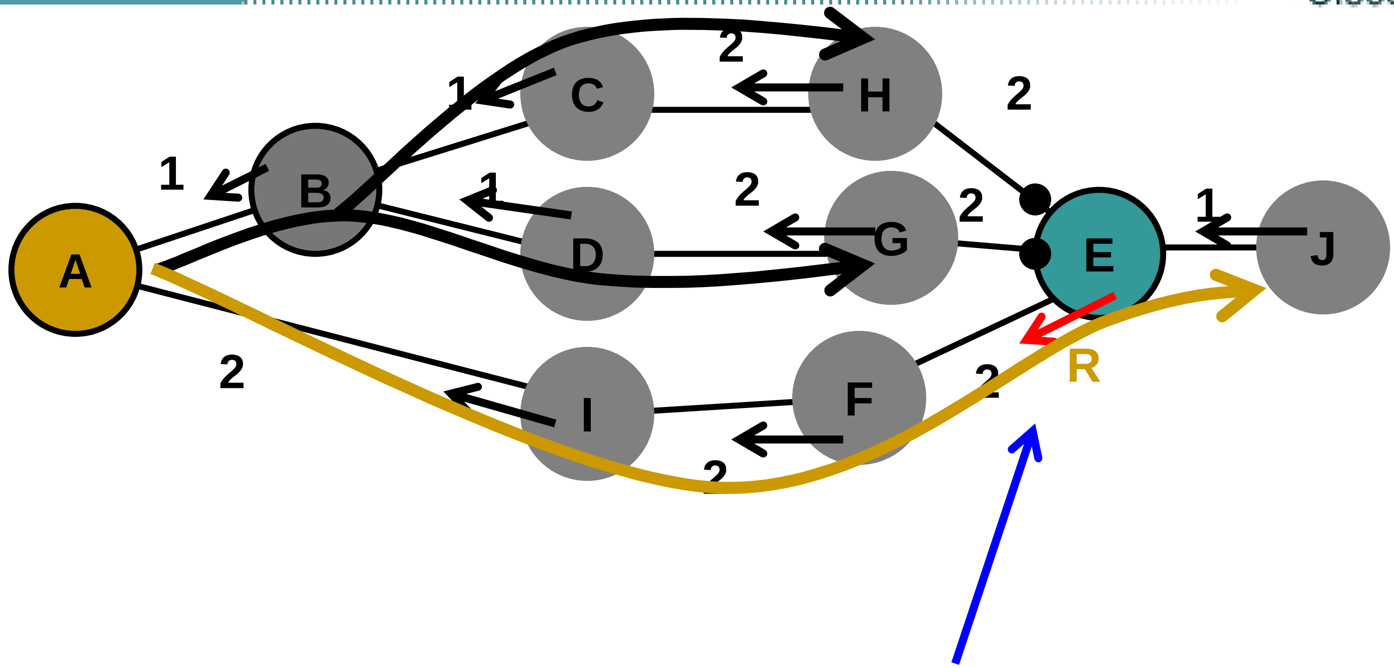
Link State Protocols

How to use Link State in Bridges

- One plan is to update your Bridges' software to calculate the Optimal Bridging multiple spanning trees using Link State information, instead of MSTP.
- Then, you get most of the claimed advantages of Routing in a Bridge.
- IS-IS is a solid protocol, and TLVs to carry MAC addresses instead of IP addresses are trivial to add.

Link State in Bridges

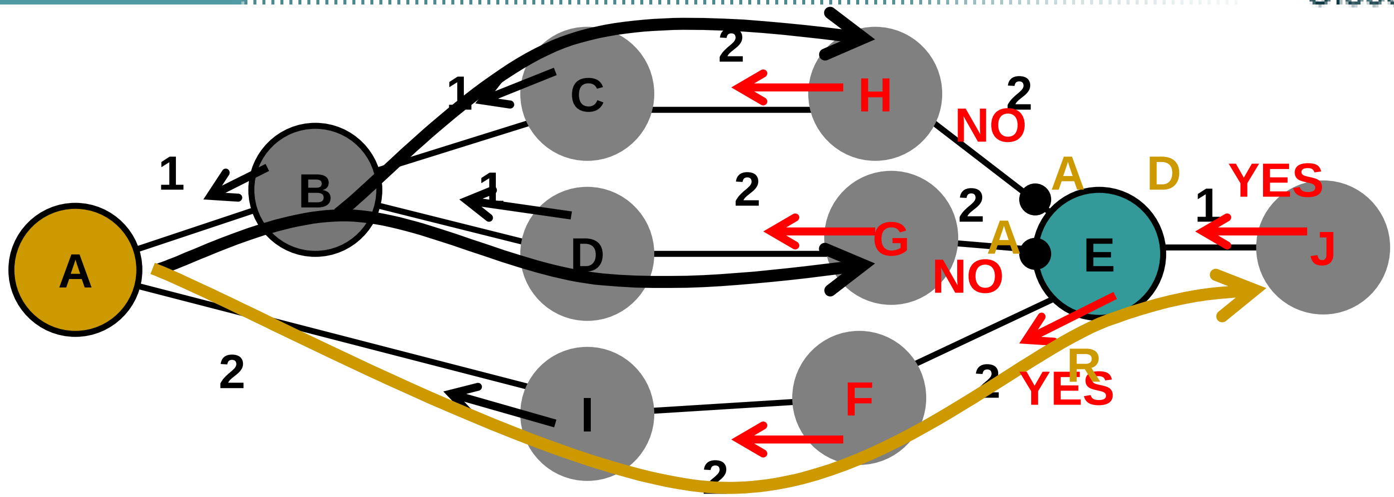
Cisco.com



- **Bridge E's (symmetrical) Dijkstra calculation of the path to Bridge A identifies the "Root Port" for the spanning tree rooted at Bridge A.**

Fast topology changes

Cisco.com



- **In order to make fast changes to the topology, Bridge E also needs to know whether the other ports are Alternate or Designated.**

  Doing a Dijkstra calculation for each of my neighbors would tell me.

  They could tell me, themselves.

Topology Changes

- But what about temporary loops during topology changes? Ethernet frames have no TTL field.
- Add interlocks between bridges, similar to MSTP, so that frames are discarded until the topology update is complete.
- These interlocks are invoked only when necessary to prevent loops, and usually are not needed.

Control Plane Issues

Control plane issues

Cisco.com

- **Two kinds of information within a cloud:**

  Information that changes when there is a topology change within the cloud.

  Information that is not tied to the cloud’s interconnect topology, and perhaps, has different timescales for changes.

Control plane issues

- **Tied to cloud topology:**

  What is the route from Bridge A to Bridge B?

  On what links should I forward this multicast or broadcast frame?

- **Tied to other things:**

  Which VLANs (Communities) are configured on this Bridge's access ports?

  Which IP multicast addresses are being requested via IGMP on this access port?

  Which stations are registered on this WAP?

Control plane issues

Cisco.com

- Use link state protocol (IS-IS?) to distribute the network topology, and to create the spanning trees.
- For information not tied directly to network topology, use a protocol suited to the expected rate of change of the information, requirements for getting that information right, etc.

Control plane issues

- When changes to one kind of information are distributed, every Bridge can calculate the changes required to its databases.
- There is no need to run a GMRP/MVRP type of protocol after the topology changes.
- This speeds convergence of multicast and VLAN distribution significantly.

Control plane issues

Cisco.com

- For example, the list of configured VLANs changes very slowly.
- The list of multicast MAC addresses has some interesting needs regarding distributing it quickly but unreliably vs. slowly and reliably.
- The unicast addresses of 802.1X and/or WAP clients have needs connected with mobility.
- One size probably does not fit all.

Plug and Play

- **As described, each bridge must be assigned one or more VIDs to identify the spanning tree instances it "owns".**
- **This can be done by configuration, and an MSTP-like "configuration name" can be used to ensure old STP connectivity while changes are made.**

But, this is a significant interruption in normal services whenever a new bridge is added.

A better scheme is needed.

Why the VID assignment problem?

- **The TRILL solution doesn’t need this VID assignment to work. Why do Bridges?**
- **In multicasts, the VID serves as the Source address in the {S,G} routing pair. This enables Bridges to do {S,G} routing with existing implementations.**
- **In unicasts and multicasts, the VID tells which frames to discard during topology changes, since there is no TTL.**

Cisco Systems