Shortest Path Bridging

An Update on Bridging Technologies

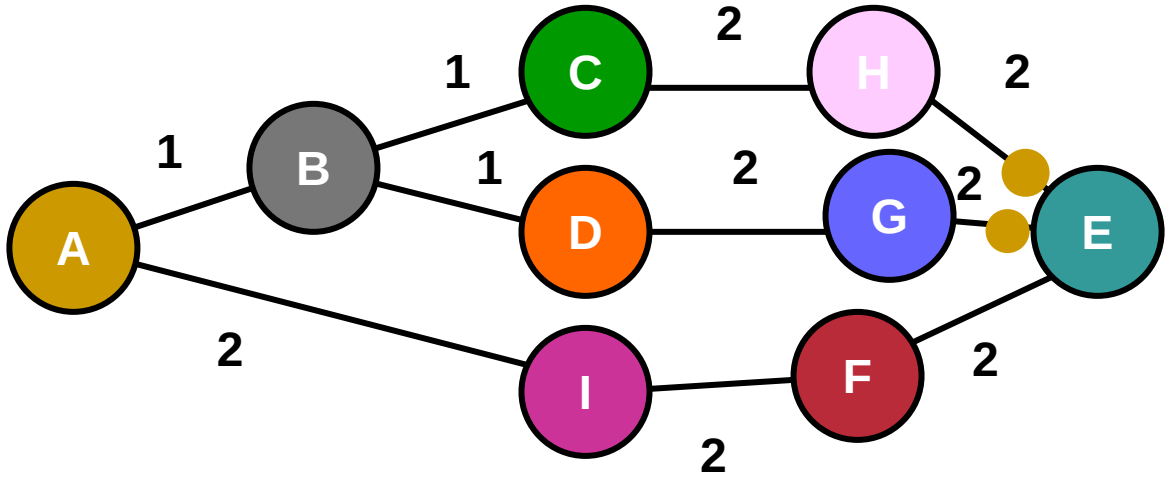
Norman Finn

IEEE Tutorial, July 18, 2005

Caveat

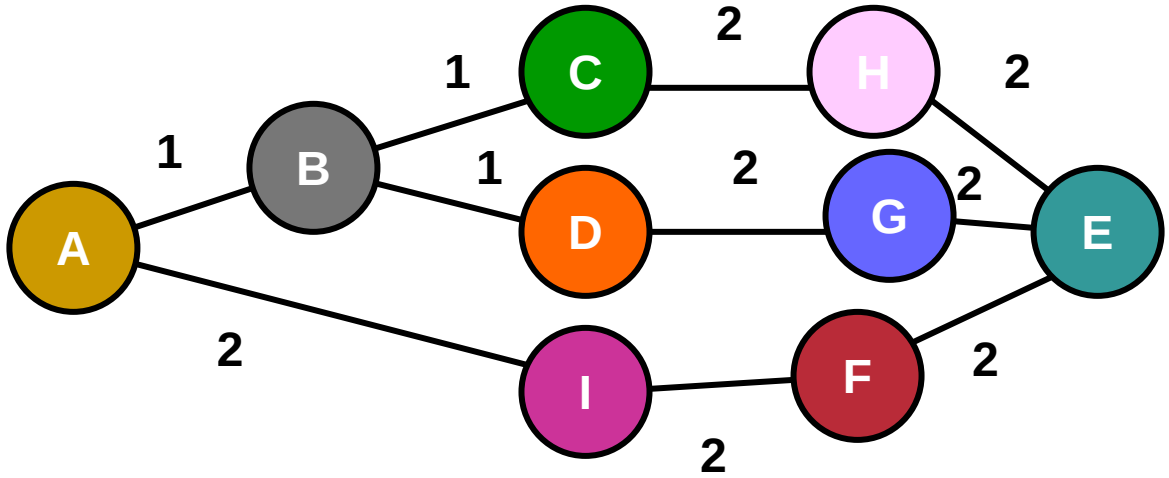
- **802.1 has not spent a significant amount of time working on P802.1ao Shortest Path Bridging.**
- **The opinions expressed in this presentation are those of the author, not Cisco, and not IEEE 802.1.**

Spanning Tree Uses Sub-Optimal Paths



- **Bridge A is the Root Bridge.**
- **Bridge E breaks the two spanning tree loops by blocking the marked ports.** ●
- **Path from E to G is E-F-I-A-B-D-G.**
- **This clearly qualifies as "sub-optimal."**

Spanning Tree Per Bridge

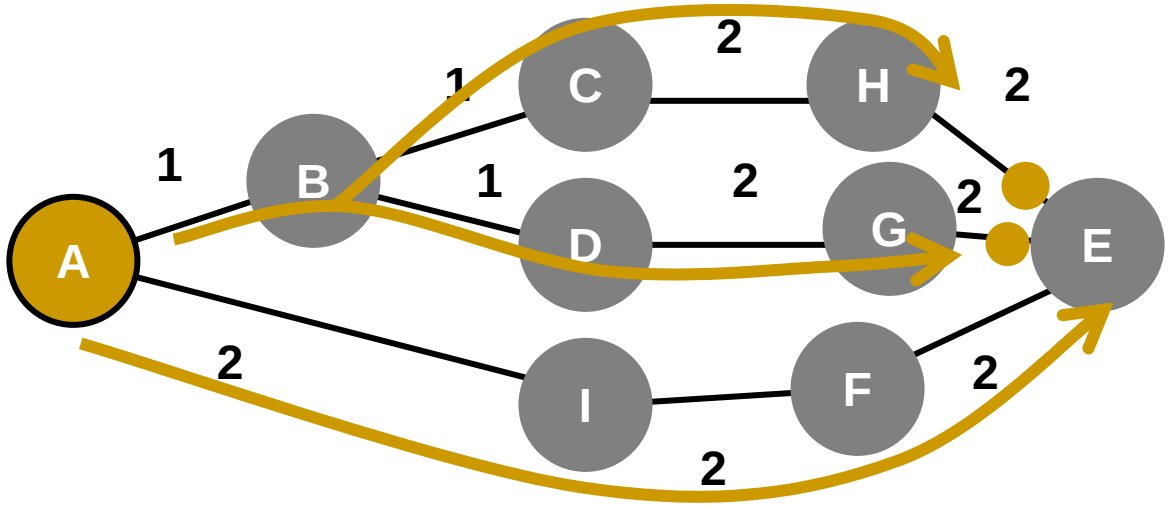


- **Instead of 1 spanning tree, we create 9 spanning trees.**

  Each bridge is the root of its own spanning tree instance (MSTI).

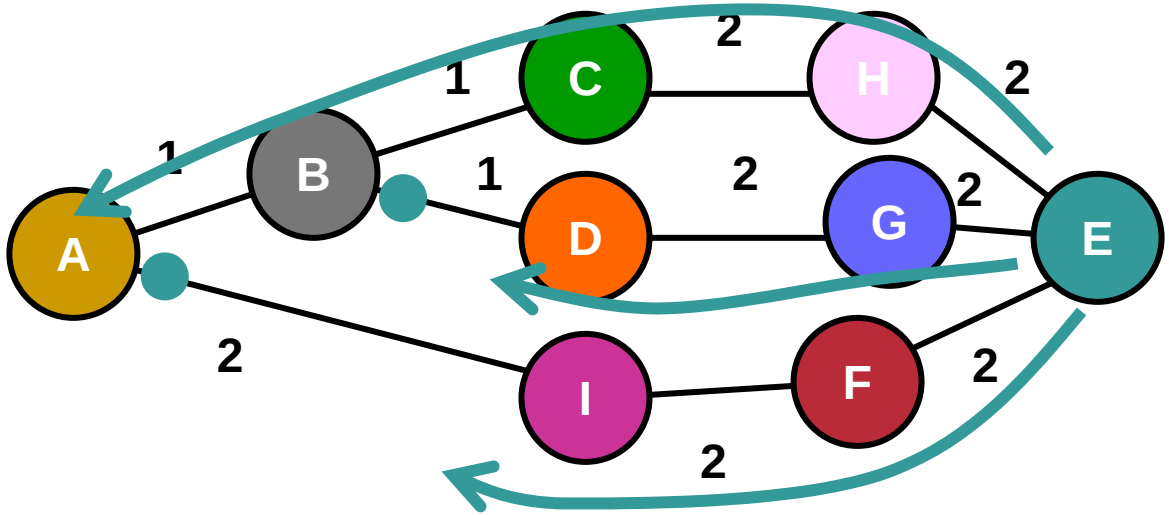
  802.1S (MSTP) supports 64 MSTIs already, 4k with some effort.

Spanning Tree Per Bridge



- **Whenever Bridge A sends a frame, it uses MSTI A.**
- **Of course, the MSTI with A as the root is the optimal path away from A.**

Spanning Tree Per Bridge

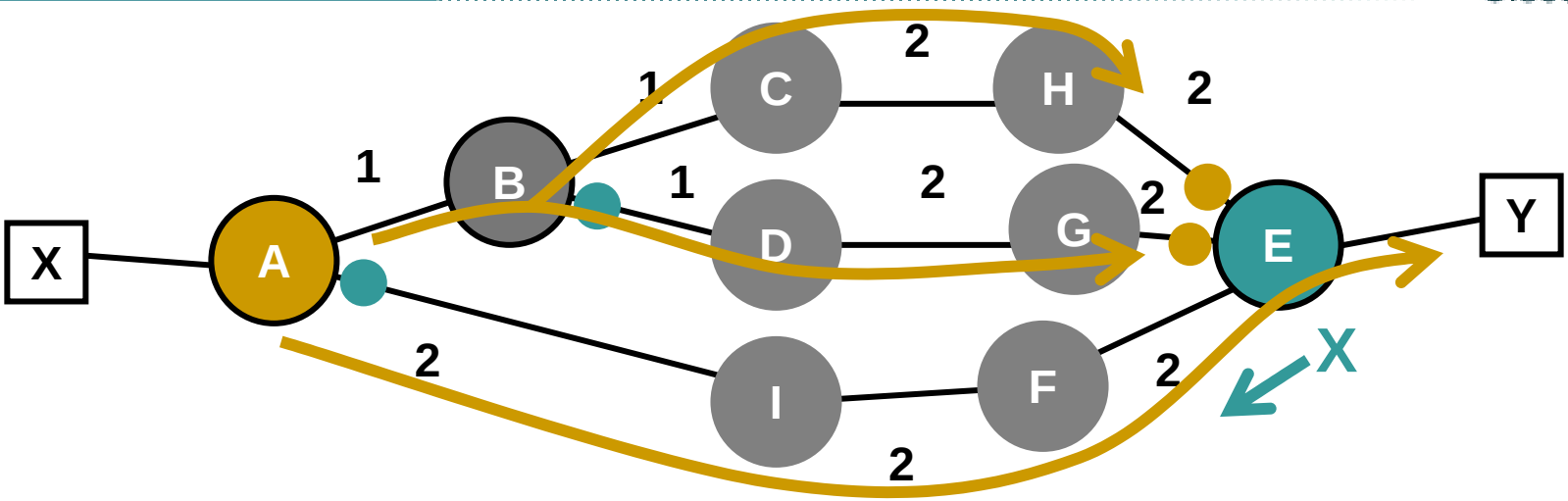


- **Similarly, traffic originating from Bridge E uses MSTI E and thus takes the optimal path from E.**

Problem solved??

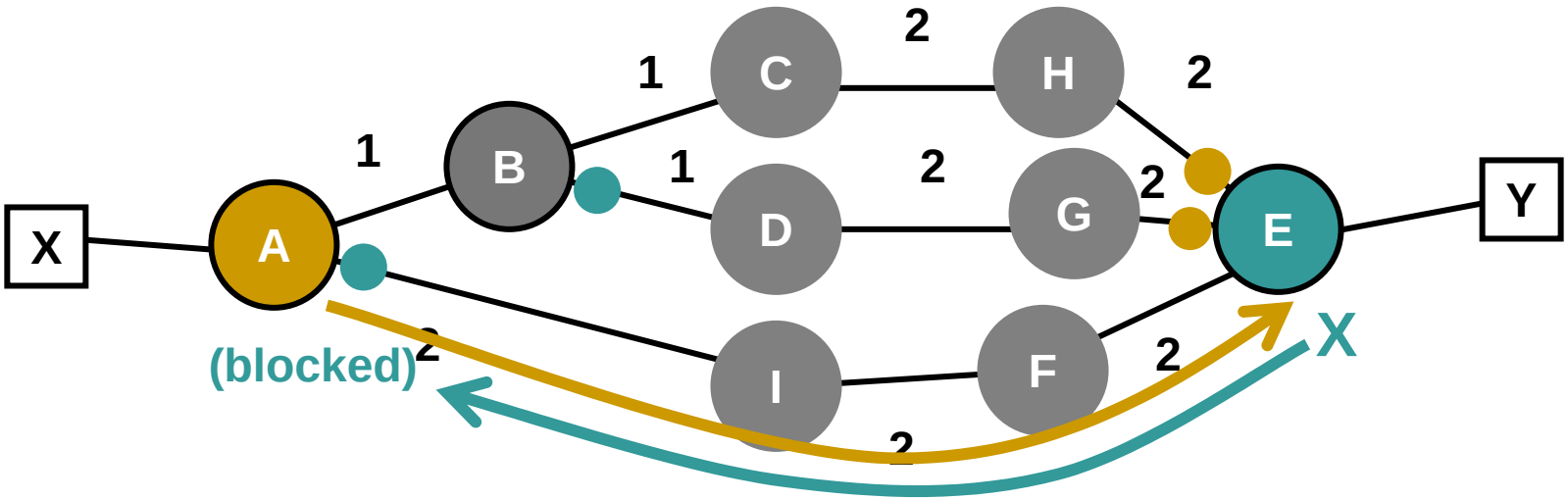
- So, every frame takes the optimal path through the network.
- Problem solved?
- Almost.

Asymmetrical Spanning Trees



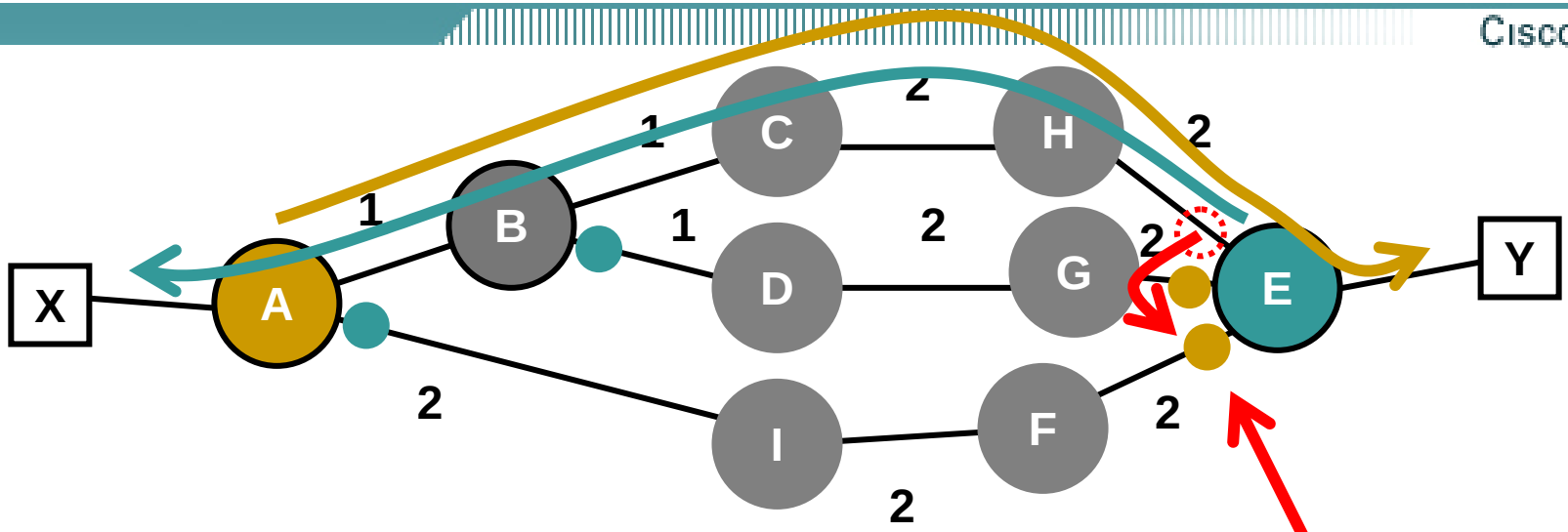
- **Station X sends a frame to Y. No bridges know where Y is, so all flood over MSTI A.**
- **All bridges learn where X is.**

Asymmetrical Spanning Trees



- **Y replies to X on MSTI E.**
- **Frame cannot get from I to A because MSTI E is blocked.**

Asymmetrical Spanning Trees



- **But, if the two spanning trees are symmetrical, then A-B-C-H-E is used for both directions, and the learned addresses work just fine.**
- **But how do we accomplish this?**

Part 1: Symmetrical Port Path Costs

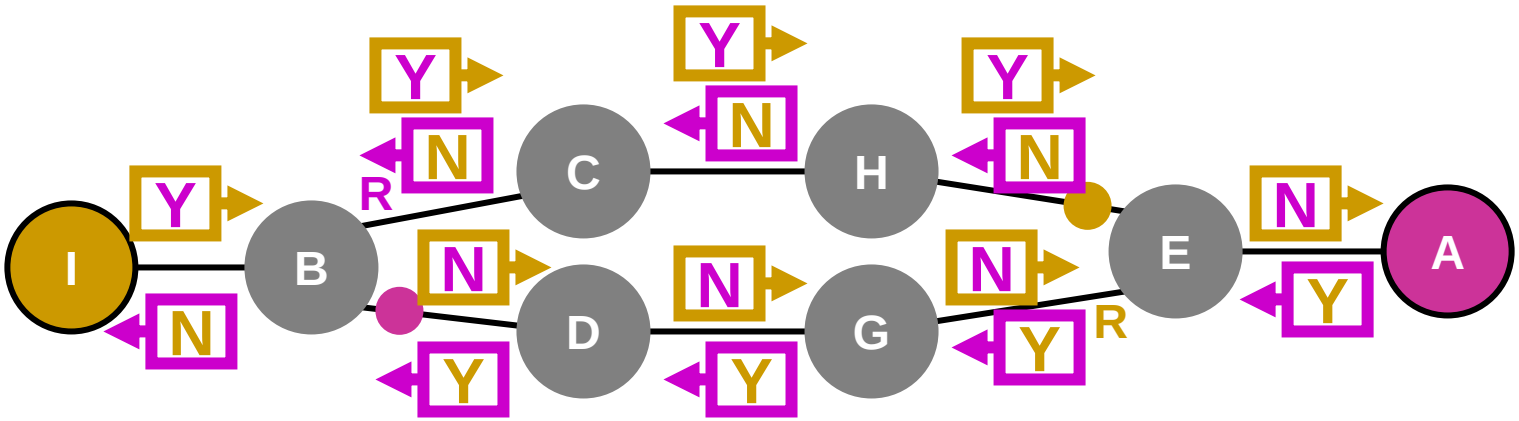
Cisco.com

- **The bridge advertises its link configured costs in BPDUs.**
- **All bridges on a given LAN use the link costs advertised by the CSTI Designated Bridge.**
- **Also, bridge's bridge priority must be the same in all STIs that must be symmetrical.**

Part 2: Reflection Vector

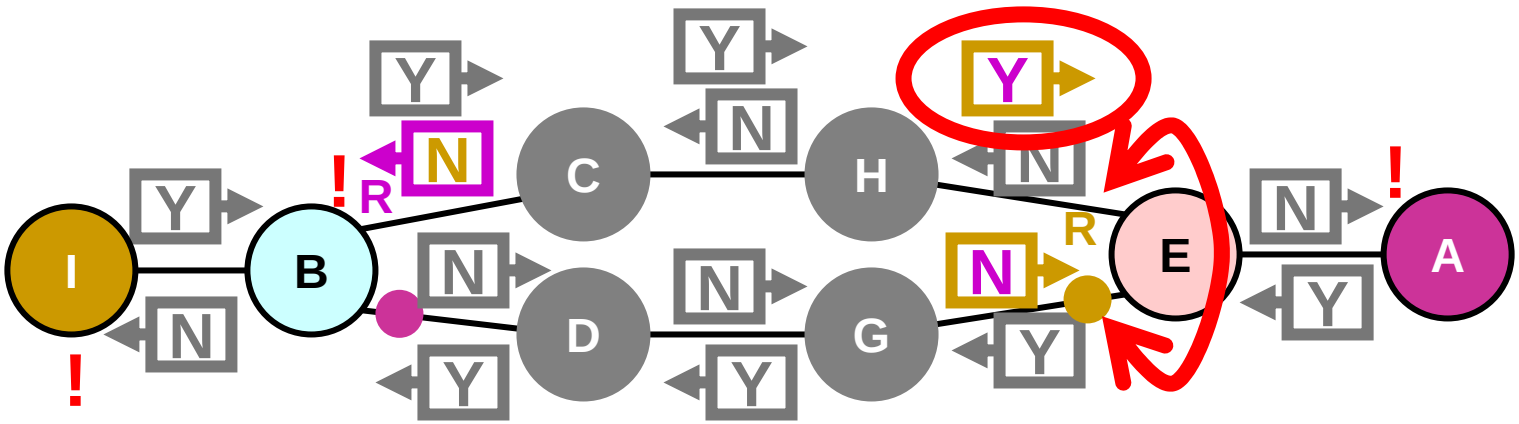
- For each MSTI *x*, BPDU carries a vector relating that MSTI to each of the other MSTIs *y*.
- For *n* MSTIs, that takes n^2 bits.
- Each bit says: "Along the path from the Regional Root of MSTI *x* to this port, every port on which this MSTI *x* information was transmitted was a Regional Root Port for MSTI *y*."

The Reflection Vector



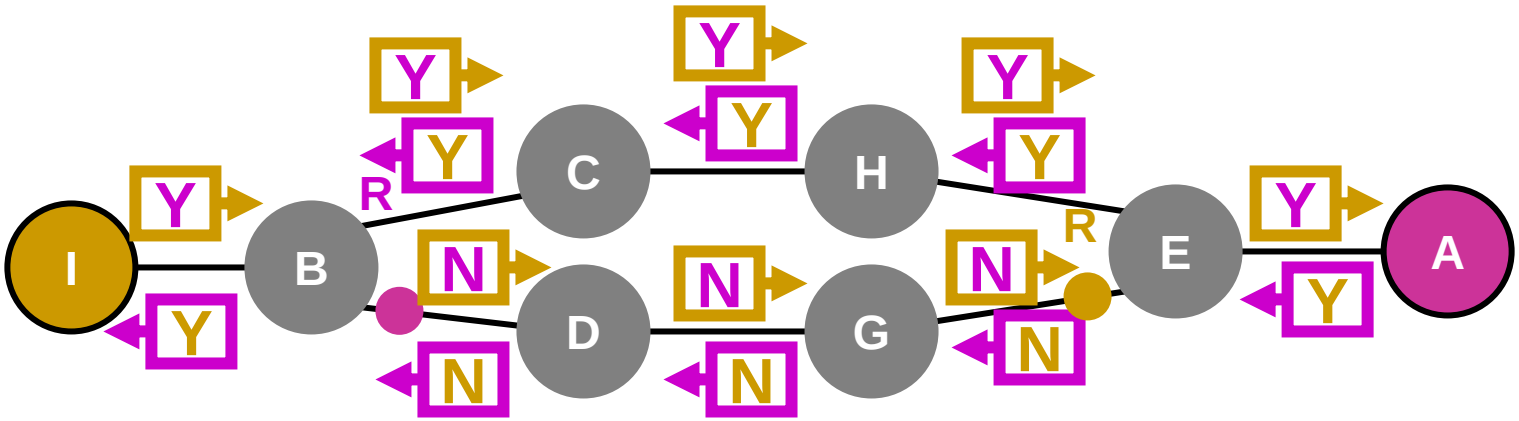
- **The box is the color of the BPDU, the letter inside shows what the vector says about the other Root's MSTI.**

The Reflection Vector



- **Bridge E** sees from the Reflection Vector that it must fix the problem. It selects a new Root Port from among its Alternate Ports.

The Reflection Vector



- **Bridge E advertises its decision.**
- **Now everything is OK and everybody knows it.**

Adding back VLANs

- **You need bits from the VLAN tag to mark frames with which MSTI they're using.**
- **VLAN tag are already used for identifying broadcast domains (the current use for VLANs).**
- **There are only 12 bits.**

  That is often enough.

  When not, P802.1ah makes more bits available.

Summary so far

Cisco.com

- **The preceding was an attempt at an existence proof that spanning tree technology can:**

  Provide the optimum routes normally associated with routing protocols;

  Provide the same network characteristics, e.g. in-order deliver, no multiple deliveries, etc., as current STP-based networks; and

  Preserve the data plane forwarding hardware currently-deployed on STP-based bridges.

Where might this go?
Some commonly held beliefs:

Cisco.com

GOOD

IS-IS or OSPF

Where might this go?
Some commonly held beliefs:

Cisco.com

GOOD

IS-IS or OSPF

RSTP

BAD

Where might this go?
Some commonly held beliefs:

Cisco.com

GOOD

IS-IS or OSPF

COMPLEX and BAD

But consider:

Cisco.com

Distance Vector	Link State
MSTP++	IS-IS or OSPF

- **MSTP augmented with Reflection Vector**

But consider:

Temporary loops (TTL needed)

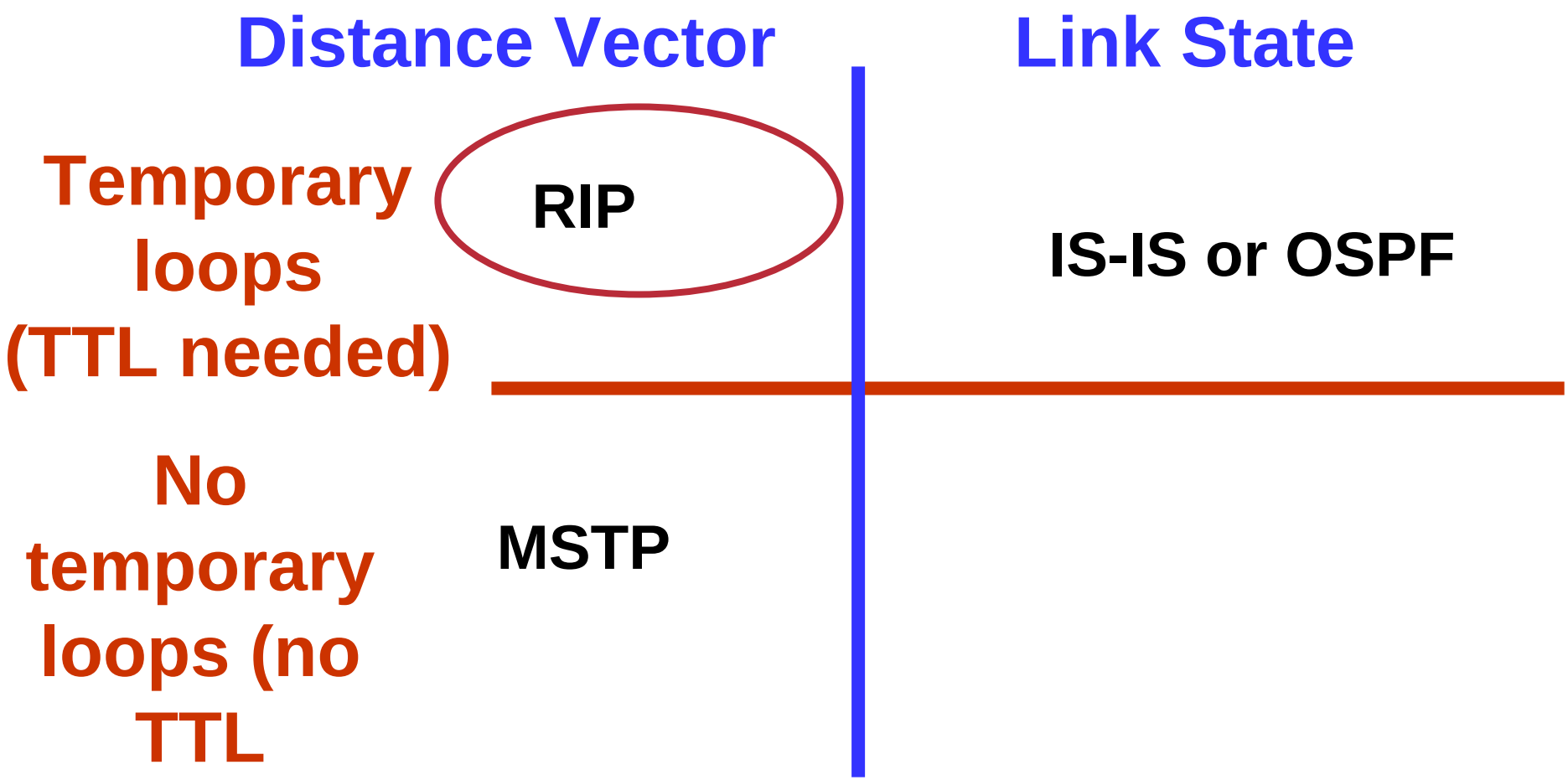
IS-IS or OSPF

No temporary loops (no TTL)

MSTP

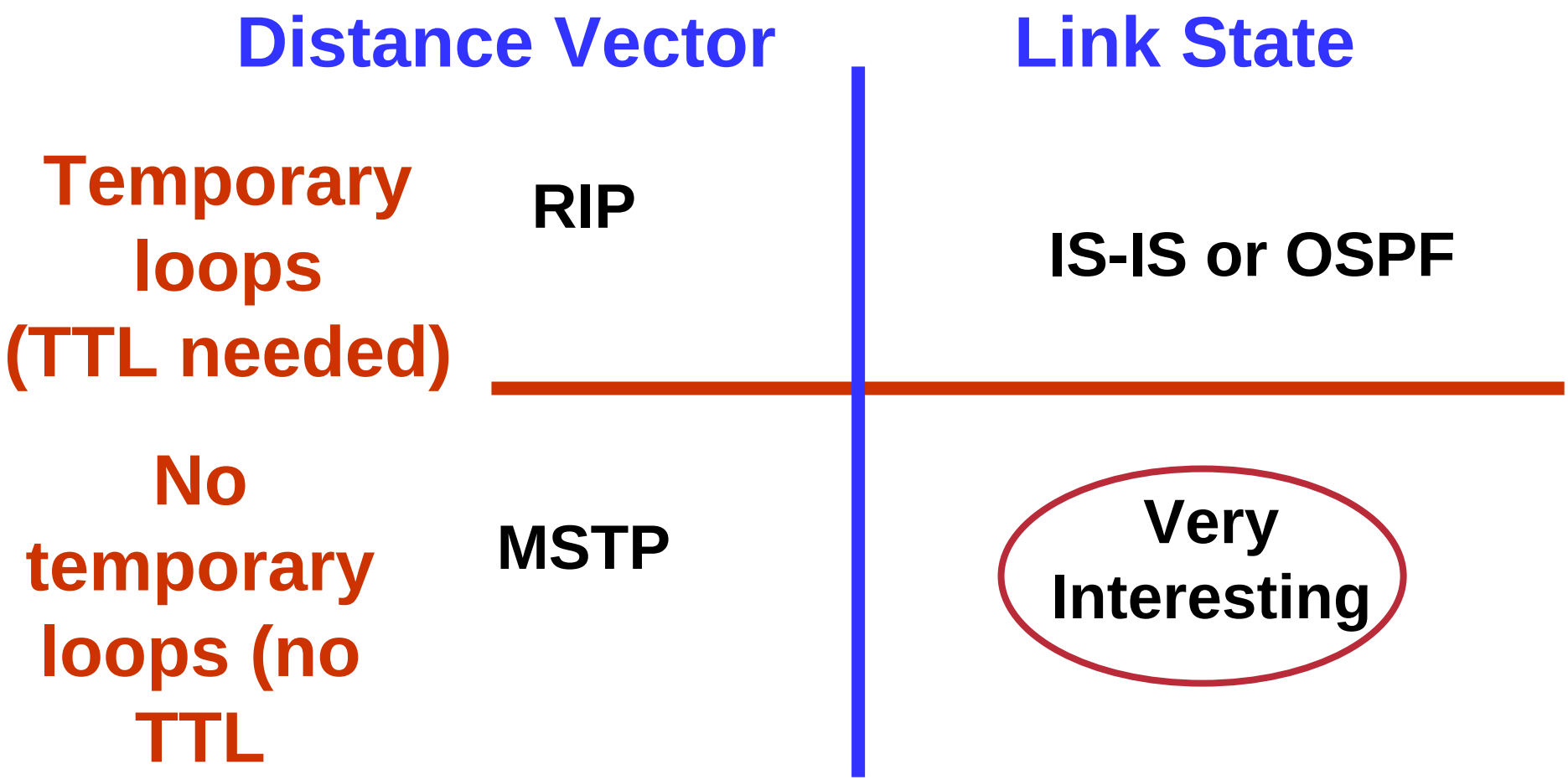
- **Another difference**

But consider:



- Combining them shows uninteresting RIP

But consider:



- **Build STIs using Link State protocol?**


CISCO SYSTEMS