MSTP issues

Multiple Spanning Tree Protocol work for the Interworking Task Group to tackle

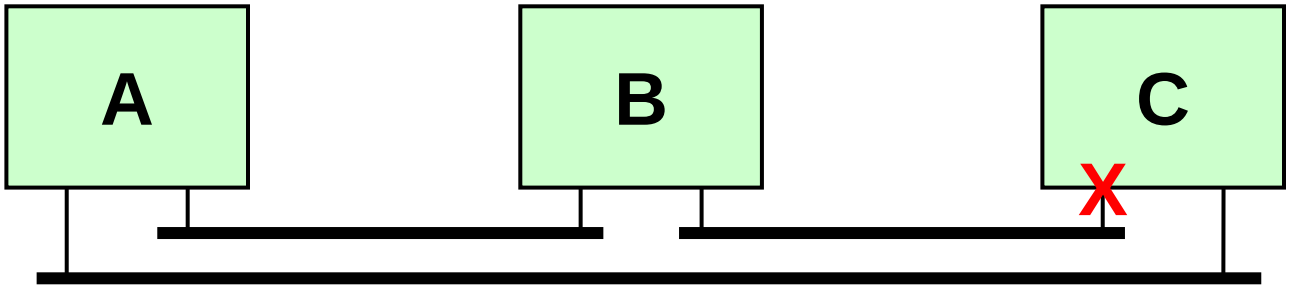
Version 1

Norman Finn

Cisco Systems

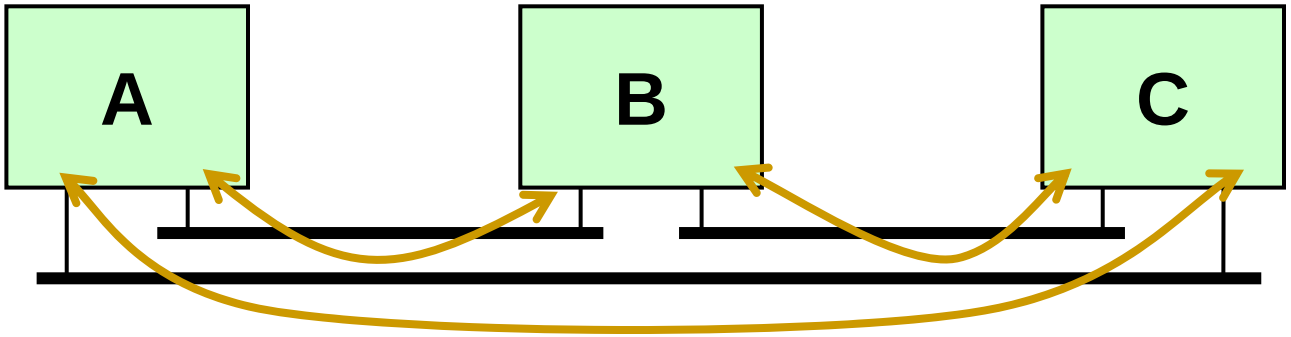
Issue 1: Brain-dead bridge

ISSUE 1: Brain-dead bridge



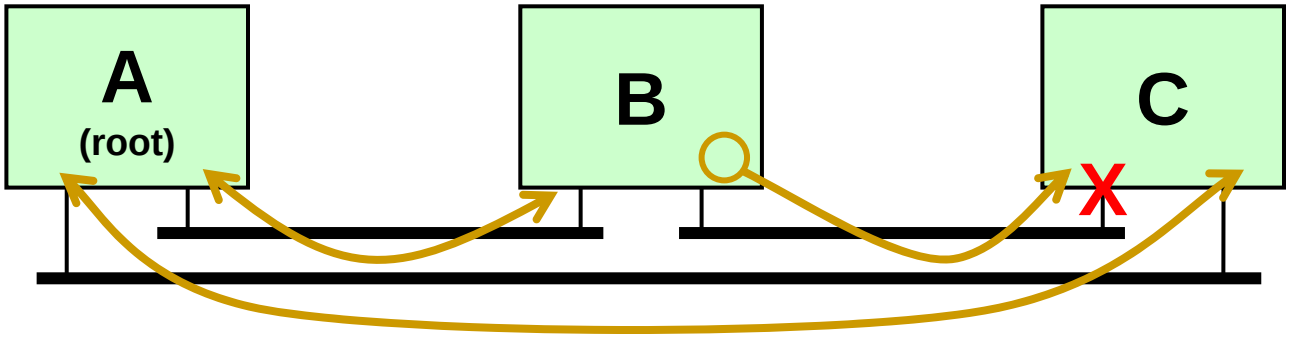
- Three bridges connected in a circle

ISSUE 1: Brain-dead bridge



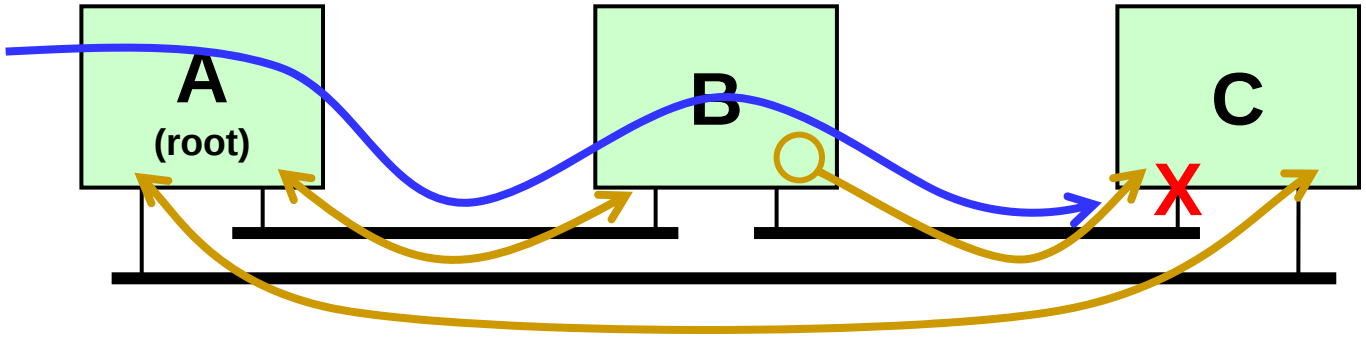
- Bridges exchange BPDUs.

ISSUE 1: Brain-dead bridge



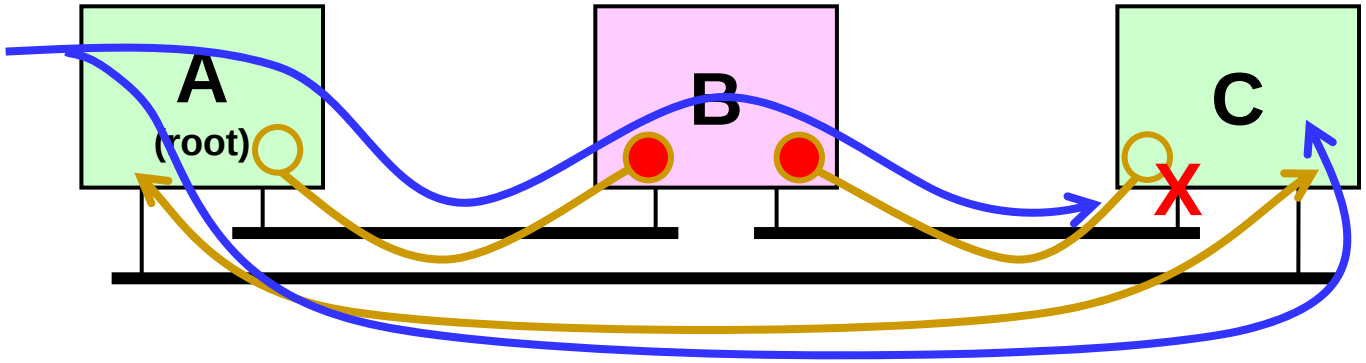
- The bridges exchange BPDUs.
- C blocks a port to prevent loops
- C stops sending BPDUs, because it is receiving superior information from B.

ISSUE 1: Brain-dead bridge



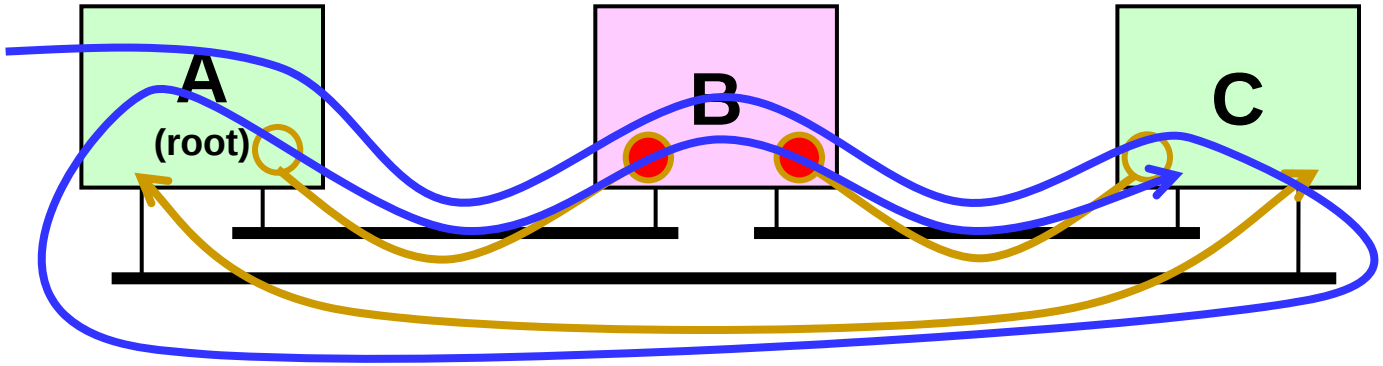
- The bridges exchange BPDUs.
- C blocks a port to prevent loops
- B still relays broadcasts to the B-C link, because bridges ensure that every **LAN** gets a broadcast, not every station. C, of course, discards them.

ISSUE 1: Brain-dead bridge



- **B goes brain dead.**
- B blocks BPDUs. (They are sent to B's brain, but not acted upon.)
- B generates no BPDUs.
- But, B's hardware still relays data frames.

ISSUE 1: Brain-dead bridge



- B does not relay BPDUs, so A and C don't know that B is relaying frames.
- So, after a while, C unblocks the port, because it is no longer receiving BPDUs.
- **Data frames loop forever.**

The basic problem

- Today, a bridge either receives BPDUs on a link, or it does not. If it is receiving BPDUs, then obviously, it is connected to another bridge, perhaps on an 802 LAN, perhaps via a BPDU-transparent device.
- If a bridge is receiving no BPDUs, it cannot distinguish whether its neighbor is:

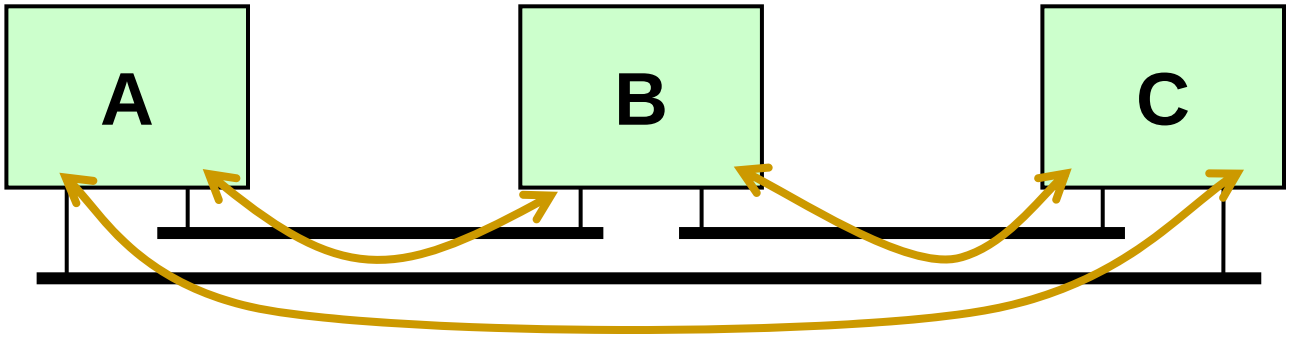
  A **brain-dead bridge**; or

  A **station** or **buffered repeater** (inexpensive STP-less bridge).
- Right now, it assumes that it is a station or buffered repeater. If it is wrong, then you get a hard loop.

Apparently simple solution

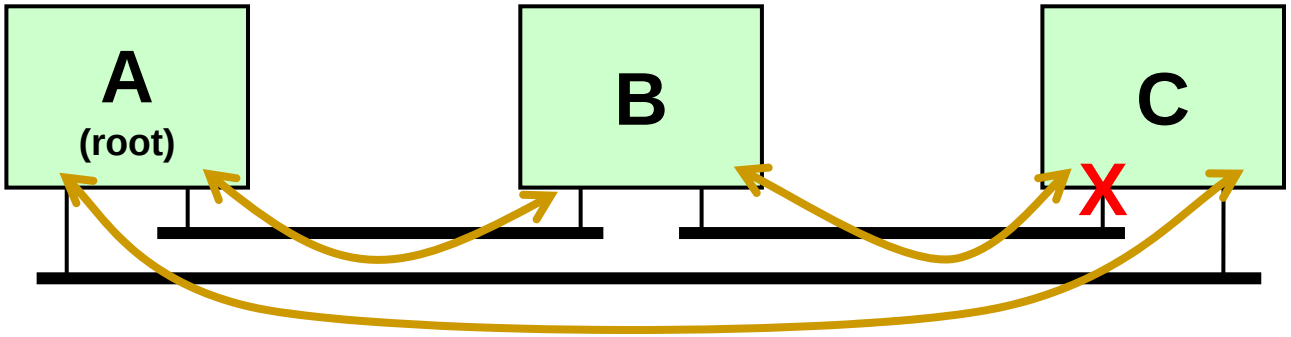
- Suppose every bridge transmits BPDUs regularly, even if it is receiving superior information.
- Suppose those BPDUs say, “I promise to keep sending BPDUs regularly, even if I am receiving superior information.”
- Then if another bridge on the link (or buffered repeater) stops receiving those BPDUs, it knows that the bridge has either disappeared from the LAN (or buffered repeater), or gone brain dead.
- The fail-safe assumption is that the bridge has gone brain dead, in which case it blocks the link.

ISSUE 1: Simple solution



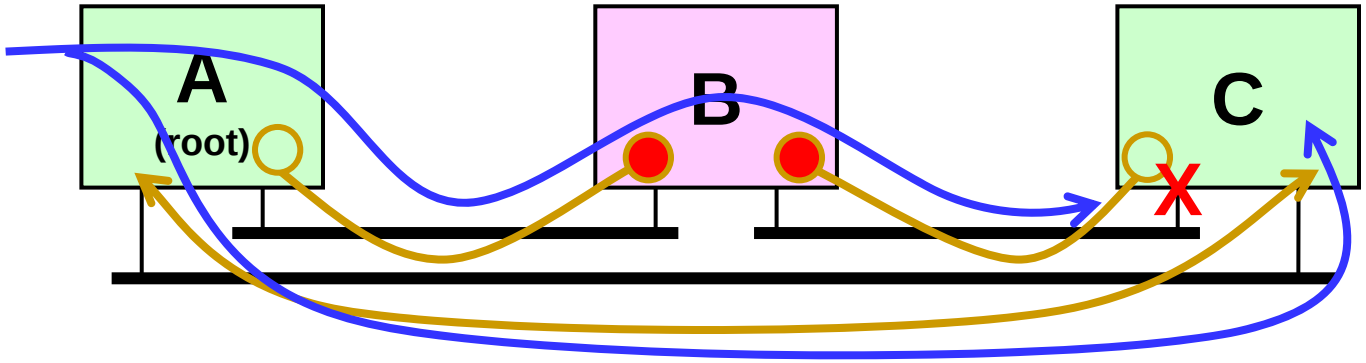
- Bridges exchange BPDUs.

ISSUE 1: Simple solution



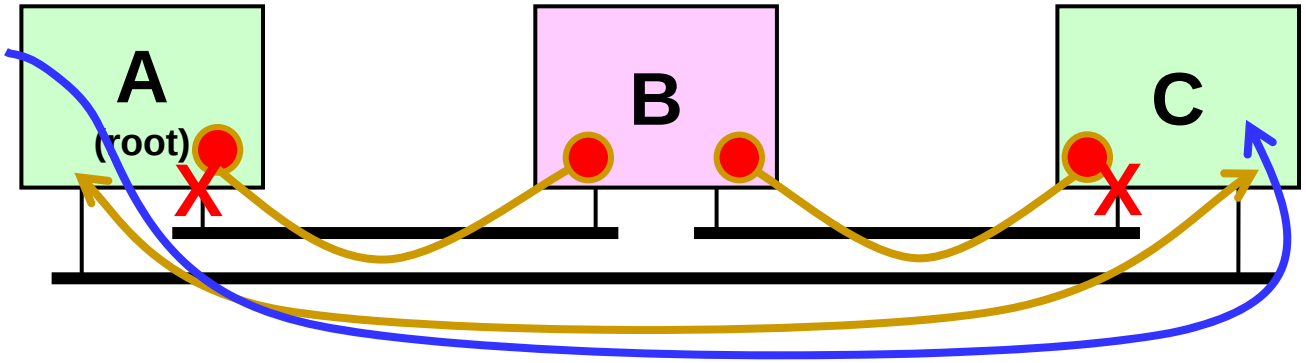
- The bridges exchange BPDUs.
- C blocks a port to prevent loops
- C continues sending BPDUs, even though it is receiving superior information from B.

ISSUE 1: Simple solution



- **B goes brain dead.**
- B blocks BPDUs. (They are sent to B's brain, but not acted upon.)
- B generates no BPDUs.
- But, B's hardware still relays data frames.

ISSUE 1: Simple solution



- A and C notice that B is brain dead, and block.
- A and C shun brain-dead B, and there are no loops.

What can go wrong, now?

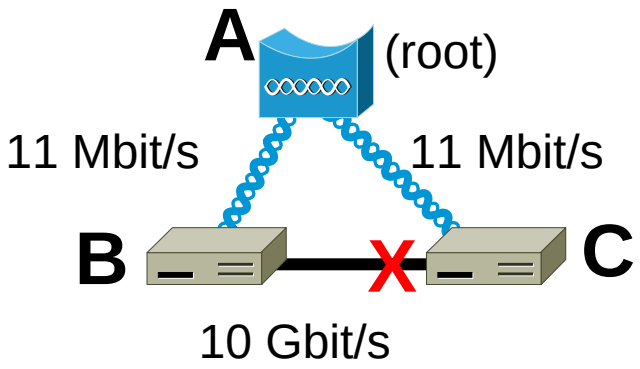
- If you are connected to a number of stations and a bridge through a buffered repeater, and disconnect the other bridge, then your fail-safe action will disconnect those stations.
  - Perhaps this is configurable per port.
- If the network powers down and back up, and that does not fix the brain-dead bridge, you will have lost the history, and you will have a hard loop, again.
  - Perhaps this is configurable per port.
  - Perhaps state is retained over a power cycle.
- Any configuration or state retention violates plug-and-play.

Summary

- The problem is **very** real.
- “Robust” and “Plug-and-play” seem to be conflicting goals, in this case.

Issue 2: Slow parallel links

ISSUE 1: AVB problem

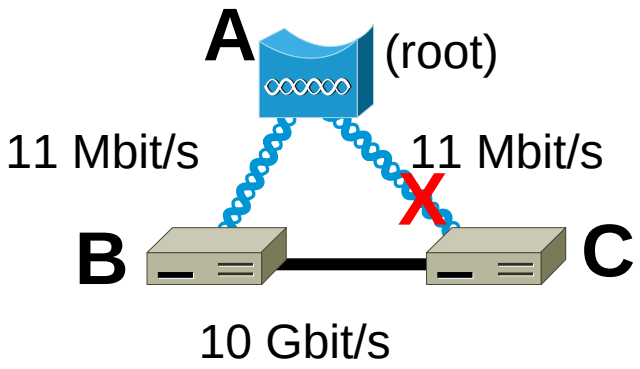


- A wireless access point is connected to two consumer electronics devices that are also **bridges**.
- C has to block a port to prevent loops.

  The cost to Root A directly is 1,000,000.

  The cost to Root A via PC B is 1,000,001.
- **So, C blocks the 10G link to B.**

ISSUE 1: Brain-dead bridge



- **In almost any situation, it would be preferable to block the wireless link.**
- Yes, traffic from “above” A will take an extra hop, but this is a very small price to pay so that B and C can communicate at full speed, instead of via the comparatively slow wireless link.

Possible solutions?

- There are a number of possible solutions to this problem, and these need to be presented.
- This part of the presentation is just a heads-up, not a PAR request!