RBridges
and the IETF TRILL Protocol

Donald E. Eastlake 3rd
d3e3e3@gmail.com, +1-508-333-2270

CONTENTS

DEFINITIONS

- TRILL – TRansparent Interconnection of Lots of Links
  - A standard specified by the IETF (Internet Engineering Task Force) TRILL Working Group co-chaired by
    - Donald E. Eastlake 3rd, Stellar Switches
    - Erik Nordmark, Sun Microsystems
- RBridge – Routing Bridge
  - A device which implements the TRILL protocol
- RBridge Campus –
  - A network of RBridges, links, and possibly intervening bridges bounded by end stations.

WHAT/WHY/WHO TRILL?

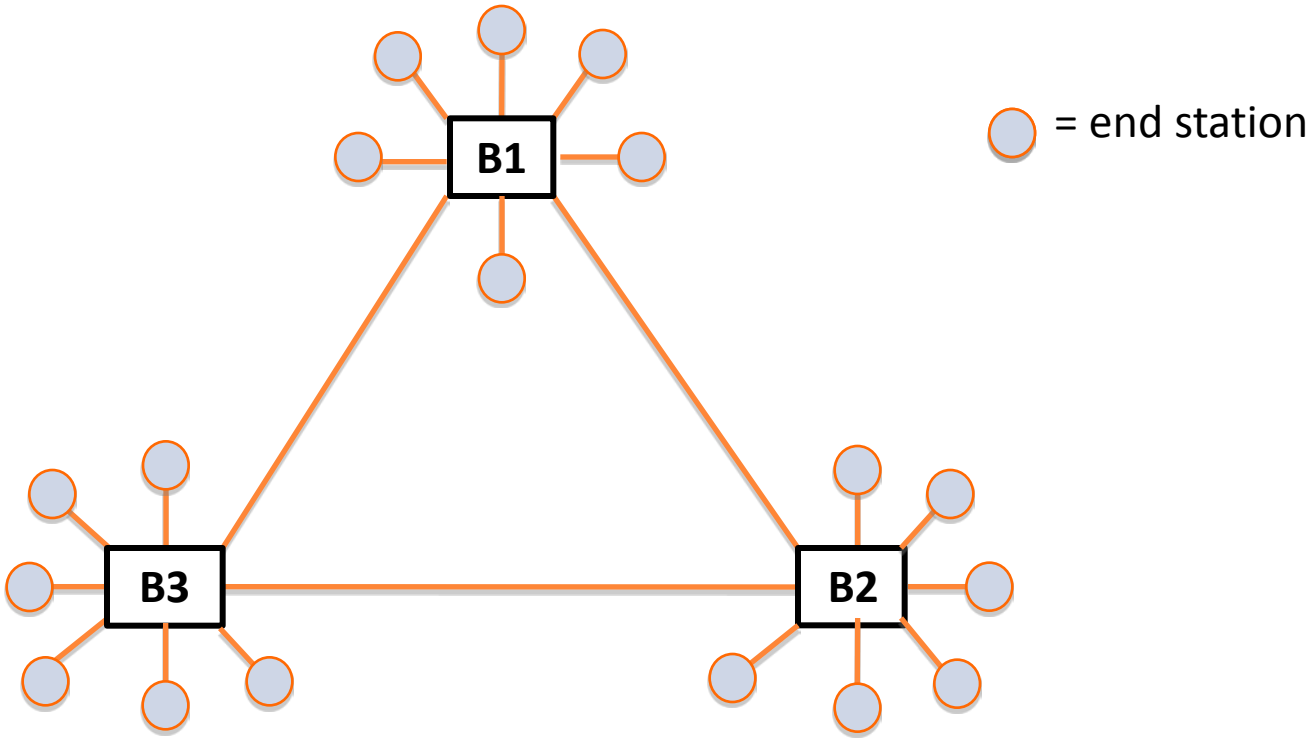
- What is TRILL?
  - TRILL is a new protocol to perform Layer 2 customer bridging with IS-IS link state routing.
- Why do TRILL?
  - Provides optimum point-to-point forwarding with zero configuration.
  - Supports multi-pathing of both unicast and multi-destination traffic.
  - Supports rapid failover.
- Who started TRILL?
  - Radia Perlman, the inventor of the Spanning Tree Protocol.

NOTE:

- This presentation is just a brief technical overview. It is not possible to include all the details in the 90+ page base protocol specification document.
- The specification is quite stable and has received substantial review. The current -15 version of the specification has been passed up from the Working Group for IETF approval on the standards track.

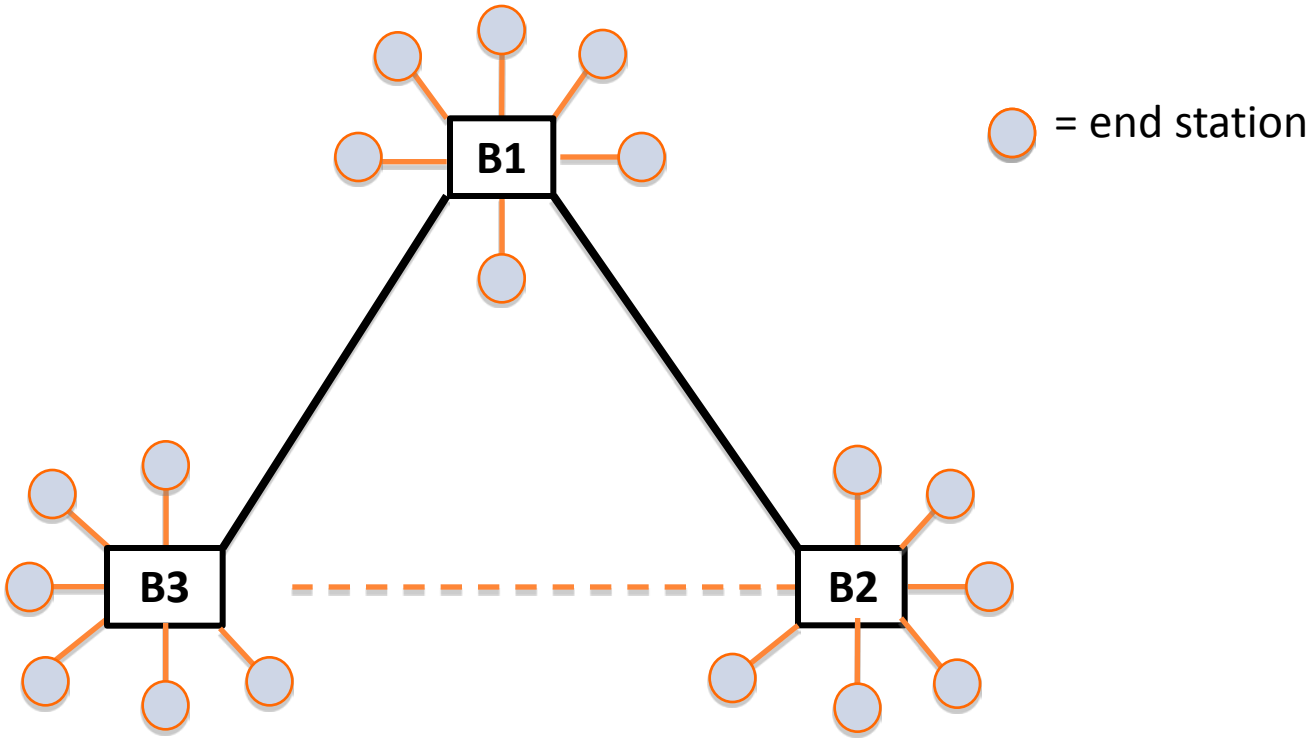
CONTENTS

OPTIMUM POINT-TO-POINT FORWARDING



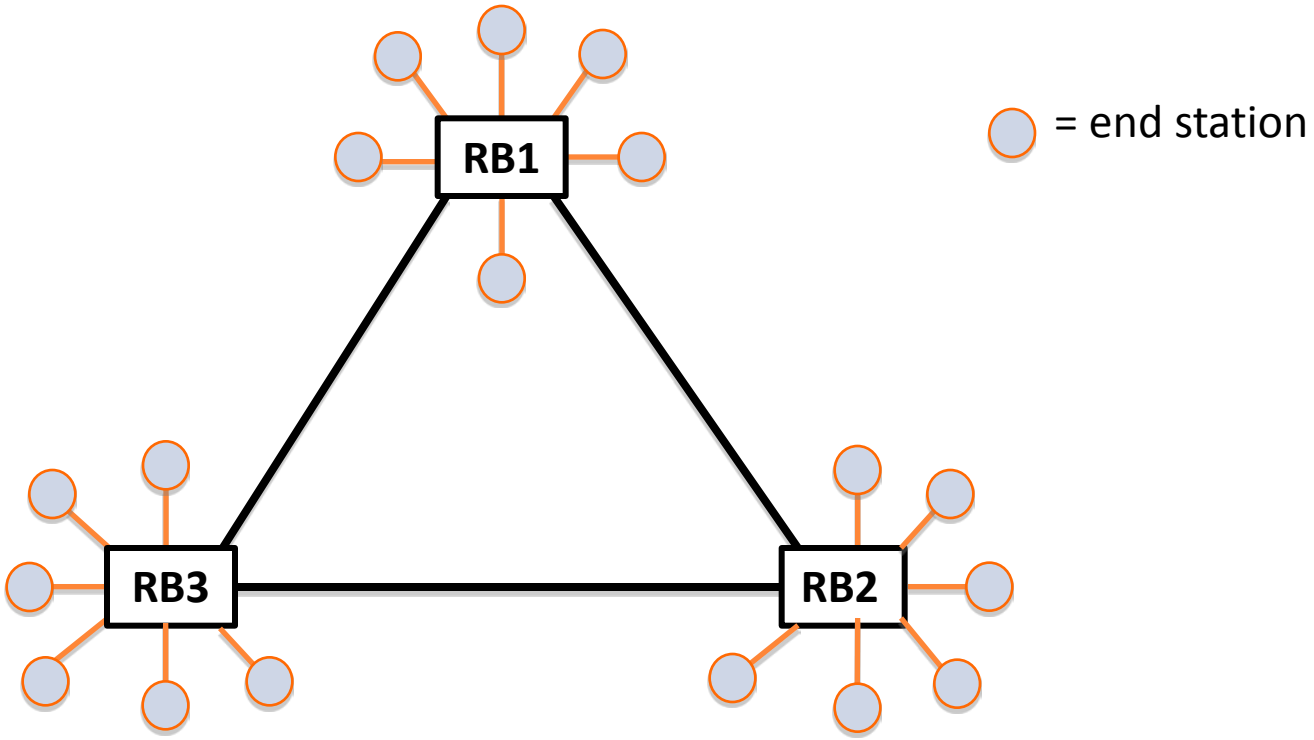
A three bridge network

OPTIMUM POINT-TO-POINT FORWARDING



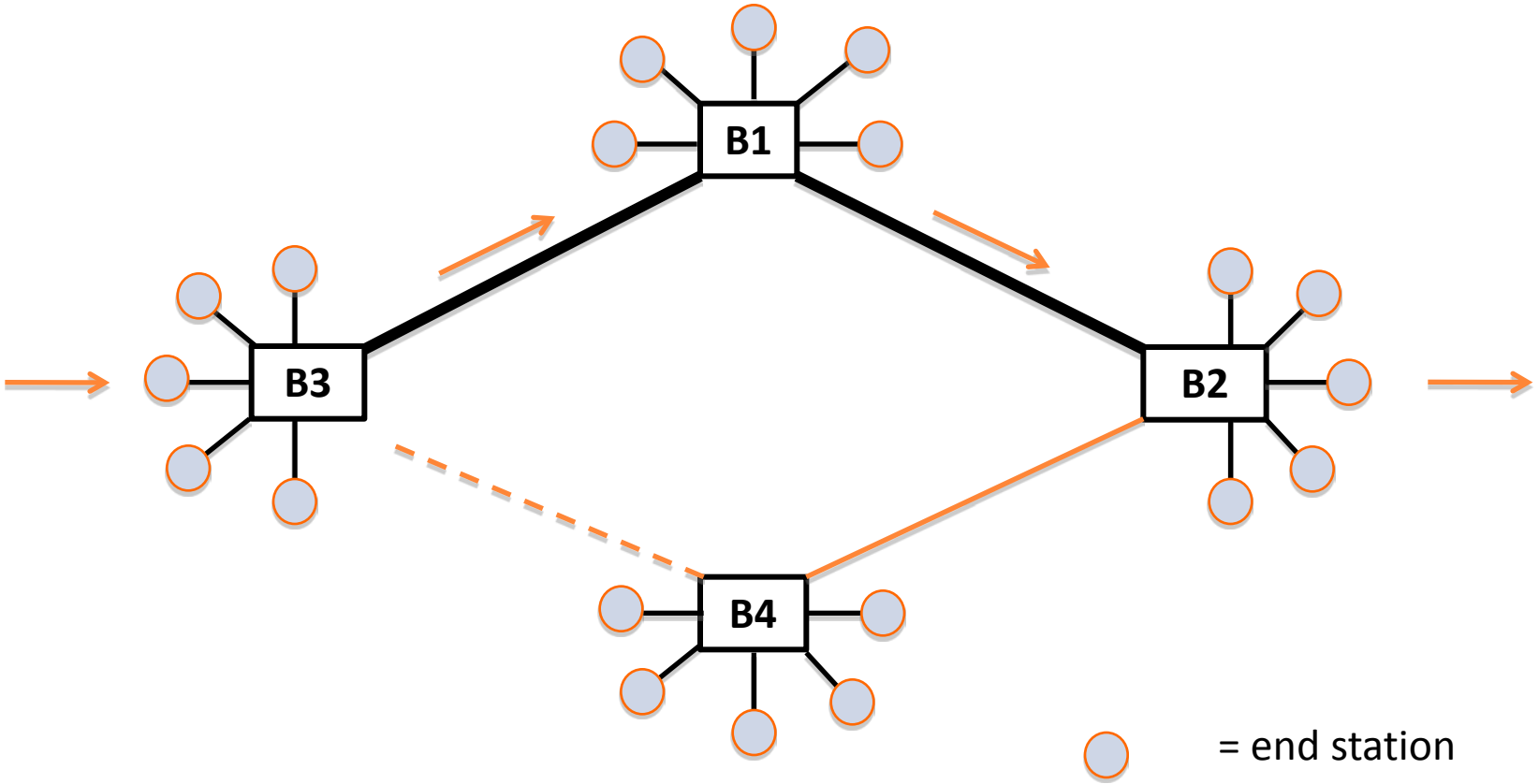
Spanning tree eliminates loops by disabling ports

OPTIMUM POINT-TO-POINT FORWARDING



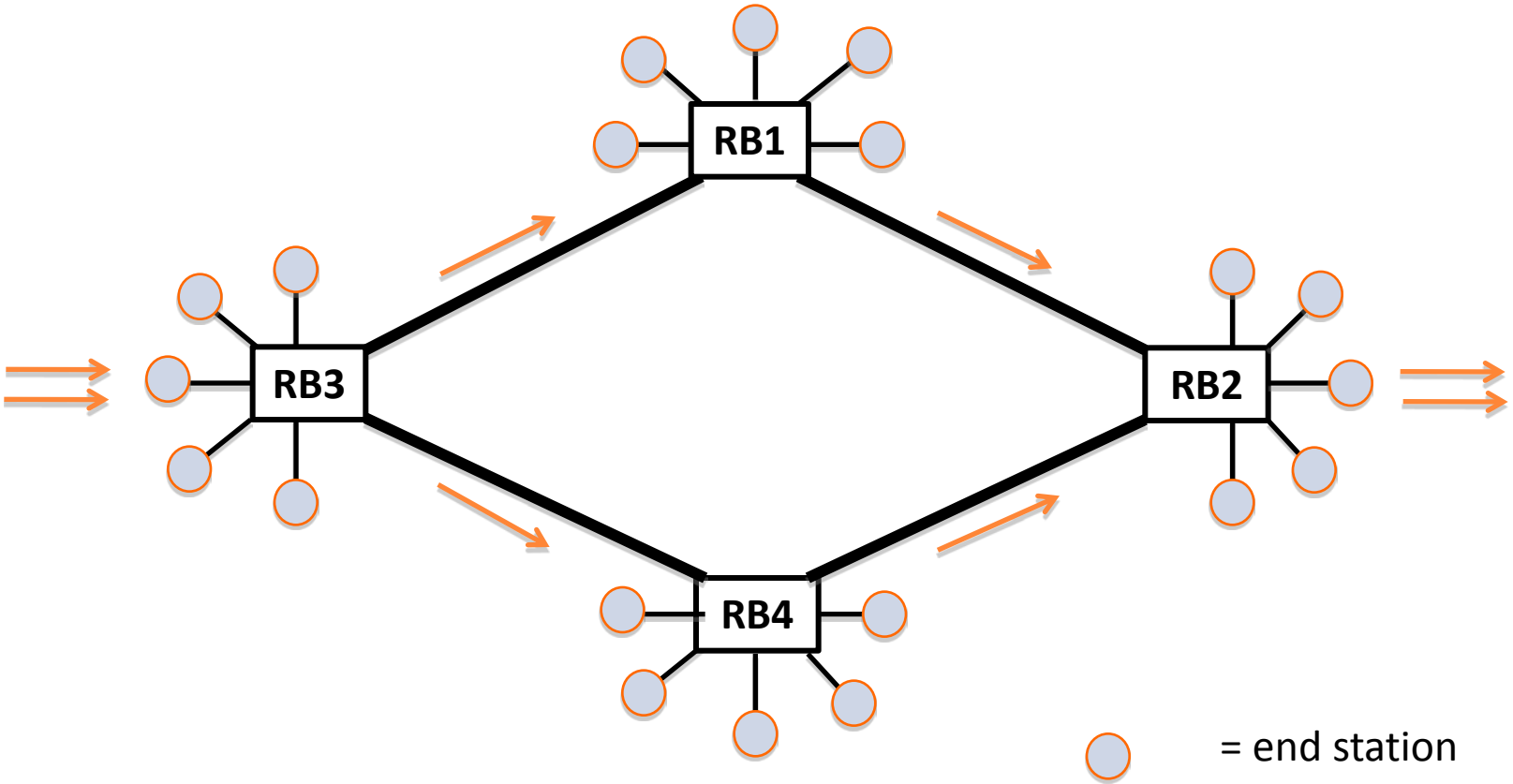
A three RBridge network: better performance using all facilities

MULTI-PATHING



Bridges limit traffic to one path

MULTI-PATHING



RBridges support
multi-path for higher throughput

Other TRILL Features

- Compatible with classic bridges. RBridges can be incrementally deployed into a bridged LAN.
- Unicast forwarding tables at transit RBridges scale with the number of RBridges, not the number of end stations. Transit RBridges do not learn end station addresses.
- A flexible options feature. RBridges know what options other RBridges support.
- Globally optimized distribution of IP derived multicast.

CONTENTS

ARE RBRIDGES BRIDGES OR ROUTERS?

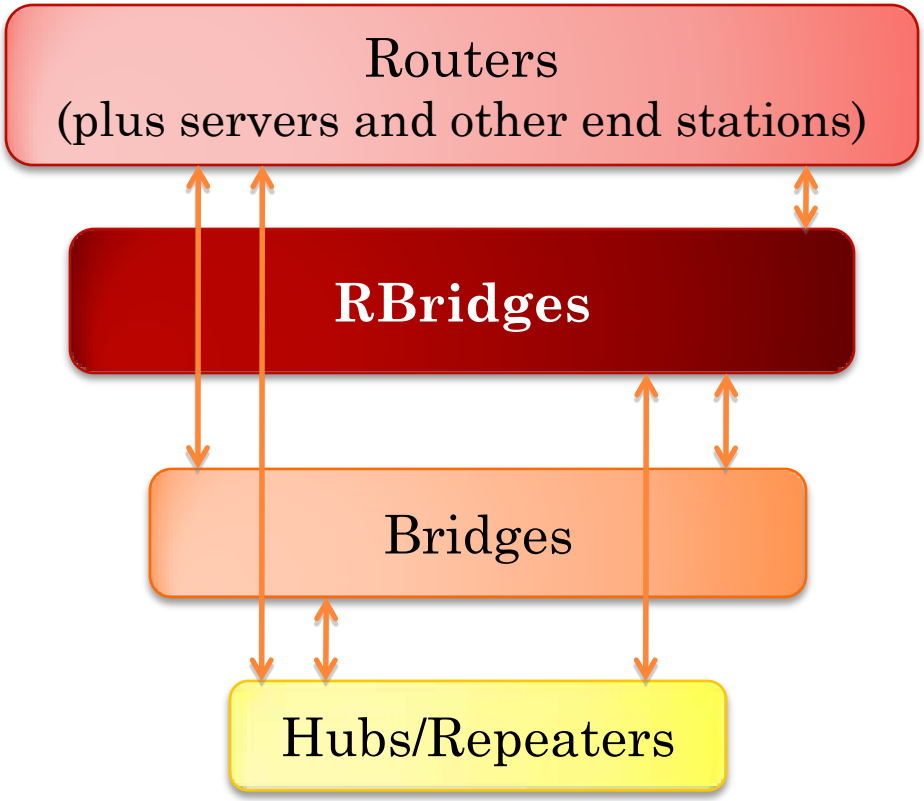
- They are obviously Bridges because
  - RBridges deliver unmodified frames from the source end station to the destination end station
  - RBridges can operate with zero configuration and auto-configure themselves
  - RBridges provide the restriction of frames to VLANs as IEEE 802.1Q-2005 bridges do
  - RBridges can support frame priorities as IEEE 802.1Q-2005 bridges do
  - RBridges, by default, learn MAC addresses from the data plane

ARE RBRIDGES BRIDGES OR ROUTERS?

- They are obviously Routers because
  - RBridges swap the outer addresses on each RBridge hop from the ingress RBridge to the egress RBridge
  - RBridges decrement a hop count in TRILL frames on each hop
  - RBridges use a routing protocol rather than the spanning tree protocol
  - RBridges optionally learn MAC addresses by distribution through the control plane
  - RBridges normally act based on IP multicast control messages (IGMP, MLD, and MRD) and restrict the distribution of IP derived multicast frames

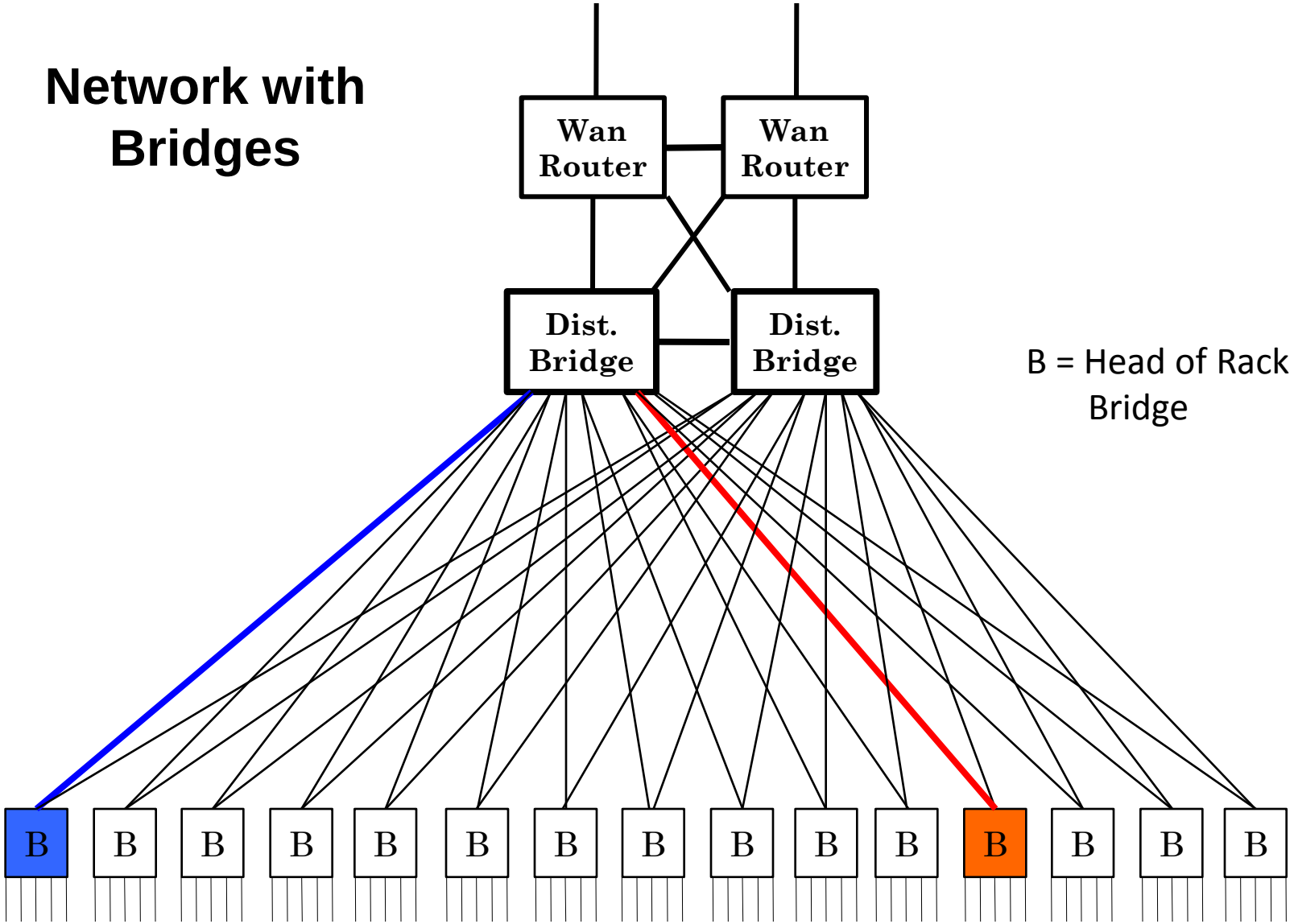
ARE RBRIDGES BRIDGES OR ROUTERS?

- Really, they are a new species, between IEEE 802.1 bridges and routers:

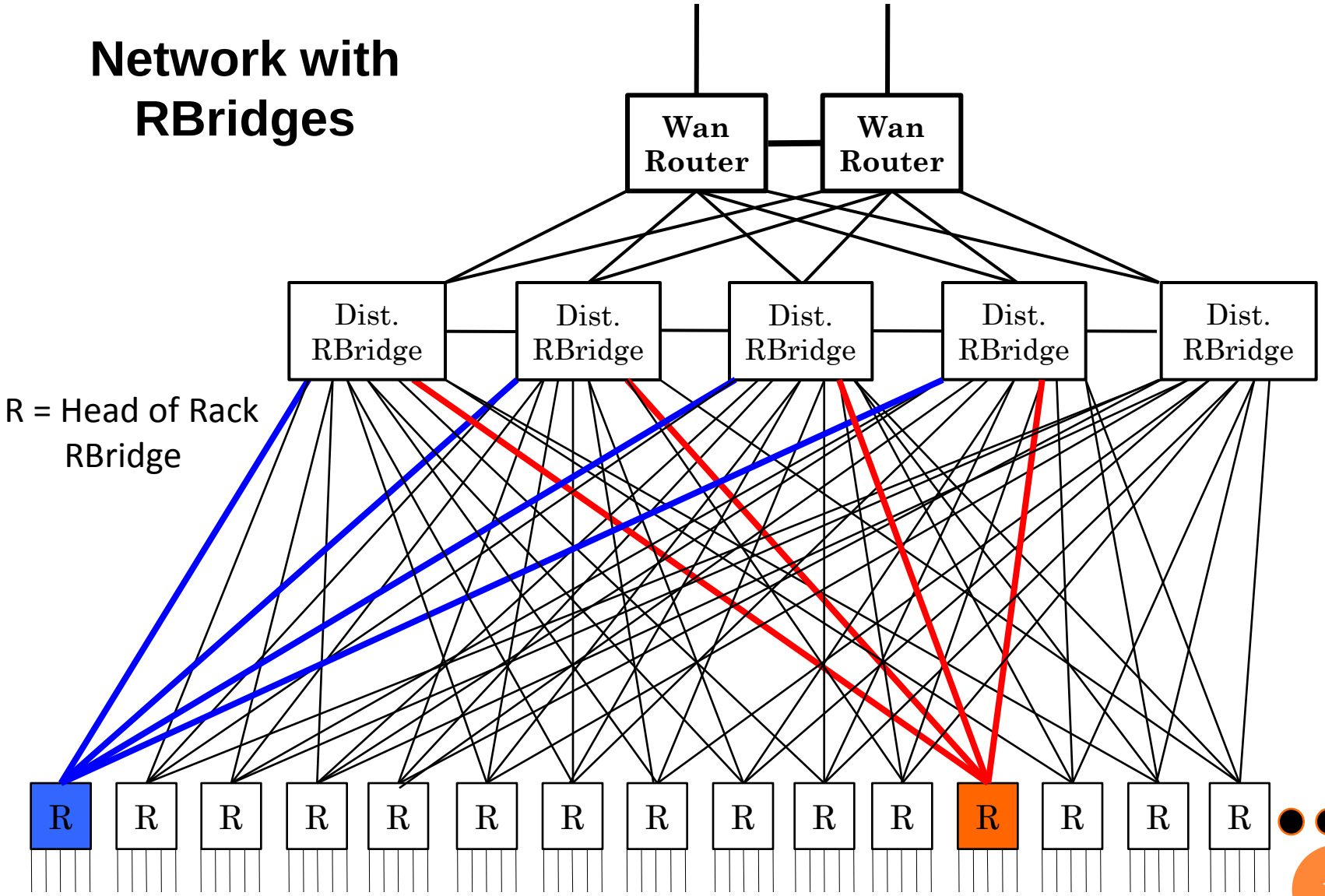


CONTENTS

Network with Bridges



Network with RBridges



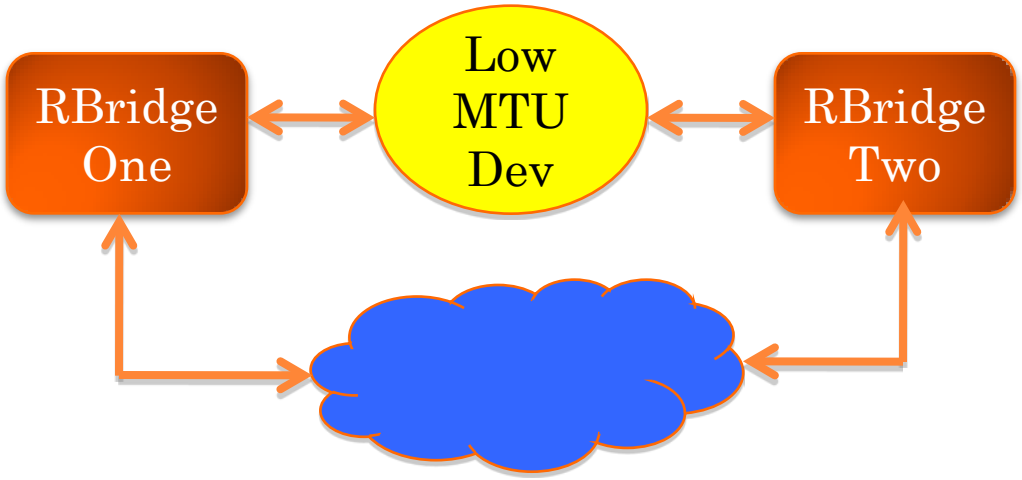
CONTENTS

HOW RBRIDGES WORK

- RBridges find each other by exchanging TRILL IS-IS Hello frames
  - Like all TRILL IS-IS control frames, TRILL Hellos are sent to the multicast address All-IS-IS-RBridges. They are transparently forwarded by bridges, dropped by end stations including routers, and are processed (but not forwarded) by RBridge ports.
  - TRILL Hellos are different from Layer 3 IS-IS LAN Hellos because they are small, unpadded, and support fragmentation of some information.
    - Separate MTU-probe and MTU-ack messages are used for MTU testing and determination.
  - Using the information exchanged in the Hellos, the RBridges on each link elect the Designated RBridge for that link (i.e., bridged LAN).

HOW RBRIDGES WORK

- TRILL Hellos are unpadded and a maximum of 1470 bytes so be sure you don't get two Designated RBridges on the same link.



HOW RBRIDGES WORK

- RBridges use the IS-IS reliable flooding protocol so that each RBridge has a copy of the global "link state" database.
  - The RBridge link state includes information beyond connectivity and link cost. Information such as VLAN connectivity, multicast listeners and multicast router attachment, claimed nickname(s), ingress-to-egress options supported, and the like.
  - The database is sufficient for each RBridge to independently and without further messages calculate optimal point-to-point paths for known unicast frames and the same distribution trees for multi-destination frames.

HOW RBRIDGES WORK

- The Designated RBridge specifies the Appointed Forwarder for each VLAN on the link (which may be itself) and the Designated VLAN for inter-RBridge communication.
- The Appointed Forwarder for VLAN-x on a link handles all native frames to/from that link in that VLAN.
  - It encapsulates frames from the link into a TRILL data frame. This is the ingress RBridge function.
  - It decapsulates native frames destined for the link from TRILL data frames. This is the egress RBridge function.

HOW RBRIDGES WORK

- TRILL Data frames with
  - known unicast ultimate destinations are forwarded RBridge hop by RBridge hop toward the egress RBridge and
  - multi-destination frames (broadcast, multicast, and unknown destination unicast) are forwarded on a tree rooted at an RBridge selected by the ingress RBridge.
    - For loop safety, a Reverse Path Forwarding Check is performed on multi-destination TRILL Data frames when received.

CONTENTS

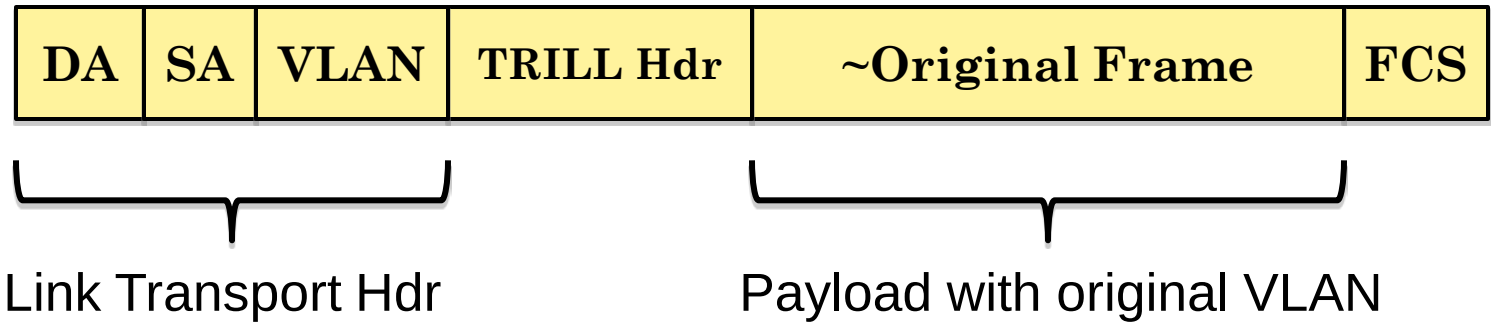
THE TRILL ENCAPSULATION AND HEADER

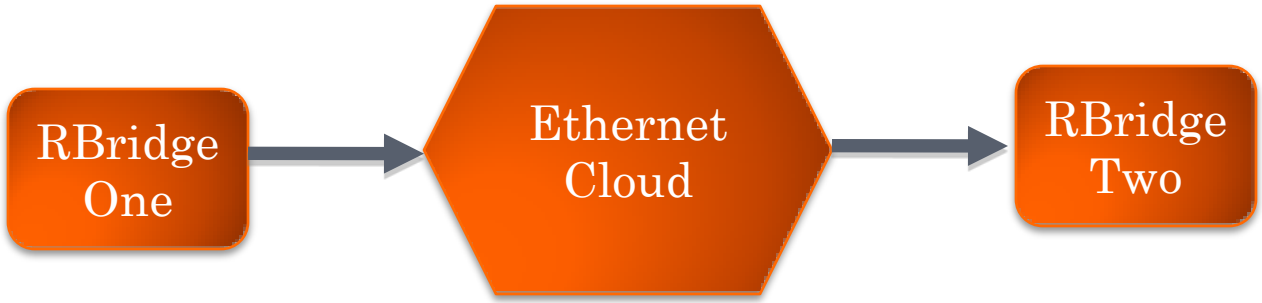
- TRILL Data frames between RBridges are encapsulated in a local link header and TRILL header.
  - The link header is addressed from the local source RBridge to the local destination RBridge for known unicast or to All-RBridges for multidestination.
  - The TRILL header gives the first/ingress RBridge and either the last/egress RBridge for known unicast or the distribution tree root for multidestination.

THE TRILL ENCAPSULATION AND HEADER

- Some reasons for encapsulation:
  - Provides a hop count to mitigate loop issues
  - To hide the original source address to avoid confusing any bridges present as might happen if multi-pathing were in use
  - To direct unicast frames toward the egress RBridge so that forwarding tables in transit RBridges need only be sized with the number of RBridges in the campus, not the number of end stations
  - To provide a separate VLAN tag for forwarding traffic between RBridges, independent of the original VLAN of the frame

THE TRILL ENCAPSULATION AND HEADER





THE TRILL ENCAPSULATION AND HEADER

- Assuming the link is Ethernet (IEEE 802.3) the encapsulation looks like:
  1. Outer Ethernet Header
     - Source RBridge One, Destination RBridge Two
  2. (Outer VLAN Tag)
  3. TRILL Header
  4. Inner Ethernet Header
     - Original Source and Destination Addresses
  5. Inner VLAN Tag
  6. Original Payload
  7. Frame Check Sequence (FCS)

THE TRILL ENCAPSULATION AND HEADER

- TRILL Header – 64 bits

TRILL Ethertype	V	R	M	OpLng	Hop
Egress RBridge Nickname	Ingress RBridge Nickname				

  - Nicknames – auto-configured 16-bit campus local names for RBridges
  - V = Version (2 bits)
  - R = Reserved (2 bits)
  - M = Multi-Destination (1 bit)
  - OpLng = Length of TRILL Options
  - Hop = Hop Limit (6 bits)

CONTENTS

ADDRESS LEARNING

- From Locally Received Native Frames
  - { VLAN, Source Address, Port }
- From Decapsulated Native Frames
  - { Inner VLAN, Inner Source Address, Ingress RBridge }
    - The Ingress RBridge learned is used as egress on sending
- Via Optional End Station Address Distribution Information protocol
  - { VLAN, Address, RBridge nickname }
- Via Layer-2 Registration protocol(s)
- By manual configuration

WHAT ABOUT RE-ORDERING?

- RBridges are required to maintain frame ordering internally, modulo flow categorization.
- When multi-pathing is used, all frames for an order-dependent flow must be sent on the same path if unicast or the same distribution tree if multi-destination.
- Re-ordering can occur briefly when a destination address transitions between being known and unknown or a topology change occurs.
  - This can be avoided with keep-alives, ESADI, or configured addresses.

WHAT ABOUT LOOPS?

- TRILL Data Frame Loops:
  - Known unicast frames have a hop count and are always unicast to the next hop RBridge.
  - Multi-destination frames must be received on a port which is part of their distribution tree, the ingress RBridge nickname must pass a Reverse Path Forwarding Check, and they have a hop count.
- Hybrid TRILL Data / Native Frame Loops:
  - TRILL takes great care to minimize the probability of there being two uninhibited appointed forwarders on the same link for the same VLAN.
- Pure Native Frame Loops: Not TRILL's problem.

ALGORHYME V2

- I hope that we shall one day see
- A graph more lovely than a tree.
- A graph to boost efficiency
- While still configuration-free.
- A network where RBridges can
- Route packets to their target LAN.
- The paths they find, to our elation,
- Are least cost paths to destination!
- With packet hop counts we now see,
- The network need not be loop-free!
- RBridges work transparently,
- Without a common spanning tree.
- - By Ray Perlner

CONTENTS

REFERENCES

- Specification Draft:
  "Rbridges: Base Protocol Specification"
  - http://tools.ietf.org/html/draft-ietf-trill-rbridge-protocol-15
- "TRILL: Problem and Applicability Statement"
  - http://www.ietf.org/rfc/rfc5556.txt
- Current TRILL WG Charter
  - http://www.ietf.org/dyn/wg/charter/trill-charter.html
- Original Paper by Radia Perlman:
  "Rbridges: Transparent Routing"
  - http://www.postel.org/rbridge/infocom04-paper.pdf

END

Donald E. Eastlake 3rd
d3e3e3@gmail.com, +1-508-333-2270