



ARP/ND and Broadcast/multicast Issues

**Joe Touch & Yu-Shun
Wang**
Postel Center
USC/ISI





Assumptions/Terminology

- Campus
 - Set of interconnected rbridges
 - MAY have internal bridges or switches that do not participate in rbridge protocols directly
- Campus edges
 - Edge rbridges surround edge of a campus, where devices attach
 - Packets MUST enter an rbridge campus via an ingress port of an edge rbridge; MUST leave campus via egress port of an edge rbridge
 - External devices MAY include conventional bridges or other rbridge campuses



Goal

- Goal: rbridge campus looks like a single bridge
 - Internal protocols hidden
 - Externally looks like a bridge
 - Works better than a set of bridges, but not necessarily better than a single bridge
- Change ports without renumbering IP
 - Assume nodes will move
 - Nodes MAY remain silent after moving



Broadcast & Multicast

- See RFC 3819 (advice to link designers)

- Broadcast **MUST** be supported
 - ARP, RARP, DHCP, ND, BOOTP, NTPv3, (RIP)
 - Broadcasts **MUST** reach all rbridge edge ports
 - Rbridge campus **MAY** optimize particular broadcast packets, but still **MUST** support broadcast for unknown payloads

- Multicast **MUST** be supported
 - ND, RIPv2, mcastIP, OSPF, (S)NTPv4
 - Multicasts **MUST** reach all Rbridge edge ports
 - Multicast **MAY** be limited to subscribing edge ports



Broadcast

- Internal broadcast tree
 - MAY support multiple trees to avoid contention
 - Spanning tree(s) managed by internal campus routing protocol
- Serial broadcast
 - Emulates spanning tree(s) via server inside campus
- Server substitute (LANE-like: RFC1577)
 - Requires all external addresses register with server
 - (not tenable for rbridge)

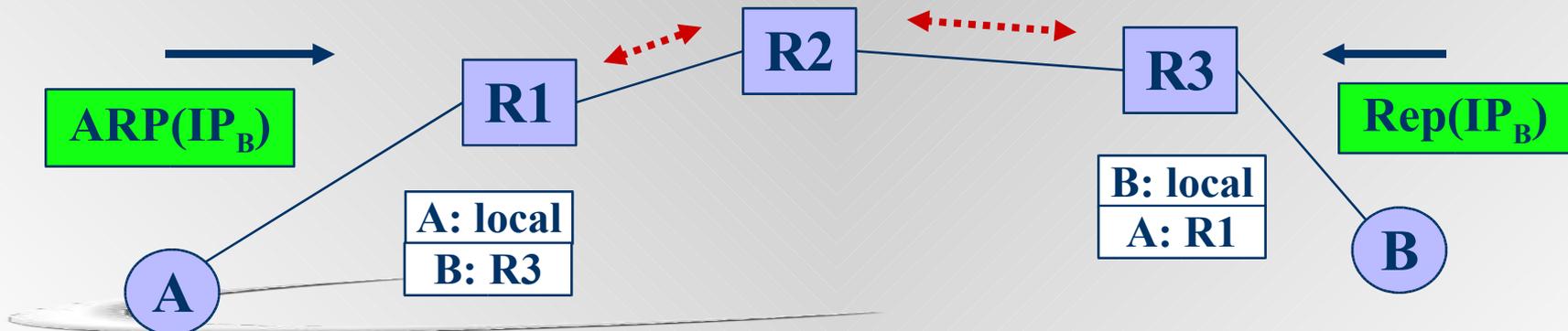


Multicast

- MAY prune hosts
 - IGMP snooping for hosts on mcastIP tree
- Cannot prune all routers
 - If snooping/acting on mcastIP routing, it becomes a mcastIP router
 - Might not prune all mcastIP routers, so MUST continue to support broadcast-based multicast

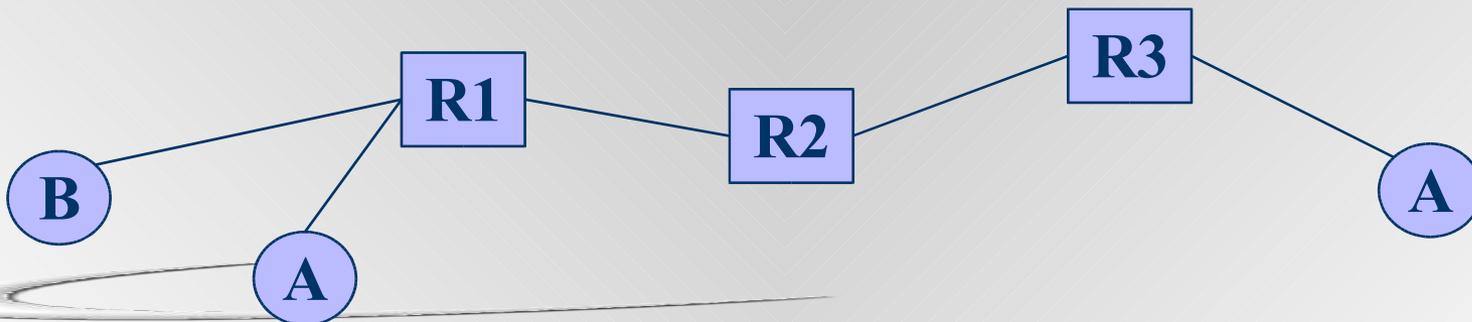
ARP/ND

- 0: Empty (Routing/Edge) Tables
- 1: ARP request/reply enter campus
 - Update edges to encapsulate packets to that ARP reply B to that ingress campus edge port



ARP/ND - Data Pkts

- 2: Packet enters campus
 - 2a: B in edge translation table
 - Encapsulate and direct to appropriate egress edge port
 - 2b: B not in edge translation table
 - (source migrates from learning bridge on campus port it shared with dest to a different campus port)
 - Broadcast packet (current behavior of learning bridge)





ARP/ND Issues

- Silent node moves to a new egress
 - Packets are misdirected until new ARP reply!
 - Same as current learning bridge!
- Keep in mind:
 - Client caching keeps traffic low anyway



Optimization Tricks

- Learn from DAD
 - MUST for v6, MAY(?) for v4
 - MUST broadcast (or won't find duplicates!), but MAY learn from it (use it to load RB routing/encaps tables)
 - Does not cover mobile case - so subsequent ARPs MUST broadcast (or won't find silent movers, or collisions with them)
- Tickle ports
 - So silent movers expose their new position - DNA?
 - Requires periodic tickling
 - Creates unacceptable broadcast storms outside campus



Optimization Tricks

- Proxy ARP
 - Originally intended to avoid IP subnets (when not supported, not to avoid renumbering)
 - Edge RBs as Proxy ARP
 - Problem with mobile nodes



Proxy ARP Variants

- Beginning: Src-IP broadcasts:
“who has Dest-IP link addr”
- Gateway says “Dest-IP is Gateway-link”
 - Original RFC 1027 Proxy ARP
- Gateway says “Dest-IP is Dest-link”
 - Gateway replies from its own ARP cache or is preconfigured with dest-IP:dest-link



Related Work

- NBMA
 - Emulates L2 subnet, not L2 device
 - Most don't support broadcast/multicast
 - Use “server” approach throughout
 - BUS (Broadcast/Unknown Server), ARP server
- L2VPN
 - Emulates L2 subnet, not L2 device
 - Proxy ARP servers between L2 “islands”
 - (Mostly) manual configurations
- ARP mediation
 - That's what routers are for!