An Update on Bridging Technologies

Change and evolution:
Why, what, and how

Mick Seaman
mick_seaman@ieee.org

Agenda

- Why has bridging evolved
- Bridging basics
- What has changed/is changing
- And how!

Why has bridging evolved - technology

- LANs have changed
  - Fewer shared media, more point-to-point links
  - Less geographic coverage w/o bridges
- Everything has got faster
  - Transmission speeds
  - Processing power
- Everything has got bigger/more numerous
  - Many many more stations
  - Many many more bridge ports

- users

- Mainstream commercial
- Telecommunications
- Residential

- expectations

- Mission critical
- Fault tolerant
- Multiple application types

Bridging basics

- Forward frames ‘unmodified’
- Learn station locations from source addresses
- Forward ‘everywhere’ if destination not known

$\Rightarrow$ Topology independent of destination location

$\Rightarrow$ Accommodates many rapid destination moves

$\Rightarrow$ Excellent multicast

Basic bridging requirements

- Symmetric traffic, A ↔ B
  - Otherwise cannot learn
- Multicast A → the same as unicast
  - Otherwise cannot learn
- Loop-free topology
  - Otherwise traffic will not be symmetric
  - Looping (not just forever) destroys bandwidth

What has changed

The way we

- Forward frames
- Define ‘everywhere’
- Learn from source addresses
- Calculate loop-free topologies

Forwarding frames

- Priorities and classes of service
- Integrated Services (IntServ)
- Differentiated Services (DiffServ)
- Drop eligibility

$\Rightarrow$ Compatible, interoperable, and interchangeable with routers adopting the same philosophy

Defining ‘everywhere’

- VLANs
  - Separating and enclosing traffic
- GARP/GVRP
  - Include me in!
- Service provider VLANs (S-VLANs)
  - Separating customers
  - Keeping out of the customers address space

Learning from source addresses

- Initially on every port
- Independently on every VLAN
- Only when it affects the forwarding outcome
  - The convergence of LANs and circuits
  - Makes bridged service provider networks scale

Calculating loop-free topologies

- Loop-free is “tree”, everywhere is “spanning”
- So “spanning tree”, but how ?
- Initial spanning tree protocol distance-vector and timer based, pseudo-static
  - Up to 40 seconds to reconfigure
- Replaced with Rapid Spanning Tree Protocol
  - Timer free reconfiguration (point-to-point links)
  - As little as a few milliseconds to reconfigure

Spanning Tree algorithm

- Root Bridge election (for network)
- Root Port selection (on each bridge)
- Designated Port decision (for each LAN)

$\Rightarrow$ A unique path to the Root for every LAN

$\Rightarrow$ A unique path to any LAN from any other LAN

$\Rightarrow$ ‘Spanning Tree’

But many possible protocols to carry and agree the necessary information

RSTP Rapid Reconfiguration

- Legacy STP was purely timer based
  - Tree structure not used to help configuration convergence
  - Port roles not communicated (!)
  - Forwarding state not communicated (!!)
- RSTP Rapid Root Transition
  - Alternate to Root Port requires no protocol at all
  - Bridges ‘above’ are still above
  - ‘Instantaneous’ and learnt information can be preserved
- RSTP Designated Port Transition
  - Do bridges ‘below’ agree they are below (recurse)
  - Proposal/Agreement of tree construction information
  - 1/3 speed of light convergence possible

RSTP fixes some problems …

- Slow reconfiguration times
- Host protocol timeouts when first connected
- One-way connectivity causing loops
- Mis-set management parameters causing loops

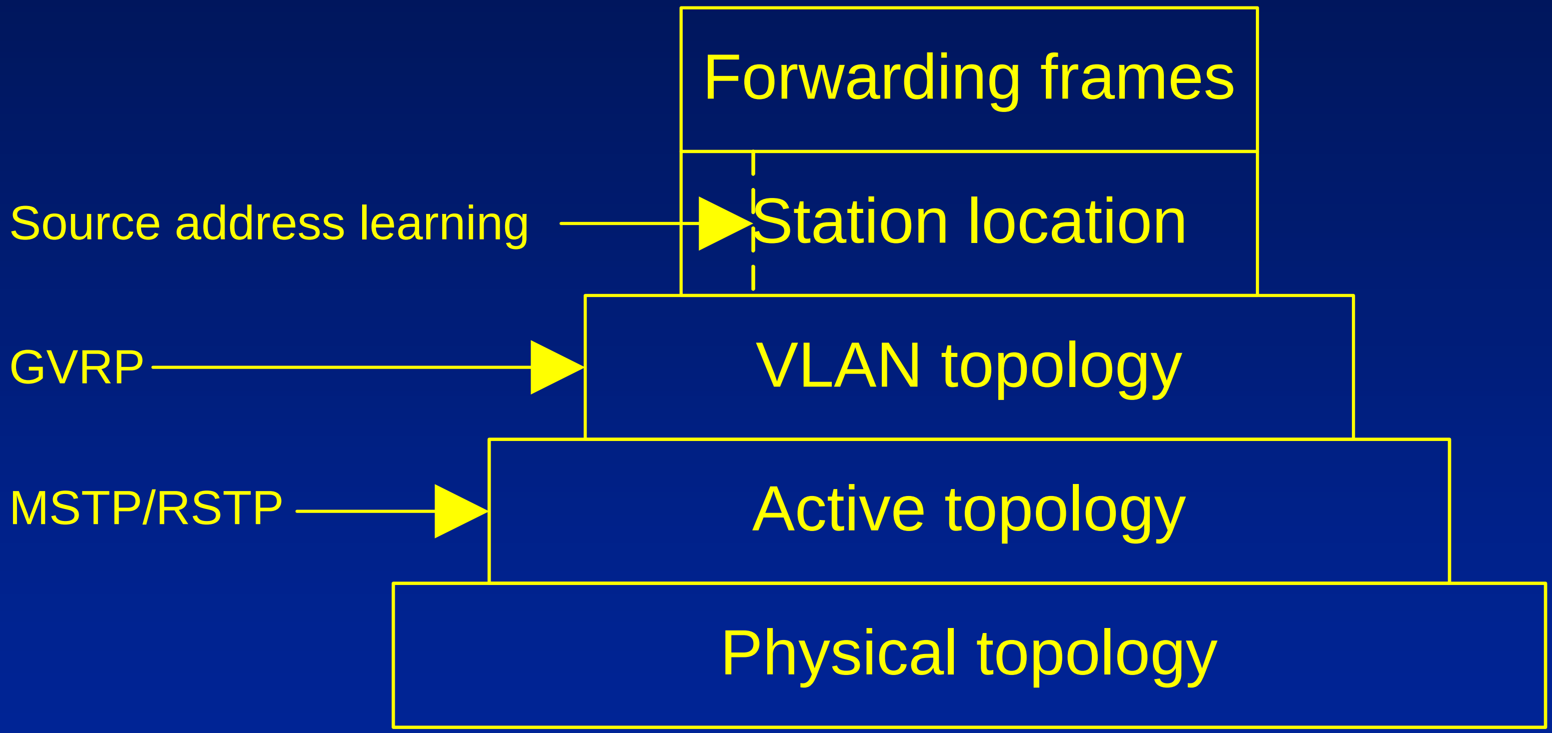
… leaves others

- A brain-dead bridge still has traffic sent to it, and can still forward traffic
  – Note: a fix is on its way
- Unused LANs and links
- Is still called “spanning tree”

Multiple Spanning Trees (MSTP)

- Uses all the RSTP improvements
- Separate trees for separate VLANs
  - Generalized N $\rightarrow$ 1 mapping
- Multiple administrative regions
  - Different N $\rightarrow$ 1 mappings
- But still ‘plug & play’
  - Loop-free communication even if mgmt parameters all wrong

Bridging Architecture (summary)



What is changing

The way we

- Scale the network
- Debug the network
- Determine ‘everywhere’
- Calculate the best loop-free topologies

Scaling the network

- At least 6 million subscribers
  - on a single bridged network
- Many more point-to-point services
- Finally a real need for address encapsulation
  - To hide VLANs!
- Provider Backbone Networks

Debugging the network

A network
- run by multiple independent operators
- offering services to many independent customers
- that is customer IP address plan agnostic
- that offers provisioned services

Connectivity Fault Management (CFM) tools
- continuity checks, link trace, loop back
- operating within and across independent domains
- Supporting an ITU frame work

Determining 'everywhere'

GVRP (VLAN topology mgmt) hardly deployed

– network usually too small
– but now sometime too big

MVRP underway

– very efficient, 4096 VLANs per legal 802.3 frame
– maximum rates, few frames per second per link
– very fast , responsive to link changes
– Communicates topology change per VLAN

Calculating loop free topologies

RSTP confines traffic to a single spanning tree
- unused LANs

MSTP splits traffic across trees by VLAN
- sub-optimal paths

But shortest path from X is on a tree routed at X

And shortest path to Y is on a tree routed at Y

$\Rightarrow$ Shortest Path Bridging

Shortest Path Bridging

Symmetric shortest paths
- Equal cost alternatives labeled by VLAN

Source routed trees labeled by VLAN ID
- VLAN ID pairs/sets shared learning (existing capability)

The hardware you have today can probably do it!
- Nothing new in frame forwarding or learning reqd.
- Controls already present in 802.1Q
- Compatible, interoperable in existing networks

Shortest Path Bridging Protocols

Simple enhancements to MSTP
- Existing frame formats & sizes
- For up to 32 core VLANs, 32 bridges in each

Link state protocol clearly a market need
- Superior convergence for tree calculation
- Modest CPU demands, no per end station component

In conclusion

Evolving technology, customers, requirement

Scale and performance we never imagined

And now,
optimal shortest paths on existing hardware!

Some references (1)

IEEE Std 802.1D

IEEE Std 802.1Q

See the 802.1 WG website for:

P802.1Q-REV

P802.1ad Provider Bridges

P802.1ag Connectivity Fault Management

P802.1ah Provider Backbone Bridges

P802.1aj Two-port MAC Relay

P802.1ak Multiple Registration Protocol (MRP/MVRP)

Some references (2)

An RSTP simulator:

www.ieee802.org/1/files/public/docs2004/RstpSimulation2402.zip

Shortest Path Bridging:

www.ieee802.org/1/files/public/docs2005/new-seaman-shortest-path-0305-02.pdf

www.ieee802.org/1/files/public/docs2005/new-nfinn-mstp-vector-0305.ppt