Converged LAN Routing
or
Routed Aggregation Domains
or
A New Architecture for Secure and Reliable Subnets
or
Some crazy idea we dreamed up while redesigning our campus network

Eric Brown

Clark Gaylord

9 February 2006

Everything you need to know…

- Perlman:
  - *If the data link layer assumes multiple hops, then it is hard to imagine what the network layer is supposed to do.*

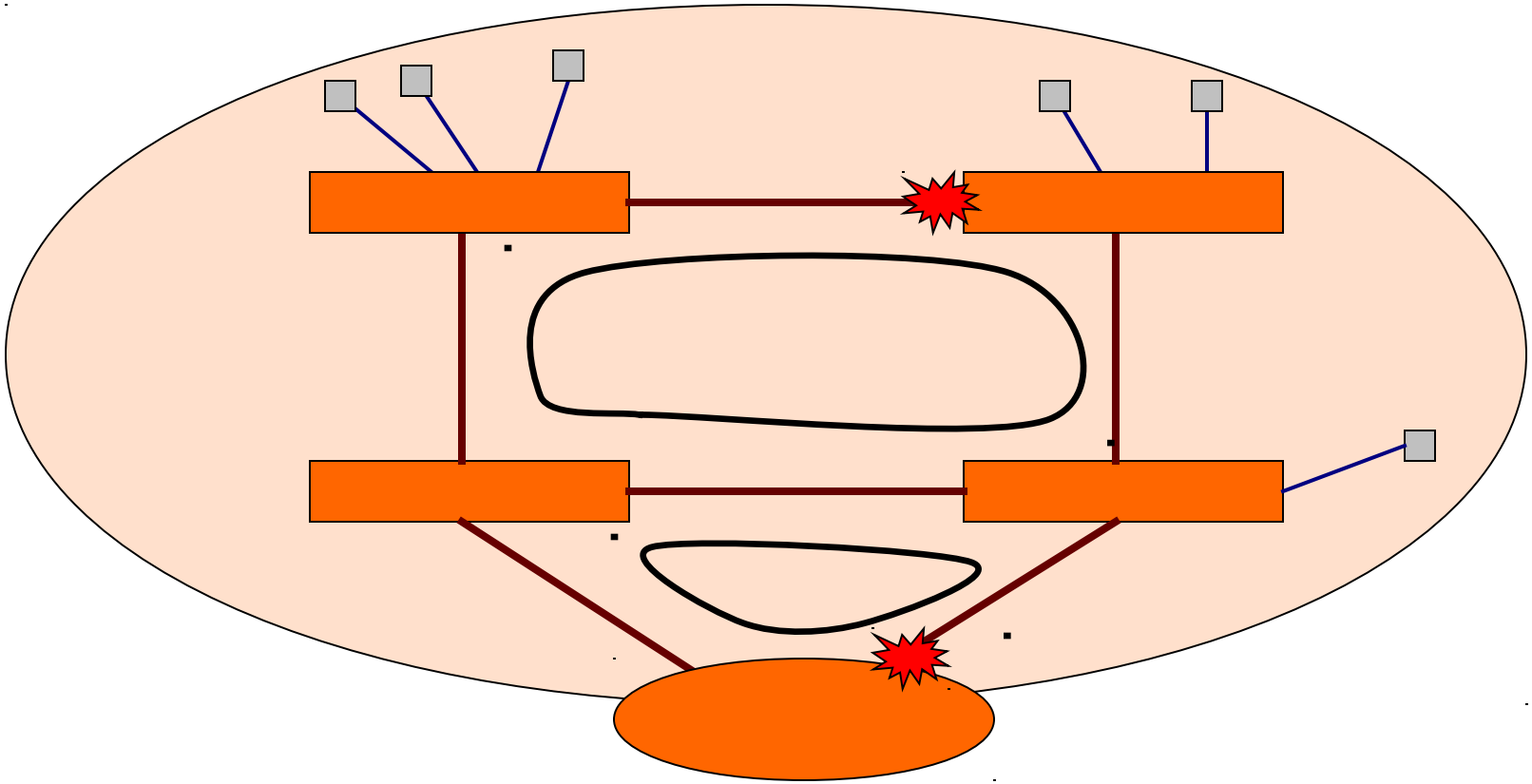
VirginiaTech
Invent the Future

User vs network interface

- A network switch is a collection of interfaces
- Interfaces are either “Network” or “User”
  - What about the access point that talks on the “user” port to another access point?

VirginiaTech
Invent the Future

Topology



VirginiaTech
Invent the Future

LAN Switching is Broken

- Do we really need to explain why we think this?
- How much work does it take us to deploy a LAN – correctly?
- How many VLANs?
  - On all links?
  - How about aggregated links?
  - Where is VLAN 1?

VirginiaTech
Invent the Future

LAN Switching is Broken

- Do we really need to explain why we think this?
- How much work does it take us to deploy a LAN – correctly?
- How many VLANs?
  - On all links?
  - How about aggregated links?
  - Where is VLAN 1?

VLAN Spaghetti!

VirginiaTech
Invent the Future

LAN Switching is Broken

- Are your VLANs included everywhere?
- What is the topology
  - physical or spanning tree?
  - Do you *really* know your spanning tree topology?
- Who's the root?
  - Which ports block?
  - How do you trace a user's path?

VirginiaTech
Invent the Future

LAN Switching is Broken

- Are your VLANs included everywhere?
- What is the topology
  - physical or spanning tree?
  - Do you *really* know your spanning tree topology?
- Who's the root?
  - Which ports block?
  - How do you trace a user's path?

Spanning kudzu

VirginiaTech
Invent the Future

Emulating a coaxial cable

- The complexity of LAN switching happens because we haven't really cut the coax
- We emulate the electrical broadcast property
  - It's called the "Ethernet" broadcast domain

VirginiaTech
Invent the Future

It's easy to build a network

1. Physical topology
2. Plug in switches and let STP do its thing
3. Put a router to gateway network layer
4. Run DHCP

- But is this really the way you want the network to work?

VirginiaTech
Invent the Future

It's not easy to build it *right*!

- ARP is hearsay – so maybe we protect bogus
- Faulty hardware makes Ethernet loops – this is really bad
- Rogue DHCP – ok, we'll protect that too
- Gateway redundancy – ok, run HSRP/VRRP
  - But that's just hearsay too
- Multicast goes everywhere – snoop again

VirginiaTech
Invent the Future

Ogres have layers

Physical & Logical Topology	Configuration

VirginiaTech
Invent the Future

Ogres have layers

Topology Management Protocol	STP and friends, Ring Protocol, Link Aggregation, BPDU protection, UDLD
Physical & Logical Topology	Configuration

VirginiaTech
Invent the Future

Ogres have layers

Link layer 2 “routing protocol”	Transparent Bridging (Learning Bridge)
Topology Management Protocol	STP and friends, Ring Protocol, Link Aggregation, BPDU protection, UDLD
Physical & Logical Topology	Configuration

VirginiaTech
Invent the Future

Ogres have layers

Link layer resolution	ARP (hearsay)
Link layer 2 “routing protocol”	Transparent Bridging (Learning Bridge)
Topology Management Protocol	STP and friends, Ring Protocol, Link Aggregation, BPDU protection, UDLD
Physical & Logical Topology	Configuration

VirginiaTech
Invent the Future

Ogres have layers

Network layer discovery	DHCP, static configuration
Link layer resolution	ARP (hearsay)
Link layer 2 “routing protocol”	Transparent Bridging (Learning Bridge)
Topology Management Protocol	STP and friends, Ring Protocol, Link Aggregation, BPDU protection, UDLD
Physical & Logical Topology	Configuration

VirginiaTech
Invent the Future

Ogres have layers

Redundant gateway	HSRP, VRRP
Network layer discovery	DHCP, static configuration
Link layer resolution	ARP (hearsay)
Link layer 2 “routing protocol”	Transparent Bridging (Learning Bridge)
Topology Management Protocol	STP and friends, Ring Protocol, Link Aggregation, BPDU protection, UDLD
Physical & Logical Topology	Configuration

VirginiaTech
Invent the Future

Ogres have layers

Link layer resolution protection	Dynamic ARP Inspection, DHCP Snooping, IP Source Guard
Redundant gateway	HSRP, VRRP
Network layer discovery	DHCP, static configuration
Link layer resolution	ARP (hearsay)
Link layer 2 “routing protocol”	Transparent Bridging (Learning Bridge)
Topology Management Protocol	STP and friends, Ring Protocol, Link Aggregation, BPDU protection, UDLD
Physical & Logical Topology	Configuration

VirginiaTech
Invent the Future

Ogres have layers

Link layer multicast pruning	IGMP Snooping
Link layer resolution protection	Dynamic ARP Inspection, DHCP Snooping, IP Source Guard
Redundant gateway	HSRP, VRRP
Network layer discovery	DHCP, static configuration
Link layer resolution	ARP (hearsay)
Link layer 2 “routing protocol”	Transparent Bridging (Learning Bridge)
Topology Management Protocol	STP and friends, Ring Protocol, Link Aggregation, BPDU protection, UDLD
Physical & Logical Topology	Configuration

VirginiaTech
Invent the Future

Security problem?

- ARP poisoning is a problem, not because of privacy, but because of availability
- Complex configurations, especially those that mostly work when done wrong, are a serious risk
- Good people and processes can't help but get this wrong!
- "Voice VLAN" considered silly

VirginiaTech
Invent the Future

Routing isn't broken

- Link state database
- Explicit forwarding
- Multiple "active" paths
- Graceful redundancy
- More manageable, well-defined configuration

VirginiaTech
Invent the Future

What do we like about LAN switching?

- Automatically learn where everyone is
- Easy to deploy in simplistic scenarios
- Readily defines the Ethernet broadcast domain

VirginiaTech
Invent the Future

Why can’t we just do routing?

- /30s on the edge port?
  - Too many wasted addresses
  - Support headache
  - DHCP scope explosion
  - No address portability
  - Maybe with IPv6 (subnet per port)

VirginiaTech
Invent the Future

Why can't we just do routing?

- /30s on the edge port?
  - Too many wasted addresses
  - Support headache
  - DHCP scope explosion
  - No address portability
  - Maybe with IPv6 (subnet per port)
- Flat Earth?
  - Not with a pseudo-broadcast technology!

VirginiaTech
Invent the Future

Why can't we just do routing?

- /30s on the edge port?
  - Too many wasted addresses
  - Support headache
  - DHCP scope explosion
  - No address portability
  - Maybe with IPv6 (subnet per port)
- Flat Earth?
  - Not with a pseudo-broadcast technology!
- /32s float around the entire campus?
  - Doesn't scale

VirginiaTech
Invent the Future

How does it work

- Instead of a learning bridge, let's use a learning router
- Currently we learn MAC addresses for the switch's forwarding table
  - Then maybe protect the ARP entry of the host
- Why not just learn the IP address
- The edge switch learns the address by traffic inference
  - All other switches learn by explicit (routing) protocol

VirginiaTech
Invent the Future

How does it work

- The “edge switch” is the gateway
- Answer “that’s me” for all ARP queries
- Learn hosts dynamically (ARP, DHCP, traffic)
  - Tell other switches about attached hosts using a real routing protocol
- Valid host addresses from some “aggregation domain”
- LAN summarizes this aggregate to core

VirginiaTech
Invent the Future

Aggregation Domain

- An aggregation domain is the “subnet” assigned to a LAN environment
- With a true subnet, you can’t have exceptions
- Needn’t be only one block
  - Does this solve address portability (at least for limited scope)?

VirginiaTech
Invent the Future

What does this solve

- No more spanning tree
- No more VLAN trunking
- No multinets
- Readily available location
  - traceroute all the way to the hosts “switch”
  - Users still think of their addresses as a block
- Portability within an aggregation domain
- Allows portability between aggregation domains – if we want to
- Better interface for policy configuration – at the host
- Doesn’t require IPv6

VirginiaTech
Invent the Future

Ogres have layers (redux)

Link layer multicast pruning	IGMP Snooping
Link layer resolution protection	Dynamic ARP Inspection, DHCP Snooping, IP Source Guard
Redundant gateway	HSRP, VRRP
Network layer discovery	DHCP, static configuration
Link layer resolution	ARP (hearsay)
Link layer 2 “routing protocol”	Transparent Bridging (Learning Bridge)
Topology Management Protocol	STP and friends, Ring Protocol, Link Aggregation, BPDU protection, UDLD
Physical & Logical Topology	Configuration

VirginiaTech
Invent the Future

More of everything you need to know

- Perlman
  - *A data link layer protocol is anything standardized by a committee chartered to standardize data link layer protocols.*

VirginiaTech
Invent the Future

Contact

Clark Gaylord

cgaylord@vt.edu

VirginiaTech
Invent the Future