Accurate yet long winded title:

Exposing network, VM edge and power topology via LLDP and BGP-LinkState -- and possible implications

Brian Field / Comcast
Applying web principles to the network
[control plane focus]

Brian Field / Comcast
Conway’s Law

• “Any organization that designs a system (defined broadly) will produce a design whose structure is a copy of the organization's communication structure.”

• What does this mean?
  – Think Reese’s peanut butter cups—chocolate and peanut butter – two great tastes that taste great together
  – Conway’s law would suggest org structures hamper creation of Reese’s peanut butter cups
What has made the web so successful?

- HTTP – open
- Apache – open source
- Linux – open source
- COTS hw – cheap

Web stack
Web model versus router model

**Web stack**: HTTP, Apache, Linux, COTS hw

**Router stack**: BGP, ISIS, PIM, etc., Proprietary CLI, Proprietary OS/chassis, Proprietary ASICs
Is this SDN?

- “open” API
- manage (“hide”) proprietary router environment
- Can we do better?
Service Providers and virtualization

- Virtualization
  - COTS hw
  - Many servers, many locations

- Apps as VMs

- App could run across multiple CPs (cloud providers)

- Agility
Applying application thinking to the network

1. Router VM in every COTS hw virtualization server

2. Rethink how we build/encode network control information

3. Bigdata thinking applied to network–per web transaction mining -> per packet mining
Background to Rethinking how we build/encode network control information

• Network/Ops:
  – What App/VMs are really running in site X right now?

• Facility:
  – Not all sites and servers are identically configured power-wise

• Application:
  – Are my VMs running in an environment that has sufficient network/server/power redundancy?
Three infrastructure topologies related to cloud applications

1. Network topology
2. VM “edge” topology
3. “power” topology

• Who owns auditing these topologies?
  – What if my app spans multiple cloud providers?
Who owns auditing these topologies?

• Past:
  – Some org or the network folks

• This isn’t a Comcast problem, this isn’t a cloud provider problem this is an application problem

• Put auditing onus on the apps

➢ *It’s hard to audit info that is dynamic and isn’t learned in real-time*
How do we dynamically expose these infrastructure topologies?

- IGP
- LLDP
- BGP LinkState
Two key protocols

• BGP LinkState
  – New AF
  – Export IGP and TE state north bound
  – Part of IETF I2RS WG
  – draft-ietf-idr-Is-distribution-02

• LLDP
  – Link Layer Discovery Protocol
  – TLV based ethernet protocol
  – Supported in linux and recent router code
Step 1: Expose network topology

- Power distribution units
- Power cords
- Virtualization servers
- Network links

Network

- Router
- BGP
- Link State
- Route reflector
- Real-time network state

Tools, Apps

5/30/13
Step 2: Expose VM edge topology
Step 3: Expose power topology

Network

Tools, Apps

router reflector

BGP Link State

real-time network state
real-time VM state
real-time power-server state

Network

Router

IGP

power distribution units
power cords
Virtualization servers
network links

PDU-42

PDU-43

PS-1 VM-X
PS-2 VM-Y
PS-1 VM-Y
PS-2 VM-Z

LLDP

LLDP

LLDP

LLDP

LLDP

LLDP

LLDP

LLDP

LLDP

LLDP

LLDP

LLDP

LLDP

LLDP

LLDP

LLDP

LLDP

LLDP

LLDP

LLDP

LLDP

LLDP

LLDP

LLDP
What are we left with?

• A real-time database of:
  – Where all services are running
  – App/VM->server->router interface mapping
  – PDU->Power Supply->server mapping

• Application Value:
  – Better VM placement
  – Enable app auditing
Lots more info to encode into BGP LinkState

• Encode this new information in bits/bytes?

• Is it time to think about something a bit more current?

• Like XML/JSON?
  – Hey, the web guys developed and use that stuff don’t they?
  – Aren’t they pretty successful at what they do?
Web thinking applied to network control protocols -> Make BGP RESTful

• Proposal:
  – Encode BGP LinkState using XML/JSON
  – Convert BGP LinkState primitives to be RESTful

• Result:
  – BGP LinkState operates like a web app
  • Apply web’s software methodology, scale and efficiency in the network control protocol space.

• Does this apply to the IGPs too?
Remember I mentioned router VMs?

Open source router code!

Power distribution units
Power cords
Virtualization servers
Network links
Which is better?

• Virtualization applied to routing is the evolutionary tipping point in the network space

• open API (SDN) vs. open source (router VM)
Applying web principles to the network [making peanut butter cups]

Thanks!

brian_field@cable.comcast.com