

NTAC Peering and Routing Working Group

12/15/2020 Call Notes

1. Agenda Bash
 - a. No bashing
2. Update on peering and I2PX (Bartig)
 - a. Not much to update on peering. Some progress with Amazon in Seattle, now have 2x 30G LAGs. Hopefully 100G's in place in 2021 there. This week working on capacity augment with Verizon EdgeCast to 20G. Working with Microsoft in Ashburn, Chicago to 2x 100G in both locations. Traffic now on decline.
 - i. Looking to hear from campus on how many will be remote vs in person.
 - ii. Oregon state – still remote
 - iii. Minnesota – hybrid. If class > 30= remote. If < 30, can be hybrid.
 - iv. Clemson – hybrid, same as Minnesota above.
 - v. Maryland – targeted classes, like labs are on campus. About the same in the spring.
 - vi. Schools in Pittsburgh area will be hybrid.
 - vii. No one else seemed to be doing anything any different than the above.
3. Network Weather Update- Trends and coming events
 - a. No discussion.
4. I2 Network update
 - a. See below.
5. Mark Brochu/ NGI VRF structure
 - a. Current and potential strategies
 - b. Pain points
 - i. Configuration cruft/complexity
 - ii. Multiple tables intermingling
 - iii. BGP routing policy variably defined
 - iv. Policies between members not defined by BGP communities
 - v. Complexity – long engineering times
 - vi. Something I missed
 - c. Things we've assumed in the past
 - i. BGP ASN = VRF
 - ii. Separation of services/routes = separate VRF
 - iii. Process of removing independent domain has taught us about ASN operation on MXes. Had to removed to make compatible with NGI gear.
 - d. Opportunity for fresh start
 - i. Don't want to reinvent the wheel
 - ii. BGP policy should be well defined for all parties. BRM/Automation/Engineering.
 - iii. Excess FIB space has allowed us to be "lazy" – VRF sprawl
 - iv. We want a consistent VRF/routing table design that's efficient – fits well in 8200 platform and doesn't pigeonhole us from using other devices.
 - v. Consistency drives efficiency and is predictable.

- vi. Need to finalize VRF structure as soon as possible as NGI models depend on it.
- e. Some ideas and principles
 - i. Considering R&E in a VRF
 - ii. Challenge idea that separation is only solvable with VRFs
 - 1. Merge Lumen transit routes with I2PX table
 - iii. Full commercial routing at every node unless there is technical reason not to.
 - iv. Can all be done on 8200 with headroom.
 - 1. Transit/Blended VRF
 - 2. R&E VRF
 - 3. I2PX VRF
 - 4. OSG VRF
 - v. Having full transit in i2PX solves some problems.
 - 1. Wallace: what happens if someone points a default to I2? Would they then not need RIPcord?
 - a. Bartig: yes, this could be abused. But they would be able to send egress traffic but would get no traffic back, so they would still need RIPcord for full connectivity. Participant routes would not be advertised to Lumen.
 - vi. Likely that will have to completely deprecate multicast with R&E in a VRF.
 - a. Farmer: many platforms only support mcast routes following unicast routes.
- f. Concerns:
 - i. Is added complexity a risk to NGI?
 - ii. Will have 4 route reflectors: each RR will handle 900K vpn4/100K vpn6 prefixes.
- g. Farmer: note that you don't have to have full routes to solve all these problems. Would be more work on front end but will require fewer router resources.
- h. Some testing of convergence: wanted to be sure it was as fast as current setup.
 - i. Junipers: 12K/second for global table 7K/sec convergence for VRFs.
 - ii. On 8201 full table convergence took 66 seconds.
 - iii. Global insert ~ 60 seconds
 - iv. Numbers overall look good on convergence.
- i. Matt Valenzisi: What is concern with RR vs no RR? Have seen better convergence with RR on MCNC.
 - i. Mark Brochu: there are 2 questions: first: if PE node has full transit, how much addl load will be put on reflectors because now have to send 800K routes to all other PE nodes? Gut: probably not an issue. Will be using pretty beefy hardware. But it is an unknown. RR's are a SPOF.
 - ii. Matt V: could split up address families.
 - iii. Mark: overall, encouraging news to hear positive comments on RR's.
 - iv. Farmer: initially were using ASR9K's now using 9001's which are pretty beefy.

- v. Farmer: with RR, don't need ASIC performance, just CPU performance. Much easier to scale processor than packet processing. Most of routing platforms even have multi-core processors.
 - j. The working group was very appreciative of Mark's presentation and offered its help as needed.
6. AOB?
- a. Discussion of Network Management solutions.
7. Adjourn at 5p EST