

PCH IPv6 Deployment

July 2009, Version 1.1

Gaurab Raj Upadhaya

gaurab@pch.net

Background

- We operate an anycast DNS Network, about 60 odd ccTLDs and gTLDs
- Spread out over 40+ locations
- We have had v6 unicast on two nodes since 2005
- But, then in early 2007 decided we should go the full way

DNS

- Enough said, and will be said

Anycast

- This is not IPv6 Anycast as defined in RFC 2526 (amongst others), but anycast as described in RFC 4786
- Same prefix announced from multiple location, so that the clients connect to the topologically closest node

Get ARIN space

- we got two /48
 - Well, a /47
 - The prefix comes from the 'critical infrastructure' ARIN space
- 2001:500:14::/47
 - One /48 for Anycasting ccTLD
 - Second /48 for Management

Address Planning

- First /48 - Anycast subnet : 2001:500:14::/48
- Second /48 - Management : 2001:500:15::/48
 - We distribute a /56 per POP, use /64 per lan/vlan
 - Mapped with POP ID
 - We use /64 for tunnel end-points, end site networks get /56s.
 - Map the bits in /56 to some of the bits in /64.

Address Management

- Idea was to make it easy for us to see v6 address and relate to other existing numbers/IDs to relate to them
- Tooling wasn't there to automate things we'd otherwise automate for v4.

Issues

- It took a long time to plan and prepare
 - But once started, deployment was quick.
- Servers needed upgrading
- We decided that we'd do global rollout only with the new hardware (OS platform was also changing)
- Routers needed the right IOS and there were changes on the router CLI, means we also had to re-tool for routers

Quagga

- Quagga is important for us
 - The servers talk to the next-hop router through Quagga, We also use quagga for some peering
 - We want to use the same versions of quagga
 - Upgrading to the latest (0.99.10) seemed to have fixed many issues
- Weird Randomness
 - IPv6 patch seem to break MD5 functionality - now fixed
 - We still set session to 'passive' to JunOS boxes
 - Still have issues with peers with old version of Quagga

BIND and NSD

- BIND and NSD were least of the problems
 - Worked almost out of the box
 - We are not yet using v6 transport for AXFR/IXFRs etc
 - None of the zone operators have even given us a v6 master address

Going Production

- First was to go public with our global nodes in Bay Area, Ashburn, London, Paris and Hongkong.
- Added transit from HE, to the one we had from NTT/Verio
 - Will soon add TATA
- Add entries in RADB (again another tooling need)

Going Live

- We put everything in place, before informing the zone operators
- Early movers identified and informed. They ran tests, confirmed things look good, informed IANA, done
- Inform all other zone operators
- For at least 12 operators, it was their first v6 secondary, for 11, it's still their only v6 server
 - Not all have updated IANA though
 - Ongoing process

Peering

- Again, takes time to get v6 address from IXP operators, it's not yet 'production' for every one
- Don't assume that because you went production, everyone else is ready
- Not all of our current IXPs will do v6
- In at least one location, we have no v6peers
- www.pch.net/peering

Getting IXP v6 addresses

- Three of those
 - Easy, just part of the regular allocation (and in fact we already had the v6 on the IXP for many years, without us knowing)
 - Moderately Easy. Hmm! this is still an 'engineering project', can you fill the form out? and then we get the address.
 - No Response, or something along the lines of 'oh Yeah, we're thinking of it.'

IXP Formulas

- Using ASN : Ten; e.g 2001:7F8:1::A500:42:1
- Using ASN in Hex : Three ; e.g 2001:7F8:4:0::2A:1
- Use v4 last octet : Six; e.g 2001:43F8:60:1::122
 - (where v4 is 198.32.144.122)
- v4 last octet, but HEX : Three ; e.g 2001:504:D::35
 - (where v4 is 198.32.176.53)
- Nothing Specific, next in Queue : Two

Side effect

- I know now that

- 2A = 42

- F10 = 3856

- 1B1B = 6939

Ongoing

- Upgrade all sites - 30 ready with v6
- Tooling for everything
- Monitoring / logging / scripting all needs to be done

Conclusion

- Planning and Preparation took longer than we thought
- Routing, DNS etc is easy
 - It's the tooling and planning that takes time
 - So it'd be good to see some nice tools for v6 (like say reverse DNS)
- There is still lack of support for extended tools for IPv6
 - Flows, SNMP

Questions

gaurab@pch.net

PCH is a 501(c) non profit organization