ESnet as ISP: Providing Internet Connectivity for Office of Science Sites

Michael Sinatra

Energy Sciences Network

ESCC – January 2012

Overview

- Evolution of ESnet as the Office of Science ISP

- Transit vs. peering

- Tier-1 vs tier-2

- De-peering events of 2011

- Tier-2 strategies

- ESnet's role as ISP in the future

Acknowledgements

This talk is Brent Draney's fault.

Evolution of ESnet as an ISP

- ESnet formed before the Internet was privatized.

- Peered at the FIXes, subsequently MAEs.

- ESnet "just peered" with everyone at the time.

- This evolved into ESnet's initial role as a tier-1 ISP.

- Unlike Internet2, ESnet provided all connectivity to sites, including commodity transit.

- Internet2 sites pay Internet2 for R&E connectivity, but must buy their own transit or get transit from a gigapop/RON.

Peering vs. transit

- ESnet was an important part of the early Internet, as the Internet *was* an R&E network!

- Organizations peered with ESnet and shared the costs of the peering. There were no measured traffic costs with these peerings.

- → This is known as settlement-free peering (SFP). I will simply refer to it as "peering" in this document.

- As the Internet evolved into a commercial venue, organizations wanted to connect without building their own wide-area networks. Internet Service Providers (ISP) allowed individuals and organizations to connect to the Internet. ISPs would then peer with other networks to provide full connectivity for their customers.

- → This is known as "transit": Customers pay the ISP for connectivity to other networks via the ISP's network.

Peering vs. transit

- Transit is usually charged based on the number of bits transferred.

  - Sometimes transit is 'measured usage': Some rate of usage (often per mbps per month measured at the 95th percentile).

  - Transit can also be 'flat rate' based on the size of the link; e.g. a 1GE link costs $1000/mo, regardless of actual usage—but larger links cost more.

  - Could also be a combination of both: A flat, committed rate, with a burst capacity above the committed rate charged at the 95th percentile.

- Peering is only flat rate, and usually only covers the cost of the port and cross-connect. Occasionally peering can be "paid peering" where one peer charges another for the "privilege" of peering.

Peering vs. transit

- The way routing is done between peering and transit is different:

  - Peering involves the exchange only of routes within the peer's network and (usually) the customers of the peer's network.

  - Transit involves the exchange of all of the above plus routes to all of the other networks to which the transit provider connects. Often, this is the rest of the Internet.

  - Transit customers can point a default route to the transit provider if they don't want to exchange routes dynamically via BGP. Otherwise, they will usually receive the full default-free Internet routing table from the transit provider.

Tier1 vs tier2

- Tier1 ISPs *only* peer. They *never* buy transit.

- Tier2 ISPs peer and buy transit. They generally buy transit from multiple tier1s, but may also buy transit from a mix of tier1s and tier2s.

- Tier3 ISPs generally only buy transit (usually only from tier2s).

- Tier1s *must* peer with all other tier1s; otherwise there will be holes in the Internet!

  - Peering disputes sometimes cause this.

  - "Hey, I thought the Internet routed around these problems." Yes, the Internet is capable of routing around problems, but economics is not.

ESnet history

- For its first 25 years, ESnet was a tier1.

  - ESnet was a key part of the Internet back when the Internet was an R&E network.

  - When peering began at the FIXes ESnet peered with everyone, as happened subsequently with the MAEs.

  - ESnet continued to peer with all other tier1s because it just made sense. ESnet had a lot of traffic compared to the rest of the Internet.

  - As the Internet became commercialized, all of the tier1s became commercial networks—except for ESnet.

  - Eventually ESnet was eclipsed (significantly!) as a traffic sink and source *to the commodity Internet.* ESnet is still a major traffic sink and source with respect to R&E networks.

ESnet history

- Maintaining settlement-free peering does have costs, and it must be "worth it" to both sides.

  - Traffic levels must be reasonable.

  - Need to peer with other tier1s *or* buy transit (which would make you a tier2).

  - → ESnet increasingly at crossroads with respect to its peers.

  - *Relative commodity* traffic levels getting lower, cost of maintaining elaborate tier1 infrastructure becoming larger.

  - ESnet resorted to some paid peerings, but generally stayed away from buying transit.

  - Frequent discussions as to whether this status was still a good thing.

June 2011

- In June 2011, ESnet was unceremoniously de-peered by Cogent Communications, which became a tier1 in the late 1990s as a result of its acquisition of PSInet.

  - No notice was given, but initially Cogent said they would restore our peering. Later, Cogent backed off this promise.

  - "…automatically dropping interconnects at around the time of your HQ's close of business while your peering co-ordinator is on vacation seems to be the new gold standard." –Nick Hilliard, NANOG mailing list.

  - Although Cogent handled the situation badly, an eventual tier1 de-peering was probably very likely.

June 2011

- ESnet had to scramble after the June 2011 de-peering to purchase transit. Because ESnet had been a tier1, this single de-peering left ESnet with a big hole in its routing table.

- We found instances of "cloud" providers who were singly connected to cogent (!). That prompted me to say "it's not a cloud if it only has a single upstream provider." (Eli came up with the more clever "it's not a cloud if it's on the other end of a straw.")

- These cloud providers couldn't get to ESnet or its sites, nor could sites get to certain cloud providers.

- We were able to push hard on purchasing and get transit via Hurricane Electric in record time.

November 2011

- On November 15, 2011, AT&T de-peered ESnet, although they provided reasons and considerable notice. They worked with us to provide an easy transition.

- ESnet had already moved toward being a tier2 provider. That meant that we would enter the market to purchase more transit service.

- We wanted transit from a tier1 provider, for various reasons.

  - Hurricane Electric is a tier2.

  - There were some bizarre situations where HE would sometimes use upstream providers that ESnet peered with, and this would create strange asymmetries.

  - There were also situations were some peers weren't seeing our routes through HE.

November 2011

- Due to the presence of existing contracts and good pricing, we were able to purchase transit from Level3 using Internet2 as a re-seller.

- Transit comes direct from Level3 at two peering locations (Sunnyvale and New York).

- The augments our HE transit at San Jose, Chicago, and Ashburn.

- ESnet now has 3.9 gbps of committed transit capacity, with burst capacities upwards of 5 gbps.

- More capacity can be added if necessary.

Lawrence Berkeley National Laboratory

U.S. Department of Energy | Office of Science

ESnet as a tier2 ISP

- Having a mix of tier1 and tier2 upstream transit allows us to maximize redundancy and minimize points of failure.

- It also insulates us against peering disputes.

- It does complicate the traffic engineering tasks, but also gives us some flexibility that we didn't have before.

- Allows us to continue to be "one-stop-shop" for connectors who don't wish to purchase their own transit (or who only want to use their own transit links for backup).

ESnet as a tier2 ISP

- There are some disadvantages of being a tier2:

  - Additional complexity for traffic engineering (but arguably more flexibility).

  - Difficulty troubleshooting—problems may be two ASes away from ESnet, but may affect sites. (E.g. GBLX – Level3 integration.)

  - Can be costlier than settlement-free-peering. (But usually cheaper than paid peering!)

  - Some entities will only peer with tier1s (e.g. other tier1s), leading to more use of transit.

  - Latencies to certain commodity sites may get worse (but not always!).

  - Well, there's the cachet of being the "world's smallest tier1"…

ESnet's ISP goals for the future

- Continue to be your "one-stop-network-shop."

    - Provide a full range of robust connectivity, ranging from ultra-high-capacity links for data intensive science, to high-capacity R&E connectivity, to robust, reliable, and redundant commodity connectivity.

    - Deliver all of this in one connection, if desired, or in a set of redundant connections (if feasible).

    - Obviate the need for sites to do their own traffic engineering, but don't preclude them from doing so if they desire.

    - Move in-step in meeting sites' needs as they move more of their services to the "cloud." (Or simply outsource them to network-connected service providers, to use less-cool lingo.)

    - Understand the growing importance of all types of networking, *and plan and execute accordingly.*

ESnet as *your* ISP

- How you can help us do our job better:

  - Let us know your needs, especially with respect to data-intensive science, R&E connectivity requirements, *and* commodity Internet needs.

  - Inform us of outsourcing initiatives that have network implications; e.g. Google Mail/Apps, on-line backups to offsite providers (e.g. Carbonite), etc.

  - When your site is considering outsourcing to a particular vendor (including RFP processes), make sure that they understand to ask about network provider redundancy. ESnet staff can help ask the right questions!

  - Let's also talk about places that need to access your sites as well, e.g. telecommuters on cable or DSL/FiOS/uverse. What providers are people using and what are their "pain points"?

ESnet as *your* ISP

- How you can help us do our job better:

  - Where do you need low latency or high bandwidth (or both)?

  - Are there problem sites or places in the Internet that give you problems, where we might be able to help?

  - In short, where are you getting the most complaints from users about "the network"?