Complexity - The Internet and the Telco Philosophies

A Somewhat Heretical View

2002.10.28

NANOG / Eugene

Randy Bush <randy@psg.com>

<http://psg.com/~randy/021028.nanog-complex.pdf>

The Internet **DOES** Work

- IP forwarding is just as fast as switching, they're all just one-table lookups
- Actual measurements show QOS is just fine due to proper provisioning, and proper provisioning is cheaper than adding a layer
- QOS is a decision of which packets to drop. Do you get paid to drop packets?
- There are reasons the Internet has taken over the data world and threatens voice

The Internet **DOES** Work (cntd)

- Reliability and Resiliency are core strengths, the Internet was designed for them (our weakness is security, as was the telco's once)

- IP routing yields as good service as MPLS Layer-2.5 switching, and better in cases of multiple extreme failures

The Only Real Problem is

Scaling

All the others inherit from that one

If you can scale, everything else must be working.

-- Mike O'Dell – Chief Technologist UUNET

Complexity is the Arch-Enemy of Scaling, Hence of the Bottom Line

- Telco culture started to glorify complexity as a competitive tactic in the '70s

- But look what it has done to OpEx

- We are all in a commodity market and buy from the same vendors as the competition

- RFC 1925 section 2(3), "With sufficient thrust, pigs fly just fine." This does not mean we can afford the fuel costs, and who wants to fly pigs anyway?

Copyright 2002, Randy Bush

'The Hitchhiker's Guide to the Galaxy … says of the Sirius Cybernetics Corporation products that `it is very easy to be blinded to the essential uselessness of them by the sense of achivement that you from getting them to work at all.

   'In other words -- and this is the rock solid principle on which the whole of the Corporation's Galaxy-wide success is founded -- their fundamental design flaws are completely hidden by their superficial design flaws.' -- Douglas Adams in "The Hitchhiker's Guide to the Galaxy"

Where the Smarts are

- Traditional Voice has stupid edge devices, phone instruments, and a very smart core

- The Internet has smart edges, computers with operating systems, applications, …, and a simple stupid core, which just does packet forwarding

- Adding an entirely new Internet service is just a matter of distributing an application to a few consenting desktops (until NATs)

- Compare that to adding a service to Voice

Where the Reliability is

- The Voice network has smart central organs which are heavily armored, have rooms of battery backup, etc.

- The Internet **assumes** major component failure and achieves reliability through redundancy in the protocol designs

- I.e., the DNS root servers can be seriously attacked without customers noticing

- The protocols simply find a working one and remember it until it fails

Carrier Class Reliability

- The famous 5ESS switch regularly has five nines in operation and has even hit six nines in the field

- We all think we want that in routers and other internet boxes

- Can we achieve this?

The Truth About How They Do It

- There is a supervisory function which runs continually cleaning up internal inconsistencies in the data structures

- Without this, the 5ESS crashes in a matter of hours

- Can you imagine this approach scaling to internet routing?

Spread it Across Layers

- RFC 1925 2(5) "It is always possible to aglutenate multiple separate problems into a single complex interdependent solution. In most cases this is a bad idea."

- This is why ATM-1 failed in the Internet, it tried to solve QOS, traffic engineering, ...

- RFC 1925 2(11) "Every old idea will be proposed again with a different name and a different presentation, regardless of whether it works."

Trade-Offs Across Layers Yield Power and Simplicity

- L2/L3 technologies such as Frame, IP, MPLS, … have costs analogous to software costs, they drop very slowly

- Fiber bandwidth costs and pricing seem to follow Moore's law, with cost per gigabit halving every 1-2 years

- Maybe we should not increase the L2/L3 costs with pseudo-muxing games like MPLS, but rather use L1's DWDM to give each their own channel/lambda

A Heritical Researchers' Approach to Operations

- With enough complexity we strongly suspect that we can operate an approximate internet in polynomial time and dollars

- We are working on a proof that operating the internet can be made to be NP hard

- And then we'll just wonder where the margins went

Copyright 2002, Randy Bush