Measuring “Next-Generation” Networks: HOPI

Matt Zekauskas
matt@internet2.edu
2006-02-07
HOPI

• Emulate a circuit-switched architecture
• Extensive connectivity to packet-switched architecture
• Explore “Hybrid Optical-Packet Infrastructure”
• Explore dynamic signaling
• It’s a testbed, not production
HOPI Node

- NLR
- RON
- Ethernet Switch
- Optical Cross Connect
- Out of Band Access
- Control
- Measurement
- Support
- Abilene
  - Abilene T-640
  - GigaPoP
  - GigaPoP
As a testbed...

• Lots of flexibility in a “circuit”
  – Entire 10GE
  – A 1GE VLAN
  – Manually setup
  – Dynamically allocated (demonstrated dynamic VLANS via GMPLS using DRAGON control plane)
    – Setup: Email, to O(minute)
    – Duration: minutes->hours->week

• Constantly evolving
But… currently all Ethernet

- Utilization
- Errors (but: if no traffic, no errors?)
- Possibility of injecting traffic parallel to other VLANs
- Possibility of passive measurement by port replication
- Packets tossed because of internal resource limits?
What do I mean by “Currently”?

- Potential for “southern path” to be OC 192
  - (Likely, want to experiment with GFP/VCAT/LCAS)
- Can optically bypass Ethernet switch
Initial Plan

• Collect control plane decisions
  – DRAGON has a “router proxy”
  – Web services in the future

• Leverage Force10 statistics (SNMP utilization, errors. Future: sFlow?)

• “Router proxy” to examine switch state

• Nagios verifying stuff “up” (TBD: informed by control plane)

• Ad-Hoc use of measurement machines (e2e replacement, alongside, or in the middle)
Initial Plan – Not yet implemented

• Measurement machines
  – continuously/periodically…
  – signal network
  – measure resulting path (throughput, latency)
  – Cycle through full mesh of 5 HOPI nodes (will initially, at least, pre-compute schedule)
  – Exercise control plane as much as verifying circuits
Problems?

- To date: underlying circuits failing completely
- Otherwise, has just worked…
I wish I had...

- More time, debugged cloning technology
- Better ability to stress circuits
  - 10GE PC’s at line rate just becoming reality; Spirent gear and their ilk expensive
- Reports from endpoints at circuit teardown
- Flexibility to passively focus on circuits
  - A-la “lambdamon” [Micheel - PAM2005]
  - And/or SCNM (Self-Configuring Network Monitor) [Teirney – PAM2003]
- Statistics: more more more…
  - caveats: I’m a {packrat, engineer, packet-switcher}
What if you’re Layer 1?

- Errors / errored frames
  - Before and after error correction?
- Light levels
- Current state (what maps to what)

- And?
  - I’m open to suggestions…
Complications

• As you glue together different technologies (L2+L1+MPLS+…) if there is a problem, finding that problem will be harder;

• If you don’t use SONET at L1, indications from network are potentially fewer (or different); GFP operations and monitoring functions?
A loss of functionality?

- divide-and-conquer by adding active equipment
- Statistics from routers (utilization, malformed data packets)
- Convenient points for passive traces

⇒ A loss of visibility

(personal bias: debugging performance problems)
Thanks

- Hopi design team, corporate advisory team, everyone I’ve forgotten, and …
- The Technical Service Center
  - Indiana University [the NOC]
  - MAX [control plane]
  - NCREN [application support]
- For more info: http://networks.internet2.edu/hopi/
References


Back Pocket Slides
HOPI Project - Overview

• We expect to see a rich set of capabilities available to network designers and end users
  – Core IP packet switched networks
  – A set of optically switched waves available for dynamic provisioning

• Examine a hybrid of shared IP packet switching and dynamically provisioned optical lambdas

• HOPI Project – Hybrid Optical and Packet Infrastructure - how does one put it all together?
  – Dynamic Provisioning - setup and teardown of optical paths
  – Hybrid Question - how do end hosts use the combined packet and circuit switched infrastructures?
  – HOPI is a testbed for experiments, not a production network
  – We will use some of the experiment results to guide the next generation of Abilene
Previous talks

- This talk is a follow-on to "(Next-Generation Network) Measurement Infrastructures BoF" at the Vancouver Joint Techs in July. Slides are expanded...

What are the right metrics?

• Use IP-based ones (packet oriented)
  – Latency
  – Loss
  – Throughput verification
• Use telephony-based ones
  – Circuit setup time
  – Errored seconds
  – Whatever the ITU has been doing for years (need to investigate, don’t have any kind of systematic or exhaustive list)
What are the right tools?

• Could some passive measurement architecture, such as Lambdamon or PIANO, get us back some visibility?
Initial Thoughts

• Collect control plane decisions (with reasoning?): state
• Can query devices for “true state”
• Collect link error indications
• Collect light levels
• Use IP metrics
• Pretest circuits before handoff (won’t catch end interfaces)
• Leverage Force10 and collect utilization
From BoF at Vancouver JT

• Use optical switch to cycle through switch ports (~transponders)
• Use optical switch or attenuator to intentionally lower light levels near minimums (“margin testing”).
• Monitor pre-FEC errors too
From BoF at Vancouver JT

- Hook into control plane/middleware: when a connection is torn down, get a report on connection (errors, jitter, performed to specification)
  - Only if paths are fairly dynamic, and application to application
From BoF at Vancouver JT

- Think about applications
  - Why are paths being used/created?
  - Bulk transport: mostly loss
  - Interactive: mostly latency
  - Augmentation of IP infrastructure?
  - How often will paths change?