Studying Black Holes on the Internet with Hubble

Ethan Katz-Bassett, Harsha V. Madhyastha, John P. John, Arvind Krishnamurthy, David Wetherall, Thomas Anderson
University of Washington
RIPE, May 2008

This work partially supported by NSF
Global Reachability

- When an address is reachable from every other address
- Most basic goal of Internet, especially BGP
  - “There is only one failure, and it is complete partition” Clarke, Design Philosophy of the DARPA Internet Protocols
- Physical path $\implies$ BGP path $\implies$ traffic reaches
- **Black hole**: BGP path, but traffic persistently does not reach
Does Internet give global reachability?

- From use, seems to usually work
- Can we assume the protocols just make it work?

“Please try to reach my network 194.9.82.0/24 from your networks…. Kindly anyone assist.”

Operator on NANOG mailing list, March 2008.
Does Internet give global reachability?
Hubble System Goal

In *real-time* on a *global scale*, *automatically monitor* long-lasting reachability problems and *classify* causes.
Problem Seen by **Hubble** on Oct. 8, 2007

1. **Target Identification** – distributed ping monitors detect when the destination becomes unreachable
Problem Seen by Hubble on Oct. 8, 2007

1. Target Identification – distributed ping monitors
2. Reachability analysis – distributed traceroutes determine the extent of unreachability
Problem Seen by **Hubble** on Oct. 8, 2007

1. Target Identification – distributed ping monitors
2. Reachability analysis – distributed traceroutes
3. Problem Classification
   a) group failed traceroutes
Problem Seen by **Hubble** on Oct. 8, 2007

1. **Target Identification** – distributed ping monitors
2. **Reachability analysis** – distributed traceroutes
3. **Problem Classification**
   a) group failed traceroutes
   b) spoofed probes to isolate direction of failure

D to Y works!
Y to D fails!

D to Z works!
Z to D fails!
Architecture: Detect Problem

- Ping prefix to check if still reachable
  - Every 2 minutes from PlanetLab
  - Report target after series of failed pings
- Maintain BGP tables from RouteViews feeds
  - Allows IP $\Rightarrow$ AS mapping
  - Identify prefixes undergoing BGP changes as targets
Architecture: Assess Extent of Problem

- Traceroutes to gather topological data
  - Keep probing while problem persists
  - Every 15 minutes from 35 PlanetLab sites
- Analyze which traceroutes reach
  - BGP table to map addresses to ASes
  - Alias information to map interfaces to routers
To aid operators in diagnosis and repair:

- Which ISP contains problem?
- Which routers?
- Which destinations?
Architecture: Classify Problem

- Real-time, automated classification
- Find common entity that explains substantial number of failed traceroutes to a prefix
- Does not have to explain all failed traceroutes
- Not necessarily pinpointing exact failure
Classifying with Current Topology

- Group failed/successful traceroutes by last AS, router

Example: Router problem

- No probes reach $P$ through router $R$
- Some reach through $R$’s AS
- 28% of classified problems
Classifying with Historical Topology

- Daily probes from PlanetLab to all prefixes
- Gives baseline view of paths before problems
- Example: “Next hop” problem
- Paths previously converged on router $R$
- Now terminate just before $R$

- 14% of classified problems
Classifying with Direction Isolation

- Traceroutes only return routers on forward path
  - Might assume last hop is problem
  - Even so, require working reverse path
  - Hard to determine reverse path
- Internet paths can be asymmetric
- Isolate forward from reverse to test individually
- Without node behind problem, use spoofed probes
  - Spoof \textit{from S} to check forward path from S
  - Spoof \textit{as S} to check reverse path back to S
Classifying with Direction Isolation

- **Hubble** deployment on RON employs spoofed probes
  - 6 of 13 RON permit source spoofing
  - PlanetLab does not support source spoofing

Example: Multi-homed provider problem

- Probes through Provider $B$ fail
- Some reach through Provider $A$
- Like Cox/USC

- 6% of classified problems
Architecture: Summary of Approach

- Synthesis of multiple information sources
  - Passive monitoring of route advertisements
  - Active monitoring from distributed vantage points
- Historical monitoring data to enable troubleshooting
- Topological classification and spoofing point at problem
Evaluation

Target Identification
- How much of the Internet does Hubble monitor?

Reachability Analysis
- What percentage of the various paths to a prefix does Hubble analyze?

Problem Classification
- How often can Hubble identify a common entity that explains the failed paths to a prefix?

For further evaluation, please see NSDI 2008 paper.
How much does **Hubble** monitor?

Every 2 minutes:

- 89% of Internet’s edge address space
- 92% of edge ASes
What % of paths does Hubble monitor?

Compare with BGP paths of 447 RIPE peers (downhill ASes)

- PlanetLab’s restricted size and homogeneity limit uphill
- 90% of our failed traceroutes terminate within 2 AS hops of prefix’s origin
What % of paths does **Hubble** monitor?

Compare with BGP paths of 447 RIPE peers (downhill ASes)

BGP ASes: { AT&T, Sprint, Gigapop, Cenic, Intel }
Also on Traceroutes: { Sprint, Gigapop, Cenic, Intel }
Coverage for Intel prefix: 4 of 5 downhill ASes = 80%
What % of paths does Hubble monitor?

Compare with BGP paths of 447 RIPE peers (downhill ASes)

Overall for prefixes monitored by Hubble
- For >60% of prefixes, traverse ALL downhill RIPE ASes
- For 90% of prefixes, traverse more than half the ASes
How often can **Hubble** classify?

- 9 classes currently
  - Based on topology
  - Point to an AS and/or router

- Results from first week of February 2008

- Automatically classified 375,775/457,960 (82%) of problems as they occurred
How long do black holes last?

- 3 week study starting September 17, 2007
- 31,000 black holes involving 10,000 prefixes
- 20% lasted at least 10 hours!
- 68% were cases of partial reachability
How long do black holes last?

- 3 week study starting September 17, 2007
- 31,000 black holes involving 10,000 prefixes
- 20% lasted at least 10 hours!
- 68% were cases of partial reachability

Partial reachability:
- Can’t be just hardware failure
- Configuration/policy

Graph showing cumulative fraction of events vs. duration of reachability event (hrs).
Other Measurement Results

- Can’t find problems using only BGP updates
  - Only 38% of problems correlate with RouteViews updates

- Multi-homing may not give resilience against failure
  - 100s of multi-homed prefixes had provider problems like COX/USC, and ALL occurred on path TO prefix

- Inconsistencies across an AS
  - For an AS responsible for partial reachability, usually some paths work and some do not

- Path changes accompany failures
  - 3/4 router problems are with routers NOT on baseline path
Summary and Future Work

- **Hubble**: working real-time system
- Lots of reachability problems, some long lasting
- Baseline/ fine-grained data enable classification

Future:

- More classification/analysis, including cross-prefix
- Expand number/diversity of vantage points
- Make this a useful tool
How **Hubble** Can Help Operators

- Access to queriable real-time and historical traceroutes and reachability analysis?
- Notification of problems?
- Other problems or causes to look for?
- Please email *ethan@cs.washington.edu*
How Operators Can Help Hubble

- Validation/explanation of specific problems to help refine our techniques
- Traceroute servers/ host **Hubble** nodes
- Please email *ethan@cs.washington.edu*

http://hubble.cs.washington.edu

Uses iPlane, MaxMind, Google Maps