Pitfalls and Problems with Internet Data

David Moore

March 18th, 2002 - IPAM Large-Scale Communication Networks

dmoore @ caida.org

www.caida.org

Outline

- **Common themes of measurement problems**
- Available tools, data, and collection projects
- Problems by measurement type
- Other resources

Common themes

- timestamps & time synchronization
- lack of atomic measurements
- asymmetric paths
- not enough data and too much data
- bizarre behavior, misconfigurations, non-RFC, attacks
- measurement tools can lie
- representative data points - there is no “typical” on the Internet
- archived data often corrupt, truncated or poorly documented

Common themes:

Time

- clocks are of varying precision, accuracy and drift
- where are timestamps applied?
  - in hardware, at start or end of packet/cell?
  - in hardware, after a queue (FORE cards, NLANR .crl)?
  - in the device driver? in the kernel? in user space?
  - after buffers and queues? after mixing with other data?
- how is time synchronized between local (hardware timestamps) and remote clocks? NTP considered harmful?

Common themes:

Atomicity

- Can not atomically measure:
  - forward-path of an active measurement
  - routing table (when using ‘sho ip bgp’, updates better)
  - routing convergence
  - topology of the Internet
  - available bandwidth, router queues, utilization, etc
  - SNMP statistics

Common themes:

Asymmetry

- forward and reverse paths are asymmetric
- intermediary hops each have different reverse paths (particularly influential on intermediary RTTs)
- hot-potato (shortest-exit) routing
- routing policy is prefix-based
- passive taps may only see one direction (prefix-based)

Common themes:

Quantity of Data

- very easy to collect terabytes of data
- much harder to collect sufficient, meaningful, and representative data
- gigabytes of data hard to manage, search and process
- very hard to decide how to sample data
- tradeoff between detail of data and quantity that can be collected
- [first 40 bytes of packets at OC-48 >> 20MB/sec]

Common themes:

Expecting the Unexpected

- lots of non-RFC compliant implementations of everything (by major vendors)
- bugs in common OSes, libraries, applications
- misconfigurations can have huge effects
- attacks and probes generate strange traffic
- the Internet is where the impossible is possible, continuously

Common themes:

Tools Lie

- passive taps can drop, reorder and duplicate packets
- so can kernels/NICs for active measurements
- tools may not be robust to unexpected data
- Internet checksums are weak
- unknown network gear: L2 switches, passive caches
- tools pretend to be atomic when they aren't
- tools can't tell corrupt data (garbage in, garbage out)

Common themes:

The Typical Internet

- there is no “typical” on the Internet
- there is no location, link or path that represents all others
- there is no globally representative set of destinations for active measurement studies
- yesterday is not today, spring break is not mid-term
- some measures highly variable, makes sampling hard

Common themes:

Historic Data Lacking

- there is a lack of historic data for many questions
- existing archived data often corrupt, truncated and poorly documented
- due to size constraints, must make decisions as to what is collected and how much is kept
- no way to easily find what you want in large archives

Outline

- Common themes of measurement problems
- **Available tools, data, and collection projects**
- Problems by measurement type
- Other resources

Available Data & Tools (1)

[incomplete]

- Active performance/bandwidth measurements:
  - IEPM PingER
  - NLANR AMP
  - Surveyor, RIPE TTM
  - CAIDA Topology Measurement Project (topology focus)
- Active topology measurements:
  - CAIDA skitter
  - ISI Mercator, Cheswick's Internet mapping
  - NLANR AMP, Surveyor, RIPE TTM (RTT and one-way delay focus)

Available Data & Tools (2)

[incomplete]

- Passive header measurements:
  - tcpdump archive
  - NLANR PMA
  - CAIDA monitors
  - tcpdump wherever you want
  - Sprint POPs, AT&T dialup
- Passive SNMP/netflow measurements:
  - everyone has tons of mrtg graphs
  - local flow data, flowscan, rrdtool data files
- Some data proprietary/sensitive

Outline

- Common themes of measurement problems
- Available tools, data, and collection projects
- **Problems by measurement type**
- Other resources

Active Measurements

- Path related:
  - paths can change during measurements
  - non-atomic path determination
  - only forward paths (unless mesh measurement)
  - load-balancing and route changes can make spurious links
  - paths are asymmetric
  - intermediary hops each have different reverse paths (particularly influential on intermediary RTTs)

Active Measurements

- Topology probing related:
  - rfc1918 or other non-routed addresses for intermediaries
  - IP addresses of interfaces, not routers
  - different interfaces (inbound/outbound) may respond to TTL exceed depending on vendor (contrary to RFC)
  - primary path followed, may not see backup paths
  - straight topology says nothing about importance, usage, or capacity
  - some methodologies completely discard prefix-based routing policy
  - resulting graph is a “time smear” over entire collection interval

Routing Tables

- because of convergence times of peers, information for even a single prefix may be inconsistent for peers
- even if the internal networks of peers have converged, there may be additional delay getting that information to the monitor box
- since BGP only publishes the "best" path at each AS hop, information about secondary paths can be lost. So building an AS topology graph may not actually capture backup paths or peering
- people occasionally announce routes they shouldn't or deaggregate prefixes due to misconfiguration

Routing Tables

- sets of peers change over time
- some peers provide different sets of information: customer, aggregated, etc
- daily (or hourly) snapshots may not be representative of the majority of the day, just that time period
- different peers have different policies for prefix filtering
- time smear across collection period for `sho ip bgp' (query walks live data structure)
- historical archives (nlanr) are often truncated due to IOS TCP bugs & collection periods were 2+ hours at times

Passive Measurements

- path asymmetries - also “what link is being monitored”
- where are timestamps applied?
  - on the card, at start or end of packet/cell?
  - on the card, after a queue (FORE cards, NLANR .crl sites)?
  - in the device driver? in the kernel? in user space?
  - after buffers and queues? after mixing with other data?
- lost/reordered/duplicated packets:
  - really occurred on network
  - by card or full memory buffers
  - by BPF implementation
  - can’t capture to disk or process fast enough

Passive Measurements

- first N bytes capture can lose TCP options, other important session information
- address encoding (NLANR or normal tcpdpriv) prevents comparison between sites or traces
- hostname, geographic and routing mapping of addresses changes over time. currently only partial information about routing archived independently.
- clock drift on single card
- clock synchronization between multiple cards in single box or between remote boxes

Passive Measurements

- historical archives often:
  - are truncated
  - have corrupt blocks, especially before clock resets
  - file formats don't usually have way of reporting loss (and if they did, they might lie anyway)
  - poor time synchronization between directions
  - "broken" timestamp design:
    - FORE cards have hw and sw clocks, hw can wrap w/o sw noticing
    - FORE cards (and any normal NIC) have queues before timestamps
  - University sites on Abilene & vBNS links
  - have IP address encoding (changes per file)
  - no SNMP (mrtg) or other any sort of calibration/sanity

Auxiliary Measurements

- Bandwidth measurements:
  - Good news: many tools to pick from (ok, not many)
  - Bad news: they may all measure different things
  - Worse news: they don't work
- Interface -> router mapping:
  - works sometimes, blows up sometimes
- Geographical placement:
  - very difficult in general
  - may internally discard wildly bad data, leaving only hard to verify bad data
  - accuracy can vary widely in geographically/economically biased ways

Outline

- Common themes of measurement problems
- Available tools, data, and collection projects
- Problems by measurement type
- **Other resources**

Useful resources

- "Some Not-So-Pretty Admissions About Dealing With Internet Measurements" - Vern Paxson
  - http://www.icir.org/vern/talks/vp-nrdm01.ps.gz
- Index/summary page about various projects:
  - http://www.caida.org/analysis/performance/measinfra/
- These slides (eventually):
  - http://www.caida.org/outreach/presentations/

Cooperative Association for Internet Data Analysis
(CAIDA)
San Diego Supercomputer Center

Dept. Computer Science & Engineering
University of California, San Diego