### Internet measurements - Why?

- **Why measure?**
  - What's the need? Do we need to measure?
  - Can we just google it?
  - What is the motivation? Who gains?

- **Characterize behavior**
- **Performance**
- **Understand growth**
- **Properties**
- **Design issues**
  - Identify problems, suggest improvements
  - Understand protocol issues
  - Understand behavior
- **Manage the network**
  - Diagnose and correct
  - Traffic patterns
  - Nielsen rating for network performance

### Internet measurements - What?

- **What to measure?**
  - need to determine X (X - a property of the internet)

- **What property?**
  - Performance
    - Latency, throughput, link B/W, jitter, ...
  - Structure
    - Topology, relationships, characteristics
  - Behavior
    - Stability, failures, pathological, misconfigurations

- **Internet measurements - How?**

- **How to measure?**
  - need techniques and tools to measure, gather data, analyze data, present data
Network Measurement

- Measurement techniques
- Passive techniques
  - Collect traces and analyze or build special tools
  - Monitor application traffic, route announcements and analyze
  - Number of point to collect, single vantage point, several vantage points
  - Across ASes, or within AS (Gao 01, labovitz 97, labovitz 10)
- Active Techniques
  - Measure by probing (ping, traceroute, pathchar or build special tools e.g., Bolot93, paxson 97, rocketfuel 02)
  - Inject probes/packets or remove packets and measure impact
  - Inject faults at various points BGP beacon

Issues in measurement

- To collect traces need mechanism to capture packets and dump information
- Design packet filters, should be able to take in a spec and collect trace
- Data collection and analysis are independent
- Share traces for different experiments
- Manipulation and storage cost for traces
- How to analyze packet traces?
  - Time-sequence plots
  - Plot sequence number as a function of time
  - Repeat for different time scales

Internet measurements-Challenge?

- Routing/announcement and data interaction
  - Control plane vs data plane
- Causality: cause and effect analysis
- Coverage and completeness
- Internet is too big, not everything is public
  - Commercial interests conflict
  - Access to data difficult (ASes hide stuff)
- Huge amounts of data

Network measurement Papers

- Jean Bolot SIGCOMM 93
- Craig Labovitz SIGCOMM 2010 Internet Inter-domain Traffic
- Faloutsos et al., Power-law relationships of internet topology SIGCOMM 99
- Additional reading: IXP paper by Avia feldman SIGCOMM 13
Active Measurement

- Common tools:
  - ping
  - traceroute
  - scriptroute
  - Pathchar/pathneck/… BW probing tools
  - Fault injection: BGP beacons

End-to-end delay and loss behavior in the Internet

- Simple experiments, interesting way of arriving at the results
  - Presentation makes all the difference
  - Simple analysis and model fitting
  - Use end-to-end measurements to derive bottleneck link behavior
  - Use active measurement tool
  - Send UDP packets at regular intervals and measure RTT. Repeat experiments for different intervals
  - Data analysis using phase plots

Time-series plot

- For every sample, plot the $rtt_n$
- Not much to gain (except some have zero or packet loss)
- Average, max $rtt_n$, min $rtt_n$
- Instead plot $rtt_n$, $rtt_{n+1}$ $(x,y)$
- If constant, point.
- $D$, $\mu$
- $RTT = Prop$ delay + Queueing delay + Transfer delay
- Where does the time go? $D + \epsilon_{n+1} + Service$ time
- Probe size is $P$
- Both see same load, one sees load, interarrival time has a significant impact on probe compression.

Delay analysis

- $s_n =$ time of $n^{th}$ probe, $r_n =$ time at which echo of the $n^{th}$ probe is received, $rtt_n = r_n - s_n$
- $\delta = S_{n+1} - S_n$
- For various values of $\delta$ determine the distribution of $rtt_n$
- Need to figure the impact of load on the network
- Assume that there is very light load
- $\epsilon_{n+1} = \epsilon_n$ or $\epsilon_{n+1} = \epsilon_n + \epsilon$
- A phase plot should give us points on or around the diagonal
**Phase plot**

- $rtt_{n+1} = rtt_n + \varepsilon$

**Delay analysis**

- Assume a bottleneck link
- $rtt_n = \text{prop delay} + \text{queuing delay}$
- $rtt_n = D + \text{wait time} + \text{service time}$
  - $rtt_n = D + w_n + P/\mu$

- Heavy cross traffic
  - $w_{n+1} = w_n + B/\mu$ or $rtt_{n+1} - rtt_n = B/\mu - \delta$ or approx $B/\mu$

- During this time $k$ probes arrive
- For these probes the waiting times are $P/\mu$
  - $rtt_{n+k} - rtt_{n+(k-1)} = P/\mu - \delta$

**Bottleneck link capacity**

- Repeated experiments for different values of $\delta$ over a different path (umd, pittsburgh)
  - For $\delta = 8$ msec, $y$ intercept = 7.8 msec, $P/\mu = 0.2$
    - Given $P = 32$ bytes, we get $m = 1.28$ Mbps (Figure 5)
  - For $\delta = 50$ msec, phase plot consists of points scattered around $y=x$ line. No probe compression
    - Figure 6

- $\delta P/\mu$ probe saturates the queue
Packet size of cross traffic

- Assume cross traffic between two probes \( \delta \) units apart
- \( w_{n+1} \) = service time for \( b_n \) and \( w_n + \) wait time for \( w_n \)
- \( B = \mu (w_{n+1} - w_n + \delta) + P \)
- Interarrival time between successive probe returns is \( (r_{n+1} - r_n) \) is \( w_{n+1} - w_n + \delta \) which is the same as \( \text{rtt}_{n+1} - \text{rtt}_n + \delta \)

Finding packet size

- \( B/\mu + \delta \)
- \( B/\mu + \delta \)
- \( B/\mu + \delta \)
- \( B/\mu + \delta \)

Three cases

- Probes that were subject to probe compression and were serviced according to probe packet size
  - Difference in Wait times should be \( P/\mu \)
  - First peak
- Probes that were not subject to probe compression
  - Difference in Wait times should be \( \delta \)
  - Second peak
- Probes that were behind one internet packet
  - Difference in Wait times should be \( (B+P)/\mu \)
  - Third peak
- Use the value of third peak in Equation 6. Knowing bottleneck link, probe size, we can calculate internet packet size

Packet loss

- \( \text{rtt}_n \) was considered to be zero for a lost packet
- Look at the number of packets that were lost for various values of \( \delta \)
  - Smaller the value of \( \delta \) larger the loss
  - Find probability of correlated losses (burst size)
    - Further, conditional loss probability > loss probability
    - Because loss depends on queue length
  - Burst loss size is about 1 for most cases
- Interesting observation
- Implications for FEC and Go-back-N
Discussion

- Simple experiments, clever analysis leads to interesting results
- Phase plots, very creative
- Assumptions: single bottleneck
  - Needed for simple analysis
- How about internet of today
  - Asymmetric links
  - Traffic patterns may not pan out as in Jean Bolot's paper
- What about packet size distributions?

Active Measurement (2)

- Using traceroute
- Every packet has TTL
- At each hop, TTL is decremented
- When TTL becomes 0, packet is dropped and an ICMP error message is sent back to the source
- Use this to determine routers along the path from source to destination
- www.traceroute.org

Traceroute

tracing path from www.net.princeton.edu to 128.6.25.18 ...
traceroute to 128.6.25.18 (128.6.25.18), 30 hops max, 40 byte packets
1 gigagate1 (128.112.128.114) 0.530 ms 0.300 ms 0.343 ms
2 vgate1 (128.112.12.22) 4.360 ms 0.328 ms 0.348 ms
3 local1.princeton.mips.net (216.27.98.113) 23.504 ms 2.490 ms 2.493 ms
4 remote2.netip.mips.net (216.27.98.92) 4.324 ms 4.217 ms 4.180 ms
5 130.156.251.206 (130.156.251.206) 28.162 ms 4.967 ms 4.706 ms
6 198.151.130.133 (198.151.130.133) 5.180 ms 22.034 ms 5.067 ms
7 www.rutgers.edu (128.6.12) 5.391 ms 5.943 ms 5.316 ms
8 * * *
9 * * *
10 * * *
11 * * *
12 rags.rutgers.edu (128.6.25.18) 6.534 ms 5.813 ms 5.650 ms

Warning: multiple IP addresses found for www.iitb.ac.in, using 203.78.217.179
tracing path from www.net.princeton.edu to 203.78.217.179
traceroute to 203.78.217.179 (203.78.217.179), 30 hops max, 40 byte packets
1 gigagate1 (128.112.128.114) 0.926 ms 0.326 ms 0.345 ms
2 vgate1 (128.112.12.22) 0.375 ms 0.567 ms 0.349 ms
3 gi-6-0-226.hse1.phlapa02.paetec.net (209.92.27.33) 2.188 ms 2.616 ms 3.231 ms
4 ge-1-1-0-311.core02.paetec.net (169.130.105.19) 6.179 ms 2.550 ms 2.501 ms
5 * * *
6 * * *
7 as0.core02.asbnva01.paetec.net (169.130.81.233) 5.919 ms 6.013 ms
8 157.130.49.209 (157.130.49.209) 5.946 ms 6.119 ms 6.859 ms
9 9-0-0-3-2-3-4.ARIZ.ALTOR.NET (152.83.41.20) 6.950 ms 6.293 ms 6.234 ms
10 9-0-0-3-3-4.ARIZ.ALTOR.NET (152.83.23.177) 14.987 ms 15.087 ms 14.796 ms
11 0-so-0-0-5-5-2-3-2-1-2-ALTOR.NET (10.80.151.141) 87.187 ms 65.895 ms 65.641 ms
12 so-0-0-2-3-3-2-3-1-2-ALTOR.NET (10.80.37.195) 209.186 ms 210.645 ms
13 so-3-0-0-8-2-3-2-1-2-ALTOR.NET (10.80.37.114) 230.359 ms 208.795 ms 211.322 ms
14 18.17.245.0.customer.alter.net (203.78.210.150) 212.256 ms 211.991 ms 210.394 ms
15 * * *
16 * * *
17 * * *
18 * * *
19 * * *
Known issues with Traceroute

- Not all routers accept ICMP messages
  - Filtered or given low priority
- Can’t identify one-way failures (***)
  - Forward link or reverse link
- IP address of “time exceeded” packet may be the outgoing interface of the return packet

pathchar

- Traceroute includes Cumulative performance
- By subtracting successive link char you can get
- Subtract Y intercepts (RTT, RTT/2 gives one way delay)
- Subtract successive slopes (P/\mu) you get slope which is 1/ link bandwidth
- Note, X-axis is packet size (y=P/\mu)

Reference: Allen Downey Using pathchar to estimate internet link characteristics-SIGCOMM 1999

Extend to multiple links

- Extend analysis and methodology of Bolt’s paper
- Send probes of varying sizes, across several links
- Use traceroute over N hops; repeat expit for various packet sizes

Active measurements

- Multiple vantage points
- Send probes from multiple destinations
- Cover different paths from various S, D pairs
Questions on internet topology

- What does the internet connectivity graph look like?
  - Stub, tree, mesh
- What are the patterns
  - Invariants, growth (edges, core)
- Growth trends
  - Scale
- Realistic topologies for simulation
  - Is it random? Uniformly distributed links realistic?

Why do we need the topology?

- Understand the macroscopic properties of the Internet physical structure
- Network management
- Topology-aware algorithms
- Simulation and topology generation tools

Active measurement

- Use Hop-limited probes from a single vantage point
- Informed random probing
  - Probe multiple IP addresses, adjacent prefixes
- Use Source Routing (specify intermediate nodes)
  - An option in the IP packet that forces routes to use the IP addresses specified in the packet
- Degree distribution
- Hop-pair distribution

Heuristics for internet map discovery, Infocom 2000

Degree, hop-pair distribution
Power-Laws and Internet

- What can be said about the structure of the internet?
- Average out-degree
- Average # of nodes within K hops
- Power law
  \[ y = k x^\alpha \log(y) = \log(k) + \alpha \log(x) \]
- On a log log scale the function is linear

March 16: WSJ article

- Why the industry is evolving this way is rooted in balance sheets. Over the past two years, Apple Inc., Oracle Corp., Google Inc., Microsoft Corp. and six other large tech companies have generated $68.5 billion in new cash, compared with just $13.5 billion for the other 65 tech companies in the S&P 500 Index combined, according to a Wall Street Journal analysis of data provided by Capital IQ.

<table>
<thead>
<tr>
<th>No.</th>
<th>Name</th>
<th>Net worth (USD)</th>
<th>Age</th>
<th>Citizenship</th>
<th>Residence</th>
<th>Sources of wealth</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Carlos Slim Helú and family</td>
<td>$53.5 billion</td>
<td>70</td>
<td>Mexico</td>
<td>Mexico</td>
<td>Telmex, América Móvil</td>
</tr>
<tr>
<td>2</td>
<td>William Henry Gates III</td>
<td>$53.0 billion</td>
<td>54</td>
<td>United States</td>
<td>United States</td>
<td>Microsoft</td>
</tr>
<tr>
<td>3</td>
<td>Warren Buffett</td>
<td>$47.0 billion</td>
<td>79</td>
<td>United States</td>
<td>United States</td>
<td>Berkshire Hathaway</td>
</tr>
<tr>
<td>4</td>
<td>Mukesh Ambani</td>
<td>$29.0 billion</td>
<td>52</td>
<td>India</td>
<td>India</td>
<td>Reliance Industries</td>
</tr>
<tr>
<td>5</td>
<td>Lakshmi Mittal</td>
<td>$26.7 billion</td>
<td>59</td>
<td>India</td>
<td>India</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Lawrence Ellison</td>
<td>$26.0 billion</td>
<td>65</td>
<td>India</td>
<td>India</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Bernard Arnault</td>
<td>$27.9 billion</td>
<td>41</td>
<td>India</td>
<td>India</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Eike Batista</td>
<td>$27.0 billion</td>
<td>52</td>
<td>Brazil</td>
<td>Brazil</td>
<td>EBX Group</td>
</tr>
<tr>
<td>9</td>
<td>Amancio Ortega</td>
<td>$25.0 billion</td>
<td>74</td>
<td>Spain</td>
<td>Spain</td>
<td>Inditex Group</td>
</tr>
<tr>
<td>10</td>
<td>Karl Albrecht</td>
<td>$23.5 billion</td>
<td>90</td>
<td>Germany</td>
<td>Germany</td>
<td>Aldi Süd</td>
</tr>
</tbody>
</table>

Power-laws

- Found in many different areas
- Economics (wealth distribution)
- File sizes, process (thread) execution times
- Sociology
- Probability distribution (pareto, weibull, log normal)
- Intuition: Large occurrences are very rare and small occurrences are very common

On power law relationships of the internet topology-SIGCOMM1999
Power law 1: rank exponent

- Study out-degree of nodes
- Definition: \( \text{rank, } rv \) of a node \( v \) is its index in the order of decreasing out-degree.

\[
\text{Exponent } = \text{slope } R = -0.74
\]

Power-law 1: rank \( R \)

- The plot is a line in log-log scale
- Rank exponent \( R \) is the slope of the line \( \text{deg vs rank} \)

Power law 2: Outdegree \( O \)

- Distribution of out-degree of the graph
- The frequency of an out-degree \( d \), \( f_d \), is the number of nodes with out-degree \( d \).
Power-law 2: outdegree \( O \)

- The plot is linear in log-log scale
- Out degree exponent \( O \) is the slope of \( f \) vs. \( O \)

\[
\text{Outdegree} \\
\text{Frequency}
\]

\( O = -2.15 \)

Power law 3: hop-plot

- Distribution of number of nodes within \( h \) hops
- Include self pairs (i,j is pair so is j,i)

<table>
<thead>
<tr>
<th>Hop</th>
<th>P(h)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>70</td>
</tr>
<tr>
<td>4</td>
<td>88</td>
</tr>
<tr>
<td>5</td>
<td>100</td>
</tr>
</tbody>
</table>

Power law 3- hop-lot

- Hop-Plot Exponent \( H \): The total number of pairs of nodes \( P(h) \) within \( h \) hops, is proportional to the number of hops to the power of a constant, \( H \), \( \delta \) is the diameter of the graph

\[
P(h) \propto h^H, \quad h \ll \delta
\]

Power-law 3: hopplot \( H \)

- Pairs of reachable nodes as a function of hops
- Simpler: neighborhood size vs hops

\[
H = 2.83
\]
Internet Inter-Domain topology

- Internet Inter-domain traffic by Craig Labovitz, et.al., SIGCOMM 2010
- Anatomy of a large European IXP by Anja Feldmann, et.al., SIGCOMM 2013
- Pictures/graphs from slides of SIGCOMM presentations

Change in carrier traffic demands

<table>
<thead>
<tr>
<th>Rank</th>
<th>2007 Top Ten</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>AT&amp;T</td>
<td>8.77</td>
</tr>
<tr>
<td>2</td>
<td>Sprint</td>
<td>2.35</td>
</tr>
<tr>
<td>3</td>
<td>Verizon</td>
<td>2.33</td>
</tr>
<tr>
<td>4</td>
<td>Qwest</td>
<td>2.77</td>
</tr>
<tr>
<td>5</td>
<td>Sprint</td>
<td>2.15</td>
</tr>
<tr>
<td>6</td>
<td>Verizon</td>
<td>2.24</td>
</tr>
<tr>
<td>7</td>
<td>Sprint</td>
<td>1.66</td>
</tr>
<tr>
<td>8</td>
<td>AT&amp;T</td>
<td>1.42</td>
</tr>
<tr>
<td>9</td>
<td>Verizon</td>
<td>1.35</td>
</tr>
<tr>
<td>10</td>
<td>Qwest</td>
<td>1.11</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Rank</th>
<th>2009 Top Ten</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>AT&amp;T</td>
<td>8.51</td>
</tr>
<tr>
<td>2</td>
<td>Google</td>
<td>3.12</td>
</tr>
<tr>
<td>3</td>
<td>AT&amp;T</td>
<td>2.12</td>
</tr>
<tr>
<td>4</td>
<td>Google</td>
<td>2.24</td>
</tr>
<tr>
<td>5</td>
<td>Verizon</td>
<td>1.32</td>
</tr>
<tr>
<td>6</td>
<td>Verizon</td>
<td>1.32</td>
</tr>
<tr>
<td>7</td>
<td>AT&amp;T</td>
<td>1.13</td>
</tr>
<tr>
<td>8</td>
<td>Verizon</td>
<td>1.13</td>
</tr>
<tr>
<td>9</td>
<td>AT&amp;T</td>
<td>1.02</td>
</tr>
<tr>
<td>10</td>
<td>Verizon</td>
<td>1.02</td>
</tr>
</tbody>
</table>

* In 2007, top ten match “Tier-1” ISPs (e.g., Wikipedia)
* In 2009, global transit carry significant traffic volumes
  - But Google and Comcast join the list
  - And a significant percentage of ISP traffic is Google transit

Impact of cloud

- In 2007, thousands of ASNs contributed 50% of content
- In 2006, 100 ASNs contribute 50% of all Internet traffic

Traditional Internet model

Exchanges and business interactions
A new Internet model

- Flatter and much more densely interconnected Internet
- Disintermediation between content and “cloud” networks
- New commercial models between content, consumer and transit

Even more recent model

- IXPs central component
- Lots of local peering – rich fabric
- Even flatter AS topology than assumed

IXP Architecture

IXP diurnal traffic patterns

- Traffic Volume: Same as Tier-1 ISPs
- Clear daily patterns
Flatter topology

Difficulties in topology discovery
- Shared media
- Backup links
- Mobile access
- Router Identification and annotation
- Alias resolution
- Completeness Validation

Currently, none of them is completely solved!

Source: BITAG