Internet Traffic and Content Consolidation

Craig Labovitz
Chief Scientist, Arbor Networks

S. Iekel-Johnson, D. McPherson
Arbor Networks, Inc.

J. Oberheide, F. Jahanian
University of Michigan
Talk Outline

- Describe two-year traffic measurement study
- The “original” Internet topology
- The emerging new Internet
- Application transport and the end of end-to-end
- A few words on IETF implications
Two Year Study of Inter-domain Traffic

- Leverage large, widely deployed commercial Internet monitoring infrastructure
- Global deployment across 110+ ISPs / Content Providers
  - Near real-time traffic and routing statistics (14 Tbps)
  - Participation voluntary and all data sources are anonymous
  - Largest study of its kind

Graphic not an accurate representation of current ATLAS deployments
Study Details

- Within a given ISP, commercial probe infrastructure
  - Monitors NetFlow / Jflow / etc and routing across possible hundreds of routers
  - Probes topology aware of ISP, backbone and customer boundaries
  - Routers typically include most of peering / transit edge
  - Some deployments include portspan / inline appliances

- Deployments send anonymous XML file to central servers
  - Includes self-categorization of primary geographic region and type
  - Data includes coarse grain anonymized traffic engineering statistics

- Introduced at NANOG 47 academic paper under review, Arbor blog provides ongoing related bits
Traffic Measurements

- **Inter-domain traffic volumes**
  - Estimate directly monitoring 25% all inter-domain traffic
  - Believe data representative of global inter-domain traffic
  - Validate predictions based on data (using 12 known ISP traffic demands)

- **Does NOT measure**
  - Number of web hits, tweets, transactions, customers, etc.
  - Internal / private customer traffic (e.g. VPNs, IPTV)
  - ISP success nor profitability

Measurement Confidence
Original Internet (1995 – 2007)

- Textbook diagram (still taught today)
- Hierarchical, relatively sparsely inter-connected Internet
- Mostly accurate until recently (modulo a few name changes over the years)
Market Forces Reshape Traffic and Connectivity

Revenue from Internet Transit
Source: Dr. Peering, Bill Norton

Revenue from Internet Advertisement
Source: Interactive Advertising Bureau
Largest Carriers: Then and Now

<table>
<thead>
<tr>
<th>Rank</th>
<th>2007 Top Ten</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ISP A</td>
<td>5.77</td>
</tr>
<tr>
<td>2</td>
<td>ISP B</td>
<td>4.55</td>
</tr>
<tr>
<td>3</td>
<td>ISP C</td>
<td>3.35</td>
</tr>
<tr>
<td>4</td>
<td>ISP D</td>
<td>3.2</td>
</tr>
<tr>
<td>5</td>
<td>ISP E</td>
<td>2.77</td>
</tr>
<tr>
<td>6</td>
<td>ISP F</td>
<td>2.6</td>
</tr>
<tr>
<td>7</td>
<td>ISP G</td>
<td>2.24</td>
</tr>
<tr>
<td>8</td>
<td>ISP H</td>
<td>1.82</td>
</tr>
<tr>
<td>9</td>
<td>ISP I</td>
<td>1.35</td>
</tr>
<tr>
<td>10</td>
<td>ISP J</td>
<td>1.23</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>ISP A</td>
<td>9.41</td>
</tr>
<tr>
<td>2</td>
<td>ISP B</td>
<td>5.7</td>
</tr>
<tr>
<td>3</td>
<td>Google</td>
<td>5.2</td>
</tr>
<tr>
<td>4</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Comcast</td>
<td>3.12</td>
</tr>
<tr>
<td>7</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>-</td>
<td></td>
</tr>
</tbody>
</table>

Based on analysis of anonymous ASN (origin/transit) data (as a weighted average % of all Internet Traffic). Top ten has NO direct relationship to study participation.

- 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 A traffic is Google transit
The New Internet

- Flatter and much more densely interconnected Internet
  - Significant routing, traffic, security, economic, implications
- Disintermediation between content and eyeball networks
- New commercial models between content, consumer and transit

Settlement Free
Pay for BW
Pay for access BW
Consolidation of Content (*Grouped Origin ASN*)

- In 2007, thousands of ASNs contributed 50% of content
- In 2009, 150 ASNs contribute 50% of all Internet traffic
- Approximates a power law distribution
Case Study: Google

- Over time Google absorbs YouTube traffic
- As of July 2009, Google accounts for 6% of all Internet inter-domain traffic
- Google the fastest growing ASN group
Google Dense Interconnection

- Over time, Google increasingly using direct peering with tier2/3 and eyeball networks
- As of February 2010, more than 60% of Google traffic does not use transit
  - Remainder largely global transit carriers
- These numbers do not include GGC
Other Case Studies

- Rapid rise of new content players, e.g.
  - CDNs
  - Facebook
  - Baidu
  - Apple / MSFT

- Change in traffic patterns and business strategies of consumer networks
What’s Happening?

- **Commoditization of IP and Hosting / CDN**
  - Drop price of wholesale transit
  - Drop price of video / CDN
  - Economics and scale drive enterprise to “cloud”

- **Consolidation**
  - Bigger get bigger (economies of scale)
  - e.g., Google, Yahoo, MSFT acquisitions

- **Success of bundling / Higher Value Services**
  - Triple and quad play, etc.

- **New economic models**
  - Paid content (ESPN 360), paid peering, etc.
  - Difficult to quantify due to NDA / commercial privacy

- **Disintermediation**
  - Direct interconnection of content and consumer
  - Driven by both cost and increasingly performance
Applications

<table>
<thead>
<tr>
<th>Rank</th>
<th>Application</th>
<th>2007</th>
<th>2009</th>
<th>Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Web</td>
<td>41.68%</td>
<td>52.00%</td>
<td>24.76%</td>
</tr>
<tr>
<td>2</td>
<td>Video</td>
<td>1.58%</td>
<td>2.64%</td>
<td>67.09%</td>
</tr>
<tr>
<td>3</td>
<td>VPN</td>
<td>1.04%</td>
<td>1.41%</td>
<td>35.58%</td>
</tr>
<tr>
<td>4</td>
<td>Email</td>
<td>1.41%</td>
<td>1.38%</td>
<td>-2.13%</td>
</tr>
<tr>
<td>5</td>
<td>News</td>
<td>1.75%</td>
<td>0.97%</td>
<td>-44.57%</td>
</tr>
<tr>
<td>6</td>
<td>P2P (*)</td>
<td>2.96%</td>
<td>0.85%</td>
<td>-71.28%</td>
</tr>
<tr>
<td>7</td>
<td>Games</td>
<td>0.38%</td>
<td>0.49%</td>
<td>28.95%</td>
</tr>
<tr>
<td>8</td>
<td>SSH</td>
<td>0.19%</td>
<td>0.28%</td>
<td>47.37%</td>
</tr>
<tr>
<td>9</td>
<td>DNS</td>
<td>0.20%</td>
<td>0.17%</td>
<td>-15.00%</td>
</tr>
<tr>
<td>10</td>
<td>FTP</td>
<td>0.21%</td>
<td>0.14%</td>
<td>-33.33%</td>
</tr>
<tr>
<td></td>
<td>Other</td>
<td>2.56%</td>
<td>2.67%</td>
<td>4.30%</td>
</tr>
<tr>
<td></td>
<td>Unclassified</td>
<td>46.03%</td>
<td>37.00%</td>
<td>-19.62%</td>
</tr>
</tbody>
</table>

(*) 2009 P2P Value based on 18% Payload Inspection
Weighted average percentage of all Internet traffic using well-known ports

- **Growing volume of Internet traffic uses port 80 / 443**
  - Includes significant video component and source of most growth
- **Unclassified includes P2P and video**
  - Payload matching suggests P2P at 18%
  - P2P is fastest declining
The End of End-to-End?

- Growing dominance of web as application front-end
- Plus burden of ubiquitous network layer security policies
- Results in growing concentration of application traffic over a decreasing number of TCP / UDP ports
  - Especially port 80
  - Especially video
P2P

- In 2006, P2P one of largest threats facing carriers
  - Significant protocol, engineering and regulatory effort / debate
- In 2010, P2P fastest declining application group
  - Trend in both well-known ports and payload based analysis
  - Still significant volumes
  - Slight differences in rate of decline by region (i.e. Asia is slower)
P2P Surpassed by Direct Download

- Normally study lacks visibility into hosting customers
- Mega [Upload|Video|Erotic] is an exception
  - Carpathia small hosting company by traffic volume in Fall 2008
  - Mega becomes Carpathia customer in November 2008
  - Carpathia Hosting grows overnight to more than 0.5% of all traffic
IPv6

- IPv6 miniscule percentage of Internet traffic (.04 %)
- Still relatively little native IPv6 peering between large carriers
- Few carriers with v6 traffic visibility (i.e. flow)
- Tunneled IPv6 shows growth since IPv6
  - Due to uTorrent
  - And Hurricane Electric global Teredo deployment (see blog)
Internet Size / Growth

- In 2009, Internet (inter-domain) roughly ~45 Tbs
  - And growing at 45% per year

- Significant, but no “Exaflood”
  - Followed MINTS methodology for AGR
  - Used 10 known ISP totals (MRTG / Flow based) to extrapolate Internet total

<table>
<thead>
<tr>
<th>Estimate</th>
<th>Observatory</th>
<th>ISP Survey</th>
<th>Cisco</th>
<th>MINTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traffic Volume Per Month</td>
<td>9 exabytes</td>
<td>N/A</td>
<td>9 exabytes</td>
<td>5-8 exabytes</td>
</tr>
<tr>
<td>Traffic Annual Growth Rate</td>
<td>44.5%</td>
<td>35-45%</td>
<td>50%</td>
<td>50-60%</td>
</tr>
</tbody>
</table>
IETF Implications

- Increasingly dense Internet and impact on routing scalability and convergence
- Slow IPv6 deployment highlights need for alternative transition mechanisms
- The “end” of end-to-end
  - Increasing impact of firewall, NAT
  - Silo’ed ecosystems
Conclusion

- Internet is at an inflection point
- Focus shifting from transmission to content
  - Battle for access to eyeballs (and control of content)
  - Transit is commoditized and devalued
  - New focus on datacenters and co-location (caches)
- New technologies reshaping definition of Internet
  - “Web” / Desktop Applications, Cloud computing, CDN
- Changes mean significant new commercial, security and engineering challenges
- This is just the beginning…