Network Traffic Evolution

Prof. Anja Feldmann, Ph.D.
The Internet

- As with any new facility, there will be a period of very light usage until the community of users experiments with the network and begins to depend upon it. One of our goals must be to stimulate the immediate and easy use by a wide class of users.

- Steve Crocker; Host Software; RFC 1; 7 April 1969.
The Internet in the early days

- Applications such as telnet, FTP, Usenet newsgroups, IRC, MUDs and mailing lists to empower a community to share information

- Marshall McLuhan foresees the global village
Internet history: ARPANET

- 10:30 PM on October 29'th, 1969 the first packets between 2 ARPANET nodes
- Sep. 1971 remote access via Terminal Interface Processor (TIP)
- Dec. 1972. 24 sites on the ARPANET
- Dec. 1973. 37 sites on the ARPANET, including satellite link from California to Hawaii.
- 1973, University College of London, England and Royal Radar Establishment, Norway connected to the ARPANET.
- Jun. 1974 62 computers connected to the ARPANET.
- Mar. 1977 111 computers on the ARPANET.
- 1983 an unclassified military only network, MILNET, split off from the ARPANET.
- 1985. ARPANET gateways to external networks across North America, Europe, and in Australia, and the Internet was global in scope.
- 1990. The ARPANET was retired in 1990. NSFNET took over for most universities.
Internet history: NSFNET 1987-1995

- As part of the NSFNET contract, Krol authored the *Hitchhiker's Guide to the Internet* to provide a help manual for NSFNET users, providing one of the first comprehensive surveys of the Internet.
Internet history: NSFNET stats

- By January, 1992, the NSFNET traffic exceeded 12 billion packets (1 trillion bytes) of traffic a month. By November the traffic had doubled, and NSFNET was connected to more than 7,500 networks, one third of which were outside the United States.

- In December, 1992, the NSFNET backbone was completely converted to a T3 or 44.736 Mbps capacity, capable of transmitting 4 and a half million characters a second.

- In 1994, the traffic on NSFNET broke the 10 trillion bytes a month level.
Today’s Internet


The New Internet
Questions to be answered

- What creates the structure of the Internet
- How does an ISP topology look like?
- How is reachability guaranteed?
- Which route does my traffic take?
  - Is it the “best” route?
  - Which of multiple servers should one choose?
  - How does traffic flow?
  - What is the traffic matrix
- Who is providing the content?
  - Are CDNs dominating?
  - Is Google dominating?
Network Dynamics: User population

- Number of Internet hosts
  - Early 1989: 80,000
  - Early 1992: 727,000
  - Oct. 1993: 2,056,000
  - Late 1996: 10,000,000
  - ...

- Internet traffic volume (Merit; Inc.)
  - March 1991: $1.3 \times 10^{12}$ bytes/month
  - March 1994: $1.1 \times 10^{13}$ bytes/month
  - ...

...
Internet users

**Internet Users in the World by Geographic Regions - 2011**

- Asia: 1016.8 million
- Europe: 500.7 million
- North America: 273.1 million
- Latin America / Caribbean: 235.8 million
- Africa: 139.9 million
- Middle East: 77.0 million
- Oceania / Australia: 23.9 million

Source: Internet World Stats - www.internetworldstats.com/stats.htm
Estimated Internet users are 2,267,233,742 on December 31, 2011
Copyright © 2012, Mniwatts Marketing Group
Internet penetration rates

World Internet Penetration Rates by Geographic Regions - 2011

- North America: 78.6%
- Oceania / Australia: 67.5%
- Europe: 61.3%
- Latin America / Caribbean: 39.5%
- Middle East: 35.6%
- World, Avg.: 32.7%
- Asia: 26.2%
- Africa: 13.5%

Source: Internet World Stats - www.internetworldstats.com/stats.htm
Penetration Rates are based on a world population of 6,930,055,154 and 2,267,233,742 estimated Internet users on December 31, 2011.
Copyright © 2012, Miniwatts Marketing Group
Application mix?

CacheLogic Research | Internet Protocol Trends 1993 to 2006

Percentage of Internet Traffic

Year

Email
FTP
P2P
Web
Netdynamics – „Killer application“

- WWW and the Internet
  - 1993: ... Hardly any WWW traffic on the Internet
  - 1994: ... About 10% of total Internet traffic is WWW
  - 95/96: ... Up to 60-70% of overall Internet traffic is WWW
  - ...???????...
## Incoming AT&T Traffic by Port

(18 hours of traffic to AT&T dial clients on July 22, 1997)

<table>
<thead>
<tr>
<th>Name</th>
<th>port</th>
<th>% bytes</th>
<th>% packets</th>
<th>bytes per packet</th>
</tr>
</thead>
<tbody>
<tr>
<td>world-wide-web</td>
<td>80</td>
<td>56.75</td>
<td>44.79</td>
<td>819</td>
</tr>
<tr>
<td>netnews</td>
<td>119</td>
<td>24.65</td>
<td>12.90</td>
<td>1235</td>
</tr>
<tr>
<td>pop-3 mail</td>
<td>110</td>
<td>1.88</td>
<td>3.17</td>
<td>384</td>
</tr>
<tr>
<td>cuseeme</td>
<td>7648</td>
<td>0.95</td>
<td>1.85</td>
<td>333</td>
</tr>
<tr>
<td>secure web</td>
<td>443</td>
<td>0.74</td>
<td>0.79</td>
<td>603</td>
</tr>
<tr>
<td>internet chat</td>
<td>6667</td>
<td>0.27</td>
<td>0.74</td>
<td>239</td>
</tr>
<tr>
<td>file transfer</td>
<td>20</td>
<td>0.65</td>
<td>0.64</td>
<td>659</td>
</tr>
<tr>
<td>domain name</td>
<td>53</td>
<td>0.19</td>
<td>0.58</td>
<td>210</td>
</tr>
</tbody>
</table>

World Wide Web traffic dominates traffic mix
# MWN traffic by port

(24 hours of traffic to/from MWN clients in 2006)

<table>
<thead>
<tr>
<th>Port</th>
<th>% Conns</th>
<th>% Success</th>
<th>% Payload</th>
</tr>
</thead>
<tbody>
<tr>
<td>Web 80</td>
<td>70.82%</td>
<td>68.13%</td>
<td>72.59%</td>
</tr>
<tr>
<td>445</td>
<td>3.53%</td>
<td>0.01%</td>
<td>0.00%</td>
</tr>
<tr>
<td>Web 443</td>
<td>2.34%</td>
<td>2.08%</td>
<td>1.29%</td>
</tr>
<tr>
<td>SSH 22</td>
<td>2.12%</td>
<td>1.75%</td>
<td>1.71%</td>
</tr>
<tr>
<td>Mail 25</td>
<td>1.85%</td>
<td>1.05%</td>
<td>1.71%</td>
</tr>
<tr>
<td>1042</td>
<td>1.66%</td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td>1433</td>
<td>1.06%</td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td>135</td>
<td>1.04%</td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td>&lt; 1024</td>
<td>83.68%</td>
<td>73.73%</td>
<td>79.05%</td>
</tr>
<tr>
<td>&gt; 1024</td>
<td>16.32%</td>
<td>4.08%</td>
<td>20.95%</td>
</tr>
</tbody>
</table>
Application mix?

CacheLogic Research | Internet Protocol Trends 1993 to 2006

Percentage of Internet Traffic

Year

Email
FTP
P2P
Web
Application mix – different locations?

P2P more likely* at 18%

*Based on limited dataset using payload inspection
A Large European IXP

- Daily traffic (2013): ~14 PB
- sFlow export
- Random Sampling 1/16K Packets
- Snaplen 128 Bytes
- Weekly Snapshosts dating back to 2011

Sample snapshot (2013-09, 496 networks, 1 week)

<table>
<thead>
<tr>
<th>packets sampled</th>
<th>bytes sampled</th>
<th>IPv4 / IPv6</th>
<th>TCP / UDP</th>
</tr>
</thead>
<tbody>
<tr>
<td>9.3B</td>
<td>5.9TB</td>
<td>99.37% / 0.63%</td>
<td>83.7% / 16.3%</td>
</tr>
</tbody>
</table>
HTTP(S) dominates ~ 67%
Other applications (e.g., RTMP, mail, news) ~ 6%
BitTorrent/BT/P2P likely ~ 22%
Unclassified ~ 5%
Application Mix: Over Time

- Diurnal patterns, e.g., P2P dominates in off-hours
- Historical view shows increasing dominance of HTTP(S) and significant HTTPS increase in 2013.
Application Mix: Network Type

- Content/Content distribution networks (CDN) almost 100% HTTP
- HTTPS increase caused by a few networks
- Peer to peer not only between Eyeballs! Hoster/IaaS too!

A closer look can result in different appmix!
Aggregate mix not representative of single link
Many links just have one dominant protocol
Business type of the AS gives hints on app mix
Application Mix: Per Link

- Aggregate mix not representative of single link
- Many links just have one dominant protocol
- Business type of the AS gives hints on app mix

content <> eyeball: HTTP
Application Mix: Per Link

- Aggregate mix not representative of single link
- Many links just have one dominant protocol
- Business type of the AS gives hints on app mix

eyeball <> eyeball: Peer to peer (p2p)
Application Mix: Per Link

- Aggregate mix not representative of single link
- Many links just have one dominant protocol
- Business type of the AS gives hints on app mix

hoster/IaaS: diverse application mix
Questions to be answered

- Why do the results vary by study?
- What does it mean to be representative?
- How can one determine the application mix?
- What about user privacy?
P2P vs. Direct Download Providers

Global P2P Trends

P2P Replaced by Direct Download
Questions to be answered

- Why do the results vary by study?
- What does it mean to be representative?
- How can one determine the application mix?
- What about user privacy?
- What about individual applications?
  - P2P
  - Gaming
  - Online Social Networks
  - ...

P2P
Gaming
Online Social Networks