Internet-scale Experimentation

The challenges of large-scale networked system experimentation and measurements
The state of affairs

- An ever growing Internet
  - ~3 billion people
  - 15 billion devices connected
  - 10 thousands ISPs
  - >52 thousands networks (ASes)

- Tons of money at play
  - Alphabet 3rd Q 2015 revenues - $18.7 billions (+13% per year)
The state of affairs

- Society’s increased dependency on ...
  - More, ever-larger Internet-scale systems
    - FB, Skype, Twitter, Google, Akamai, Amazon, Netflix ...
  - Facebook’s 1.44 billion monthly users
    - Average time in FB 20’/day
    - Or 20% of all online time

- Yet, we still
  - Can’t predict these systems’ behaviors
  - or trust their security, performance, resilience, ...
  - Don’t know how the network underneath looks like
  - ...
Experimentation

- Observe, measure, build and test ideas in working systems
  - To test our theories and pose new questions
  - To validate our assumptions
  - To understand our large and complex systems
  - ...

- But ...
  - How to do experimentation at Internet-scale?
  - What’s representative? reproducible? ethical? ...

“Experiments ... the source of most questions, the final test for all answers”
~ R. Feynman
Our goal and road map

- Experiments in today’s network
- Strategies and good practices
- Edge network perspective: Network positioning
- Application performance: Public DNS and CDNs
- Moving up the stack: Broadband reliability
A bit of history, for context – Early days

- ~1960 ARPA sponsored research on computer networking to let researchers share computers remotely
  - Electronic computers were scarce resources
  - Renting an IBM System/360 - $5k/month ($35k/month 2016)

- 1969 – First four ARPANET nodes connected
  - UCLA, Stanford Research Institute, UCSB, U. of Utah
  - Key design decision – packet switching
A bit of history – Early days

- From 1975 to 1980s
  - Successful ARPANET ~ 100 nodes
  - ARPA research on packet switching over radio and satellite
  - New LANs connected via gateways
  - TCP/IP conversion in 1983
  - Autonomous Systems and backbone AS for scalability
A bit of history – NSF takes over

- Late 1980s NSF takes over
  - NSF work on expanding the backbone
- NSF encourage development of regional networks
  - Three tiers: backbone, regional, enterprise

- Enterprises were building TCP/IP networks and wanted to connect them
  - NSF charter prohibited them from using NSFNET
  - 1987 first commercial ISP, many follow shortly
A bit of history – Commercial operation

- By 1990 service providers where interconnected
  - Congress lets NSFNET interconnect with commercial networks
  - By 1995, NSFNET was retired
    - No single default backbone anymore
    - Many backbones interconnected trough Network Access Points

- ~1995 Web
  - Easier to use Internet
  - Million of non-academic users

- Now …
  - Large ISPs interconnected, regional ISPs, mid-size ISP and eyeballs
Internet as a set of ASes

- **Internet**
  - A collection of separately, usually competing, managed networks

- **Autonomous system (AS)**
  - Set of network elements under a single organization’s control
  - 1 ISP, can operate N ASes; no AS is managed by >1 ISP

- **Ases exchange traffic at peering points**
  - Connections – a link between “gateway” routers in each AS
Classical Internet model
Updated Internet model

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

Labovitz et al., SIGCOMM 2010
Some key principles inferred from early design decisions

**Decentralized design and operation**
- A loose interconnection of networks, not really “one” network
- Connecting a node to the Internet does not require the consent of any global entity

**IP hourglass or IP over everything**
- Internet overarching goal – to provide connectivity – IP is key
- Easy to incorporate new applications and new communication media
Design principles of the Internet

- **Stateless switching**
  - Switches are expected to be stateless wrt connections
  - Forward decision based on packet IP’s header and routing table
  - Results in very simple routers, … related to …

- **End-to-end**
  - Insight – many network functions require cooperation from end-systems for correct and complete operation
    - So, don’t try to do it within the network
  - Challenges to end-to-end: untrustworthy world, more demanding apps (use of CDNs), less sophisticated users, …
Design principles and measurements

- Decentralized design and operation
  - Hard to learn the current configuration of the Internet
- IP over everything
  - Complicates measuring hiding details of physical medium
- Stateless switching
  - … routers don’t capture or track anything of the traffic going by
- End-to-end argument
  - Lack of instrumentation at many points in the network, as it encourages the design of network elements with minimal functionality
In sum

- A decentralized and distributed architecture
- Without support for third-party measurements

So, measurement efforts

- have limited visibility (and shrinking)
- rely on hacks, rarely validated
- More often that not … what we can measure is not what we want to measure and, worst, what we think we are measuring
Measurement and experimentation

Given this overall picture …

- Where should we place our vantage points?
- At what layers of the stack?
- Can we get measurement control & scalability?
- … repeatability & an end-user’s perspective?
Where do we measure?

• But measurement at a single or few locations are hard to generalize from …
• Measurements across the wide-area
  – Vantage points in the same places, but across a wider area
  – Distributed platforms for coordinated measurements
And at what layer?

- Network infrastructure and routing
- Traffic
- Applications
- The user up-the-stack

Higher layers, different concerns
- Censorship
- Ethical considerations
Outline

- Experiments in today’s network
- Strategies and good practices
- Edge network perspective: Network positioning
- Application performance: Public DNS and CDNs
- Moving up the stack: Broadband reliability
On sound measurements

Do the results derived from our measurement support the claims made?

- Key question for validation of measurement-based research, but no standards
A Socratic approach*

- **Q1**: Are the measurements being used of good enough quality for the purpose of the study? Need metadata!

- **Q2**: Is the level of statistical rigor used in the analysis commensurate with the quality of the measurements?

- **Q3**: Have alternative models been considered and what criteria have been used to rule them out?

- **Q4**: Does model validation reduce to showing that the proposed model can reproduce certain statistics of the data?

*B. Krishnamurthy, W. Willinger*
Topography as an example

- Internet topology – *Why do we care?*
  - Performance of networks critically dependent on topology
  - Modeling of topology needed to generate test topologies
  - ...

- Internet topology at different levels
  - Router-level reflect physical connectivity
    - Nodes = routers
    - From tools like traceroute or public measurement projects like CAIDA’s Ark
  - AS-level reflects relationships between service providers
    - Nodes = AS
    - From inter-domain routers that run BGP and public projects like Oregon Route Views
Trends in topology modeling

(Observation ➔ modeling approach)

- Long-range links are expensive
  - Random graph (Waxman ’88)

- Real nets are not random, but have obvious hierarchies
  - Structural models (GT-ITM, Zegura et al. ’96)

- Internet topologies exhibit power law degree distributions (Faloutsos et al., ’99)
  - Degree-based models replicate power-law degree sequences

- Physical networks have hard technological (and economic) constraints
  - Optimization-driven models topologies consistent with design tradeoffs of network engineers
Power laws and Internet topology

• “On power-law relationships of the Internet topology,” Faloutsos et al. (SIGCOMM ’99)

From Faloutsos et al. ‘99
Degree-based models and the Internet

• “Error and attack tolerance of complex networks”, R. Albert et al. (Nature 2000)
  – Degree sequence follows a power law (by construction)
  – High-degree nodes correspond to highly connected central “hubs”, crucial to the system
  – Achilles’ heel: robust to random failure, fragile to specific attack

• Does the Internet have these features?
  – No … emphasis on degree distribution, ignoring structure
  – Real Internet very structured
  – Evolution of graph is highly constrained

Preferential Attachment
• (Q1) Are the measurements good enough …. 
  – Router data – original goal to “collect some experimental data on the shape of multicast trees”
    • Collected with traceroute …
  – Inter-domain connectivity data – BGP is about routing …

• (Q2) Given the answer to Q1, fitting a particular parameterized distribution is overkill
Life persistent questions ...

- ...

- (Q3) There are other models, consistent with the data, with different features
  - Seek a theory for Internet topology that is explanatory and not merely descriptive

- (Q4) Yes – model validation reduced to showing that the proposed model can reproduce certain statistics of the available data
Outline

- Experiments in today’s network
- Strategies and good practices
- Edge network perspective: Network positioning
- Application performance: Public DNS and CDNs
- Moving up the stack: Broadband reliability
Network positioning – what for?

- How to pick among alternative hosts?
  - To locate closest game server
  - To pick a content replica
  - To select a nearby peer in BitTorrent
  - ...

- Determine relative location of hosts
  - Landmark-based network coordinates (e.g. GNP)
  - Landmark-free network coordinates (e.g. Vivaldi)
  - Direct measurement (e.g. Meridian)
  - Measurement reuse (CRP)
GNP and NPS implementation*

- Model the Internet as a geometric space, a host position = a point in this space
- Network distance between nodes can be predicted by the modeled geometric distance
- For scalable computation of coordinates – landmarks

*T.S. Eugene et al., A Network Positioning System for the Internet, USENIX ATC 2004
GNP and NPS implementation*

- How do you test this? Simulation
  - Controlled experiments in a simulator using a topology generator based on Faloutsos et al. ’99

- On a global testbed - PlanetLab
  - Large set of vantage points …
  - Programmable
  - Testbeds provide wide-area network paths
PlanetLab

- A global research network to support the development of new network services
  - Distributed storage, network mapping, P2P, DHT, ...
- Each research project has a "slice", or virtual machine access to a subset of the nodes

Currently 1353 nodes at 717 sites
NPS Evaluation

- Operational on PL – use a 20hr operation period
- Using 127 nodes, 100 RTT samples per path, all-to-all
  - Select 15 distributed noes as landmarks, others as regular nodes

From T.S. Eugene et al., …
... adding the last mile via P2P clients ...

- Between PL and Azureus nodes (PL-to-P2P)
  - Ledlie et al, NSDI’07
- Between BitTorrent nodes (P2P) –
  - Choffnes et al, INFOCOM’10 (median latency 2x Ledlie’s)
Cost of error to applications

- RALP, latency penalty for an app from using network positioning, compared to optimal selection
  - Compare top 10 selected nodes ordered by estimated distance

\[
\frac{(\text{selected} - \text{optimal})}{\text{optimal}}
\]

27 times worse than optimal!
Access networks – missing piece

- Access networks not capture by existing testbeds
- Ignoring …
  - High latency variance, last-mile issues, TIV
  - Internet bottlenecks (most in access networks)
  - High heterogeneity (LTE, 802.11, satellite, Cable, Fiber …)

*Dischinger et al, SIGCOMM’08*
Growing current testbeds is not enough

- More academic network nodes doesn’t help
- Need to capture the larger Internet

![Graph showing number of unique inter-AS links](image)

- 280 PlanetLab nodes in U.S. and Europe
- 27 end nodes in U.S. and Europe

*Dischinger et al, SIGCOMM’08*
SatelliteLab – challenge

- Add nodes at the edge while preserving the benefits of existing testbeds
  - Stable software environment
  - Complete management of private virtual slices
  - Extensive API for distributed services to be built upon

- Problem with edge nodes
  - Not dedicated testbed nodes
  - Limited storage and processing resources
  - Often located behind middle boxes
SatelliteLab – key ideas

- Delegate code execution to the planets
- Send traffic through satellites to capture access link
- Detour traffic through planets to avoid complaints and work around NATs or firewalls
Outline

- Experiments in today’s network
- Strategies and good practices
- Edge network perspective: Network positioning
- Application performance: Public DNS and CDNs
- Moving up the stack: Broadband reliability
Internet experimentation by example

- 34 DNS lookups
- 204 HTTP requests
- 520 KB of data downloaded
Ubiquity of Content Delivery Networks

And it’s not just CNN

- 90% of top 50 Alexa’s sites
- 74% of top 1000 Alexa’s sites

56% of domains resolve to a CDN
Earlier this month, Google announced that it had become the largest public DNS service in the world, handling an average of more than 70 billion requests a day. From Google’s point of view, this was great. As it pointed out in its official blog, a good DNS service helps make the Web faster and more secure. That’s true. But is a giant DNS in Google’s hands really good for the world?
Industry proposed solution – Extend DNS

- To avoid impact on Web performance, add client information to DNS requests
  - A EDNS0 extension “edns-client-subnet”
  - Resolver adds client’s location (IP prefix) to request
  - *Needs CDN and public DNS to comply*
The value of experimentation

- *What is the impact of DNS server location on Web performance?*
  - No straight answer

- A complex system requires observation and experimentation to be studied and understood
  - Where is the content hosted?
  - Where are the DNS server?
  - Where is the user?
  - What is the impact of the user’s last-mile?
  - …
An experimentalist’s questions

- *Does it matter?* Do you experience a slower Web with public DNS?
  - Maybe not if public DNS servers are everywhere
  - Or if content is hosted in very few locations
An experimentalist’s questions

- If it does matter, *does the EDNS ECS extension solve it*?
- If it solves it, *is it being adopted by services*?
- If it is not being adopted, *can an end-host solution address it*?
- *How would such a solution compare*?
- *

- *What would you need to explore this*?
  - An experimentation platform at the Internet’s edge
The value of experimental platforms

• An experimental platform at the network’s edge
  – Large set of vantage points …
  – In access networks worldwide
  – Programmable
  – Can’t you not use SatelliteLab?

• Today’s platforms
  – Lack the diversity of the larger Internet
  – Assume experimenters == people hosting the platform
  – Or rely on the “common good” argument
    • DIMES, since 2004 – 453 active users
    • Even SETI@Home– 152k active users, since 1999
Experiments at the edge – goals/challenges

- Host by end users and grow organically
  - How to reach the Internet’s edge?

- Efficient use of resources, but not intrusive
  - As many experiments as possible, but not at arbitrary times or from any location

- Easy to use and easy to manage
  - How to program for thousands of nodes?

- Safe for experimenters and users
  - Extensible and safe? We can’t run arbitrary experiments
DASU pushing experiments to the edge

- Aligned end-users’ & experimenters’ objectives
  - Dasu: broadband characterization as incentive
    - Are you getting the service you are paying for?
- Software-based and hardware-informed
  - As a BitTorrent extension and a standalone client, with the router’s help
- Easy to use by experimenters
  - A rule-based model with powerful, extensible primitives
- Secure for end-users and networks
  - Controlling experiments’ run and their impact
Dasu – Getting to the edge

- Aligned the goals of experimenters and those hosting the platform
  - Characterize users’ broadband services
    Are you getting what you are paying for?
  - Support experimentation from the edge

<table>
<thead>
<tr>
<th></th>
<th>End-user</th>
<th>Experimenter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coverage</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>Availability</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>At the edge</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>Extensibility</td>
<td>✔️</td>
<td>✔️</td>
</tr>
</tbody>
</table>
Dasu in the world

- 100,118 users
- 166 countries
- 2,431 networks
Dasu – Easy to use for experimenters

- Declarative language for experiments
  - Clear, concise experiments
  - Easy to check
  - Easy to extend

```java
rule "(2) Handle DNS lookup result"
when $dnsResult:
    FactDnsResult(toLookup=="eg.com")
then
    String ip = $dnsResult.getSimpleResponse();
    addProbeTask(ProbeType.PING, ip);
end
```
Design – System components

- Experiment Admin Service
- Configuration Service
- Registration
- Configuration
- Experiment Task
- Measurement Activity
- Experiment Lease
- Coordination Service
- Data Service
- Experiment Report
Dasu – Running from the edge

• Secure the platform
  – Sandboxed experiments
  – Resource profiling
  – Secure communication

• Large-scale platform ➔ large-scale impact
  – Controlled aggregated impact of experiments with leases and elastic budgets
  – …
Dasu – Running from the edge

- Minimal impact on user’s performance
  - Limit probes to low-utilization periods
  - Pre-defined probe rates
  - Restricted aggregate bandwidth consumption

- Facing the complexity of home networks
  - Increasingly complex home networks
  - No dedicated (cross-traffic)
Complexity in number of devices

Number of networked devices found

4.6k home networks

65% of homes have at least one device

16% of homes have 3 or more
But not all devices play the same role

- Gateways
- External-facing: talks to the outside world
- Internal-facing: talks within the home network

Internal-facing (58%)
External-facing (5%)
With complexity, externally-facing devices...
The good news ...

- Complexity drives UPnP adoption to simplify home-network management

- UPnP-enabled gateway to infer cross-traffic
  - For network experimentation and broadband characterization from home
  - (the “hardware-assisted” part)
With more devices, UPnP-enabled gateways

As # of devices increases so does the likelihood home gateway supports UPnP
Many opportunities for experimentations

“who else is out there”

For 85% locations device is alone 10% of time

For 20% of samples the host is alone

For 50% of samples no other external device is present!
Usage rather than presence (microdynamics)

- For broadband characterization
  - No cross-traffic
  - Local cross-traffic from other applications in the host
  - Cross-traffic from other devices

- UPnP-enabled gateways help identify different network usage scenarios inside the home
Usage rather than presence (microdynamics)

Local cross-traffic from other applications in the host
Cross-traffic from other devices

No cross-traffic
Not alone, but you can tell

- Cross-traffic from other devices

![Graph showing downloaded data over time with different traffic types: BitTorrent, Netstat, and UPnP.](image-url)
Many opportunities to measure

- Access link shared with other devices in the network

For 83% users fraction of time access-link shared is less than 1/2 network
Dasu – Load-control and experiments

For 85% of peers, scheduled probes can be launched immediately.

- 80% download utilization
- 0% upload utilization
Back to our motivating example

- Different DNS $\Rightarrow$ different performance
  - How different (worst)?

In median case, 65% penalty

Data from >10,000 hosts in 99 countries and 752 ASes
The potential of the EDNS approach

- Where public DNS impacts performance ...

- 45% performance improvement
- But very limited adoption*
  - 3% of top 1-million Alexa’s sites
  - +10% enabled but not in use

*Streibelt et al., Exploring EDNS-Client-Subnet Adopters in your Free Time, IMC13
An alternative end-host solution

- No need to wait for CDN/DNS support
- Don’t reveal user’s location, just “move” DNS resolver close to the user
  - Run a DNS proxy on the user’s machine
  - Use Direct Resolution to improve redirection
    - Recursive DNS to get CDN authoritative server
    - End host directly queries for CDN redirection

http://www.aqualab.cs.northwestern.edu/projects/namehelp
Available now – works with all CDNs and DNS services

Within 16% of potential

Improves performance in 76% of locations

Today, ~145,000 in 168 countries
Outline

- Experiments in today’s network
- Strategies and good practices
- Edge network perspective: Network positioning
- Application performance: Public DNS and CDNs
- Moving up the stack: Broadband reliability
Broadband and its rapid growth

- Instrumental for social & economic development
Broadband and its rapid growth

- Instrumental for social & economic development
- 70+ countries with majority of population online
- 30% higher connection speeds per year, globally

*Akamai’s State of Internet Report, Q1 2015
The importance of being always on

- With higher capacities, a migration to “over-the-top” home services

- And higher expectations of service reliability
  - Main complain, from a UK Ofcom survey (71%)*

*Ofcom, UK broadband speed, 2014
Broadband reliability challenges

- What does “failure” mean in best-effort networks? What metrics for reliability should we use? What datasets?
- What determines your reliability? ISPs, services within it, technologies, geography, …?
- What can we do now to improve reliability?
- But, first, do users care? Does it impact their quality of experience?
Importance of reliability

- *How do we measure reliability impact on users’ experience? At scale?*
- Ideally – a classical controlled experiments
  - Control and treatment groups, randomly selected
  - Some treated with lower/higher reliability
  - Difference in outcome likely due to treatment
Importance of reliability

• But …
  – Heisenberg effect – change in user behavior
  – Practical issues – control over people’s networks
  – Degrading connections in home routers, would require consensus (and deter participants); doing it without consent will be unethical
Natural rather than control experiments

• Natural experiments and related study designs
  – Common in epidemiology and economics
    • E.g., Snow, pump location and the 1854 cholera epidemic in London
  – Participants assignments to treatment is *as-if random*

• Network demand as a measurable metric likely correlated with user experience
  – Change on network usage ≈ change on user behavior

• Look for network conditions that occur spontaneously, control for confounding factors
A brief note on our datasets

- Broadband performance and usage
  - From FCC/SamKnows *Measuring Broadband America*
    - Collected from home routers, including capacity, loss, latency, network usage
    - ~8k gateways in the US

- To identify source of issues
  - AquaLab’s Namehelp
    - Collected from end devices, including traceroutes
    - A subset of 6k end-hosts from 75 countries
Impact of lossy links

• Hypothesis – *Higher packet loss rates result in lower network demand*

• Experiment
  – Split users based on overall packet loss rate
    • Control group loss rate < 0.06%
  – Select users from *control* and *treatment* groups with similar regions and services (download/upload rate)
    • If usage and reliability are not related, *H* should hold ~50%

<table>
<thead>
<tr>
<th>Treatment group</th>
<th>% <em>H</em> holds</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0.5%, 1%)</td>
<td>48.1</td>
<td>0.792</td>
</tr>
<tr>
<td>(1%, 2%)</td>
<td>57.7</td>
<td>0.0356</td>
</tr>
<tr>
<td>&gt;2%</td>
<td>60.4</td>
<td>0.00862</td>
</tr>
</tbody>
</table>
Impact of frequent periods of high loss

- Hypothesis – *High frequency of high packet loss rates (>5%) result in lower network demand*
- Experiment
  - Users grouped by frequency of periods, 0-0.1% of measurements, 0.1-0.5% of measurements …
  - ...

<table>
<thead>
<tr>
<th>Control group</th>
<th>Treatment group</th>
<th>% $H$ holds</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0.5%, 1%)</td>
<td>(1%, 10%)</td>
<td>54.2</td>
<td>0.00143</td>
</tr>
<tr>
<td>(0.1%, 0.5%)</td>
<td>(1%, 10%)</td>
<td>53.2</td>
<td>0.0143</td>
</tr>
<tr>
<td>(0%, 0.1%)</td>
<td>(1%, 10%)</td>
<td>54.8</td>
<td>0.000421</td>
</tr>
<tr>
<td>(0.5%, 1%)</td>
<td>&gt;10%</td>
<td>70</td>
<td>6.95x10^-6</td>
</tr>
<tr>
<td>(0.1%, 0.5%)</td>
<td>&gt;10%</td>
<td>70.8</td>
<td>2.87x10^-6</td>
</tr>
<tr>
<td>(0%, 0.1%)</td>
<td>&gt;10%</td>
<td>72.5</td>
<td>4.34x10^-7</td>
</tr>
</tbody>
</table>
Broadband reliability challenges

- **Do users care? Does it impact their quality of experience?**
  - First empirical demonstration of its importance

- **What does “failure” mean in best-effort networks? What metrics for reliability should we use? What datasets?**

- **What determines your reliability? ISPs, services within it, technologies, geography, ...?**
  - An approach for characterizing reliability
Characterizing reliability

- To capture different service providers, service tier, access technology, ...
- An approach that uses datasets from national broadband measurement studies
  - e.g., US, UK, Canada, EU, Singapore ...
  - Some resulting constraints (e.g., number, location of vantage points, measurement granularity)
  - But can be readily applied and may inform future designs
Some classical metrics for now

- Classical reliability metrics: Mean Time Between Failures (MTBF) and Mean Down Time (MDT)

\[
MTBF = \frac{\sum \text{Total}_\text{uptime}}{\text{# of}_\text{Failures}} \quad \text{MDT} = \frac{\sum \text{Total}_\text{downtime}}{\text{# of}_\text{Failures}}
\]

- Availability defined based on MTBF and MDT

\[
A = \frac{MTBF}{MTBF + MDT}
\]

- Key to them, a definition of “failure”
A definition of failure

- What is failure is an open issue
- We use packet loss rate
  - Key to throughput and overall performance
    - VoIP can become unstable at 2% [Xu et al, IMC12]

Different distribution of loss rate, we use 1, 5 and 10% for analysis

Cox ~ Insight 27.5hr MTBF

Cox >> Insight 150/94hr MTBF!
Characterizing reliability

- Apply this approach to US FCC broadband data
  - Different tech: 55% cable, 35% DSL, 7% fiber …
  - Different ISPs, large and small, AT&T, Comcast and ViaSat/Exede
  - Every US state with between 0.2% (North Dakota) and 11.5% of boxes (California)

- **How does reliability varies across ...?**
  - Providers
  - Technologies
  - Tier services
  - Geography
  - What’s the role of DNS?
### Top 4 best/worst providers on availability

<table>
<thead>
<tr>
<th>ISP</th>
<th>Average availability</th>
<th>Average downtime</th>
</tr>
</thead>
<tbody>
<tr>
<td>Verizon (Fiber)</td>
<td>99.18</td>
<td>99.80 72</td>
</tr>
<tr>
<td>Frontier (Fiber)</td>
<td>98.58</td>
<td>99.77 124</td>
</tr>
<tr>
<td>Comcast (Cable)</td>
<td>98.48</td>
<td>99.66 134</td>
</tr>
<tr>
<td>TimeWarner (Cable)</td>
<td>98.47</td>
<td>99.69 134</td>
</tr>
<tr>
<td>Frontier (DSL)</td>
<td>93.69</td>
<td>98.87 553</td>
</tr>
<tr>
<td>Clearwire (Wireless)</td>
<td>88.95</td>
<td>98.13 968</td>
</tr>
<tr>
<td>Hughes (Satellite)</td>
<td>73.16</td>
<td>94.84 2350</td>
</tr>
<tr>
<td>Windblue/Viasat (Satellite)</td>
<td>72.27</td>
<td>96.37 2430</td>
</tr>
</tbody>
</table>

At best, 2 9s
Compare with 5 9s of telephone service

Only 1 9s, even with a 10% loss rate threshold
But not all failures are the same

Avg. number of bytes sent/received per hour

Volume of traffic (MB)

Hour of the day

Saturday
Sunday
Weekday
### Top 4 best/worst … at peak hour

**Peak hour: 7PM – 11PM**

<table>
<thead>
<tr>
<th>ISP</th>
<th>Availability</th>
<th>% change U</th>
<th>Availability</th>
<th>% change U</th>
</tr>
</thead>
<tbody>
<tr>
<td>Verizon (Fiber)</td>
<td>99.11</td>
<td>+8.7</td>
<td>99.83</td>
<td>-14.7</td>
</tr>
<tr>
<td>Frontier (Fiber)</td>
<td>98.56</td>
<td>+8.7</td>
<td>99.78</td>
<td>-4.6</td>
</tr>
<tr>
<td>Comcast (Cable)</td>
<td>98.39</td>
<td>+5.3</td>
<td>99.70</td>
<td>-11.7</td>
</tr>
<tr>
<td>Time Warner (Cable)</td>
<td>98.03</td>
<td>+5.3</td>
<td></td>
<td>+1.3</td>
</tr>
<tr>
<td>Frontier (DSL)</td>
<td>87.98</td>
<td>+90.4</td>
<td>98.42</td>
<td>+39.9</td>
</tr>
<tr>
<td>Clearwire (Wireless)</td>
<td>86.35</td>
<td>+23.6</td>
<td>97.57</td>
<td>+29.9</td>
</tr>
<tr>
<td>Hughes (Satellite)</td>
<td>60.97</td>
<td>+45.4</td>
<td>91.38</td>
<td>+66.9</td>
</tr>
<tr>
<td>Windblue/Viasat (Satellite)</td>
<td>69.44</td>
<td>+10.2</td>
<td>94.14</td>
<td>+61.2</td>
</tr>
</tbody>
</table>

Some improvements for fiber and cable

Worst for the others; scheduled and unscheduled downtime?
For most ISPs, MTBF > 200hr, but for wireless and satellite, Typical MDT < 2hr, but for wireless and satellite.
Impact of access technology

Technology – After ISP, the most informative feature for predicting availability

Access technology is the biggest factor in reliability
Impact of access technology

- To separate the impact of ISP from technology
  - Same providers, different technology
Reliability across service class

Business and residential services offer similar reliability

Service class has little effect on reliability
What about service reliability?

- For users, DNS or net failures are indistinguishable
  - But their reliability are not always correlated

Top 6 ISPs by connection and DNS availability

<table>
<thead>
<tr>
<th>ISP</th>
<th>Availability @ 5%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Verizon Fiber</td>
<td>99.67</td>
</tr>
<tr>
<td>Cablevision</td>
<td>99.53</td>
</tr>
<tr>
<td>Frontier Fiber</td>
<td>99.47</td>
</tr>
<tr>
<td>Comcast</td>
<td>99.45</td>
</tr>
<tr>
<td>Charter</td>
<td>99.29</td>
</tr>
<tr>
<td>Bright House</td>
<td>99.28</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ISP</th>
<th>DNS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Insight</td>
<td>99.97</td>
</tr>
<tr>
<td>Windstream</td>
<td>99.90</td>
</tr>
<tr>
<td>Qwest</td>
<td>99.90</td>
</tr>
<tr>
<td>Hughes</td>
<td>99.90</td>
</tr>
<tr>
<td>Frontier Fiber</td>
<td>99.90</td>
</tr>
<tr>
<td>Cox</td>
<td>99.90</td>
</tr>
</tbody>
</table>

Only one ISP in common
Improving reliability

- Target availability for telephone services
  - Five 9s (99.999%) ~ 5.26 minutes per year

- The best you can get on US broadband
  - Two 9s or ~17 hours per year
  - Setting loss rate threshold at 1%, only one provider

- Clearly we need something … key requirements
  - Easy to deploy
  - Transparent to end users
  - Improving resilience at the network level
Where do reliability issues occur?

- Experiment with 6,000 Namehelp
  - Run pings and DNS query (to Google public DNS) at 30sec intervals, traceroute upon failure

76% of issues are connecting to or going through the provider’s network
Improving reliability

- Two options
  - Improve the technology’s failure rate
  - Add redundancy

- Observation: Most users in urban setting “could” connect to multiple WiFi networks

- An approach: *End-system multihoming*
  - Neighbors lending each others networks as backup
  - Perhaps with limits on time or traffic
Estimating the potential of multihoming

- Using FCC data, group users
  - Per census block, the smallest geographical unit
  - Time online, online during the same period

Multihoming with a different ISP adds two “9”s

Multihoming with the same ISP adds one “9”
How many neighboring networks?

- Namehelp again, one month measurement

90.2% of cases, 1+ additional networks
Connecting to neighboring networks

- Look at signal strength

![Signal strength CDF graph](image)

- 40% or higher for ~83%
A system for multihoming

- *How to fail over to a neighbor’s network without interrupting open connections?*
  - Multipath TCP for reliability
  - Gateway creates a VPN to a MPTCP proxy
  - Proxy in the cloud (or Planetlab)
Multihoming at home

- A simple experiment in two scenarios
  - Client runs `iperf`, a second interruption

In both cases, a fast recovery

![Graphs showing transfer rates over time for University 100Mbps, Comcast 75Mbps, and ATT 3Mbps.]
Some closing thoughts

- **Success of networked systems**
  - An integral part of everyday life, critical for modern society
  - Evidence of the success and broader impact of our field
  - But with clear complications for experimentalists

- **How can we experiment with critical, global scale systems, how can we provide evidence of the effects of interventions?**

- **Internet-scale experimentation is still in its infancy**
  - Need new platforms, methodologies, standards, legal and ethical guidelines, …
  - And we need help, we can’t do it alone
Acknowledgements

• Graduate students
  – David Choffnes (graduated)
  – John Otto
  – Mario Sanchez
  – Zach Bischof
  – John Rula
  – Ted Stein

• Collaborators
  – Bala Krishnamurthy (AT&T)
  – Walter Willinger (AT&T)
  – Nick Feamster (Princeton U.)

• Funding sources
  – National Science Foundation
  – Google
Internet-scale Experimentation

The challenges of large-scale networked system experimentation and measurements