RIPE 57
DUBAI, 26-30 OCTOBER 2008

Google™

Global IPv6 statistics

Measuring the current state of IPv6 for ordinary users

Steinar H. Gunderson
Software Engineer

Motivation

Google

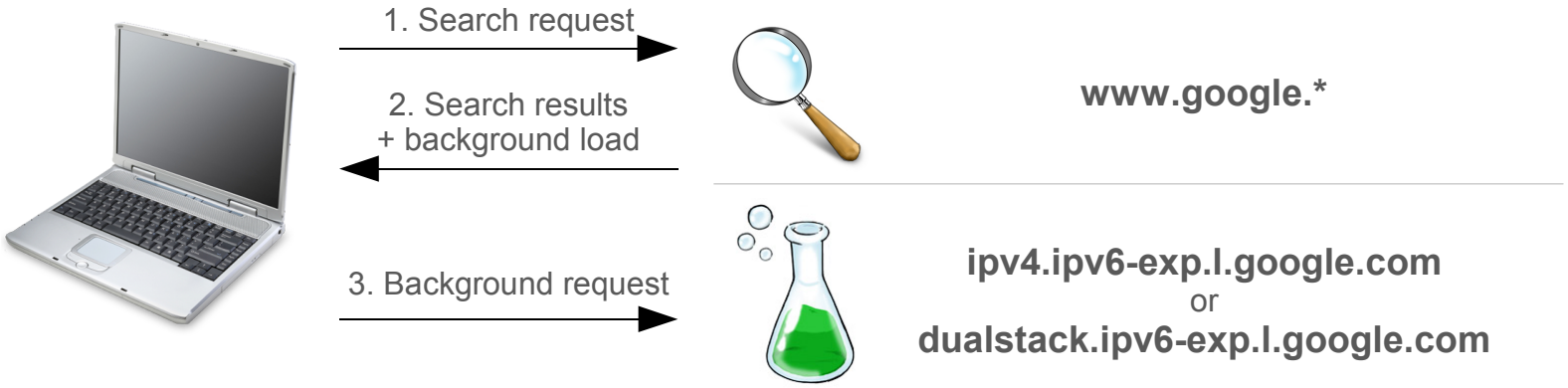
- There is too little data about IPv6 among clients
  - Existing measurements mostly on a small scale and/or only indirectly related to client IPv6 availability (e.g., IPv6 traffic percentage, IPv6-enabled ASNs)
  - Best existing number is probably **0.086%** (Kevin Day, March 2008)

- General worry that turning on IPv6 can cause all sorts of brokenness
  - Tunnels that someone forgot
  - Suboptimal routing
  - Home routers doing evil things to AAAA queries

- We need to figure out **how common** IPv6 is among our users, how prevalent **brokenness** is, and how we can best serve our IPv6 users
  - Our question: What is the impact of adding an AAAA record to a web site?

Methodology

- Enroll a small fraction of ordinary Google users into an “IPv6 experiment”, where their browser is asked to perform a background request
  - Involves users from all datacenters equally, but background request goes to one of two datacenters (one in the US, one in Europe)
  - Cryptographically signed to avoid easy injection of false data



- Recorded information:
  - IPv4 and IPv6 addresses, as applicable
  - Image request latency
  - Browser/OS details (User-Agent string)

Key figures

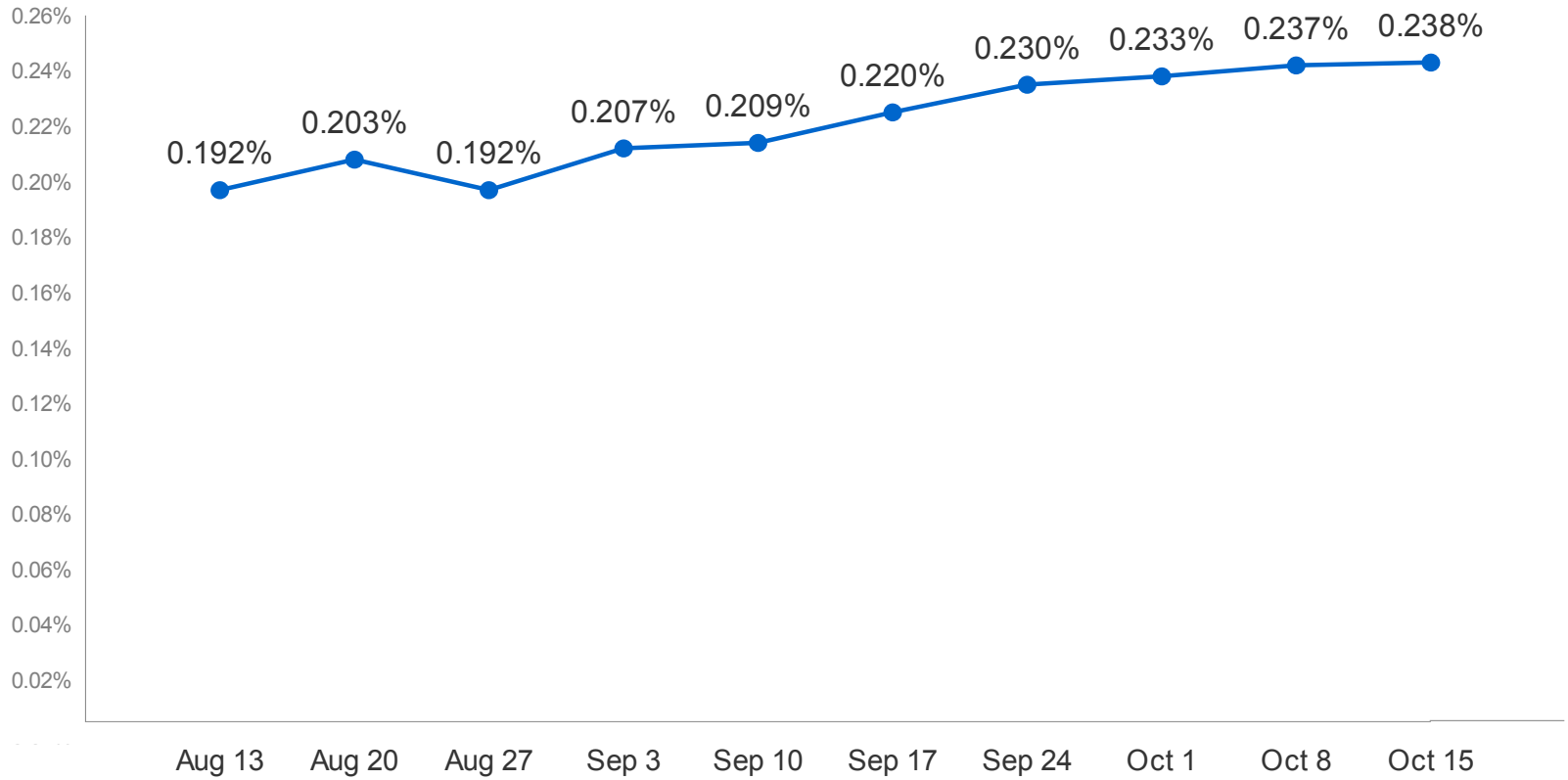
Overview of connectivity and latency data

Connectivity

Google

- **0.238%** of users have useful IPv6 connectivity (and prefer IPv6)
- **0.09%** of users have *broken* IPv6 connectivity
  - That is, adding an AAAA record will make these users unable to view your site
  - Due to statistical issues, this is a much less accurate figure (could easily be 0.06% or 0.12%), so take it with a grain of salt
- Probably at least **a million** distinct IPv6 hosts out there
  - Again, a number with statistical caveats

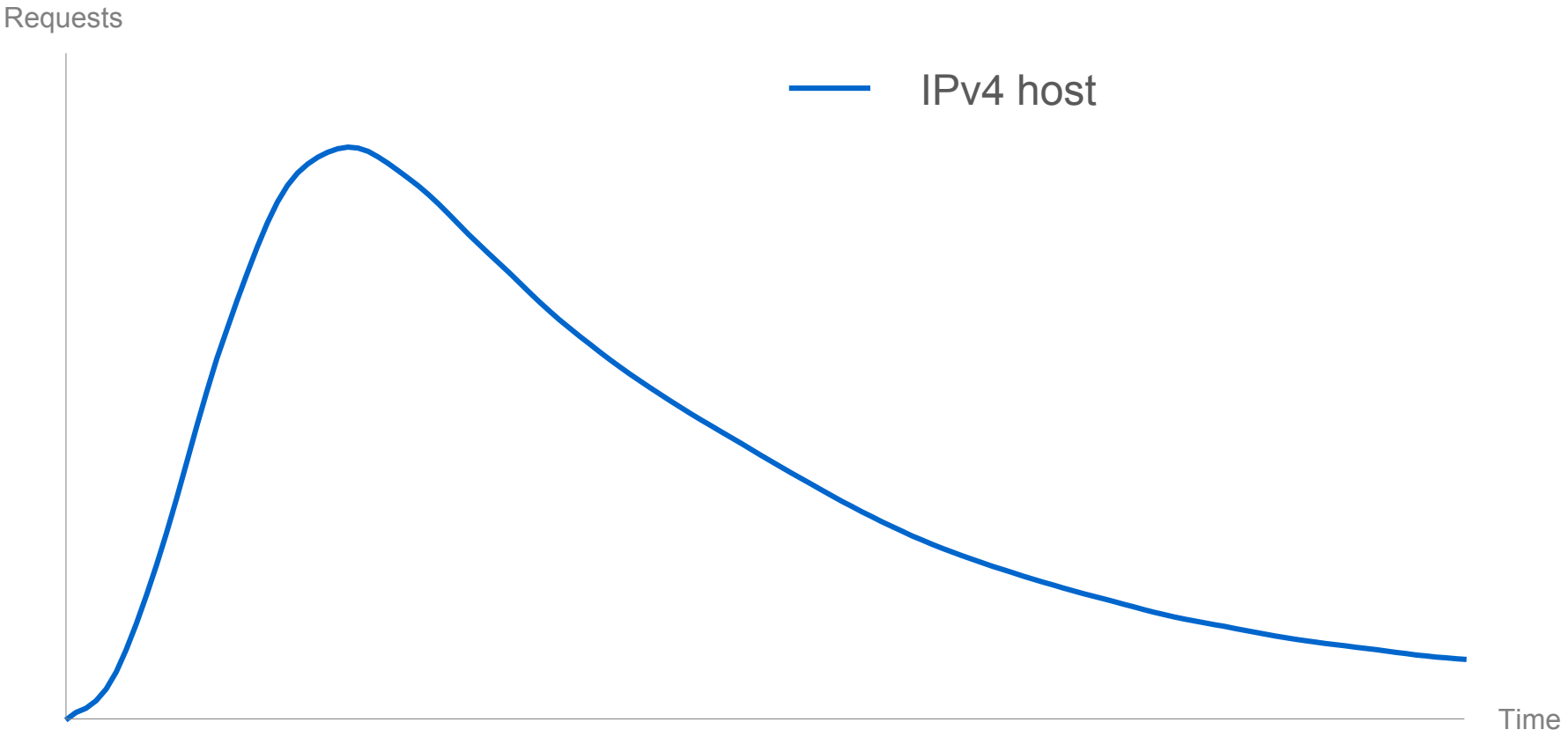
Connectivity development over time



Week averages, 2008

Latency

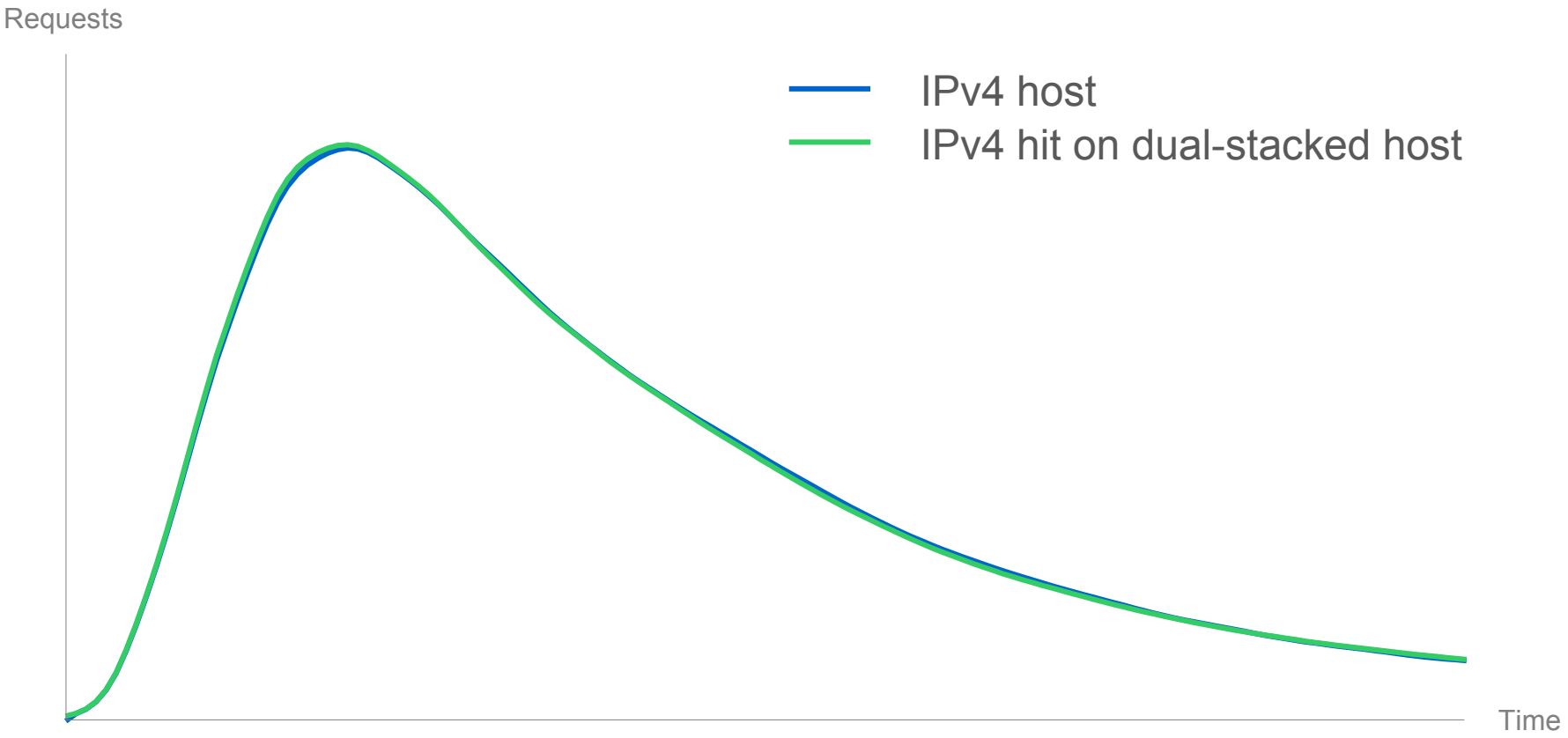
Latency distribution function, clients visiting ipv4.ipv6-exp.l.google.com



Note: This graph is *not* indicative of ordinary Google service latency

Combined data, Aug–Oct 2008

Latency



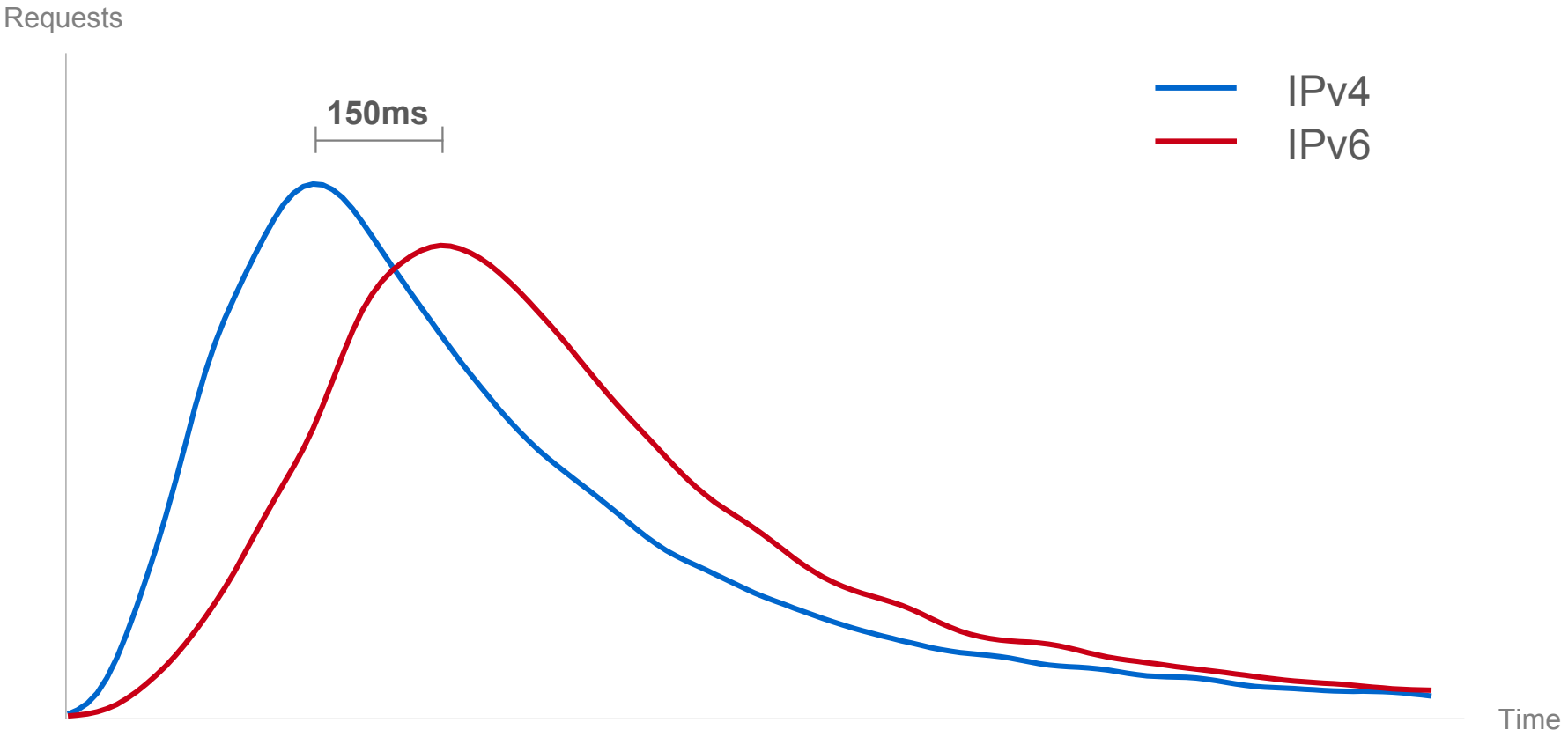
Combined data, Aug–Oct 2008

Latency, continued

Google

- We cannot directly graph IPv4 vs. IPv6 latency
  - IPv6-enabled hosts are likely to have faster network connectivity overall (universities, power users, etc.)
  - Need a way to remove inherent bias
- Solution: Find pairs of hits from the same /24 IPv4 network, discard all other data
  - Gives comparable (paired) data sets
- This means we are measuring relative latency for a *different set* of users, but the data is still indicative of what you can expect today

Relative IPv4/IPv6 latency (paired data)

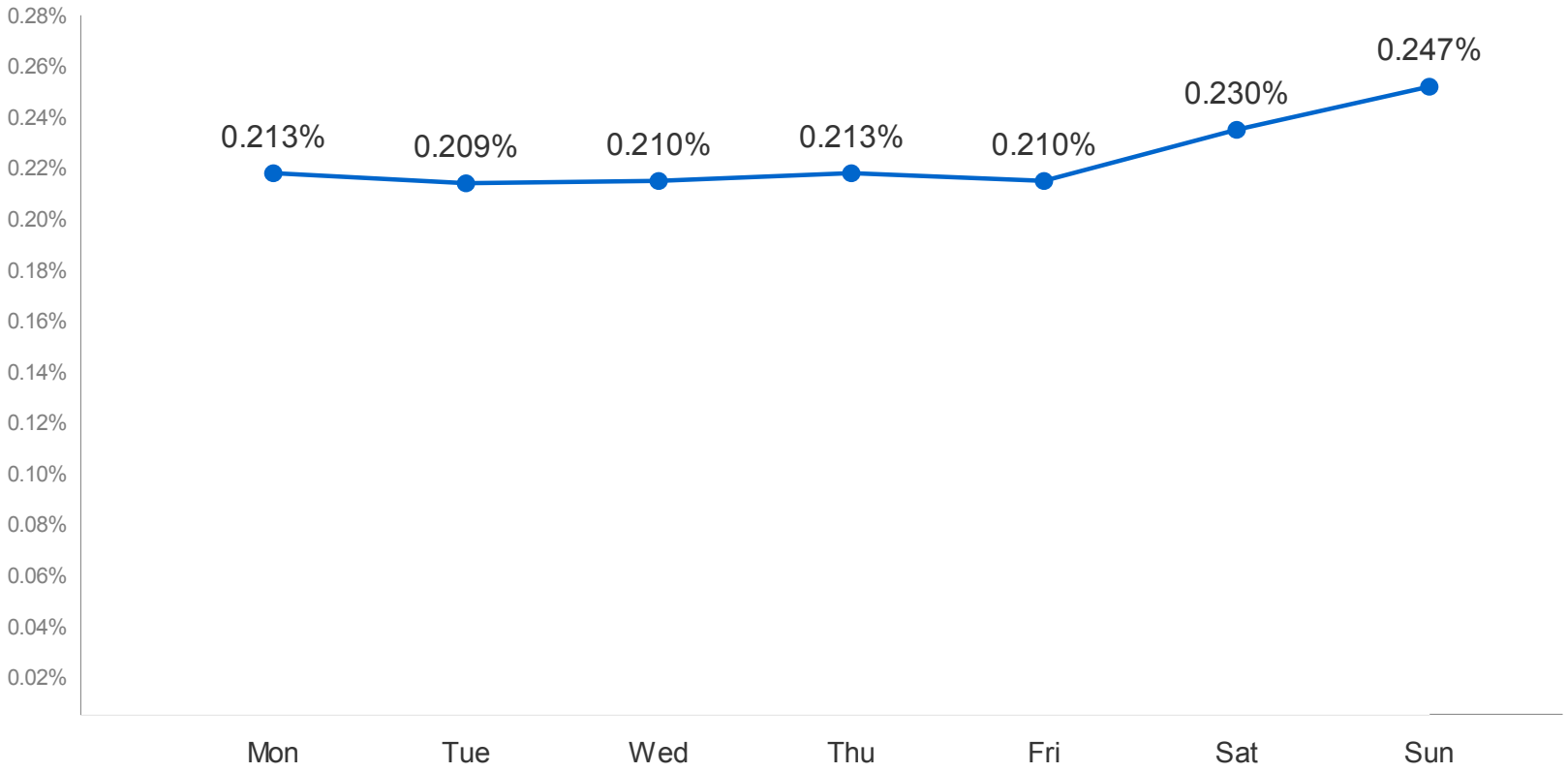


Combined data, Aug–Oct 2008

Data breakdowns

Drilling in to get a more detailed look

Connectivity by weekday (UTC)



Combined data, Aug–Oct 2008

Connectivity by country

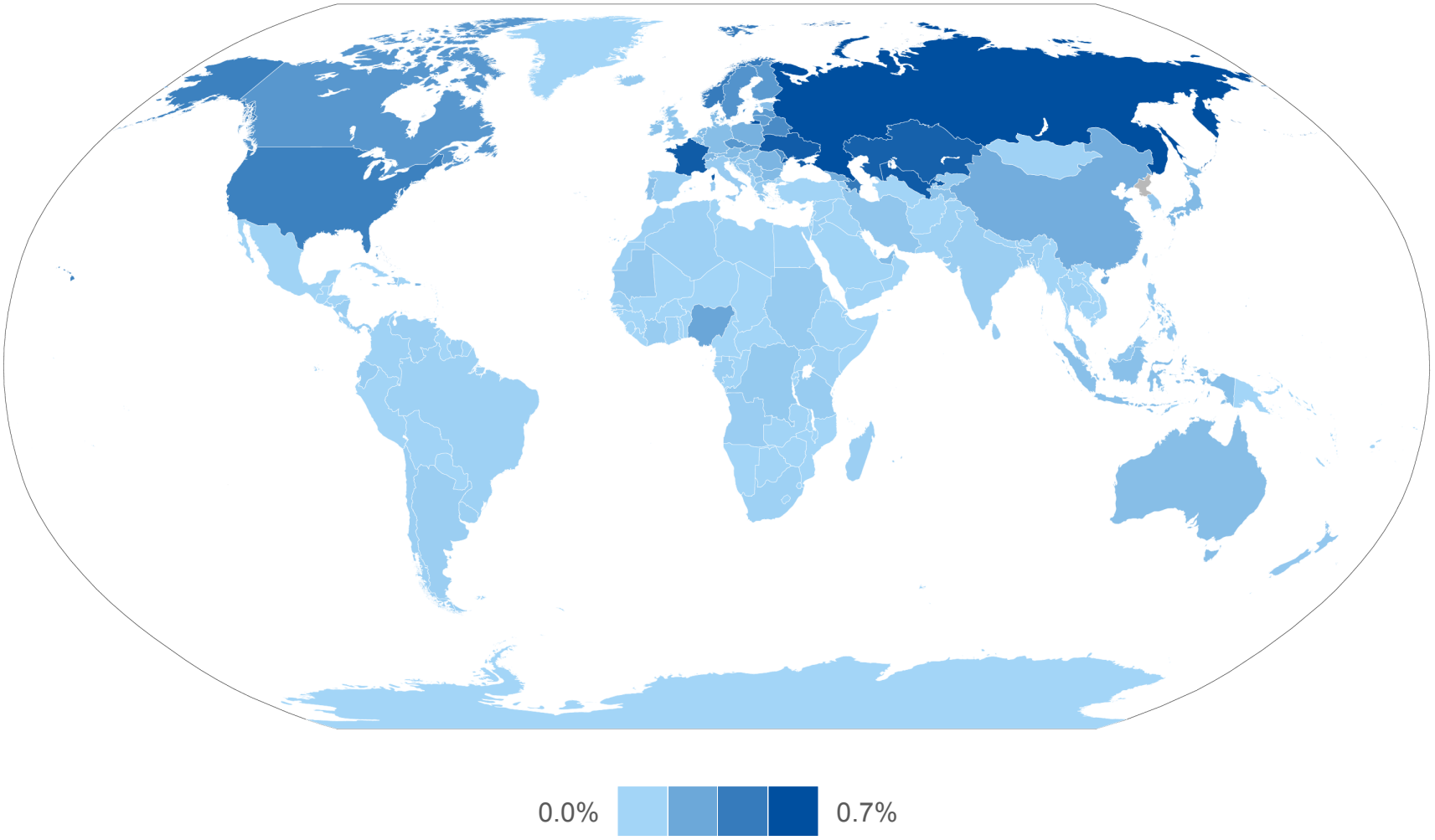
Google

- Based on the IPv4 address, geolocate the user, then group by country
  - Some countries with relatively little Internet traffic removed

Country	IPv6 penetration
Russia	**0.76%**
France	0.65%
Ukraine	0.64%
Norway	0.49%
United States	0.45%
…	
China	0.24%
Japan	0.15%

Combined data, Aug–Oct 2008

Connectivity by country



Combined data, Aug–Oct 2008, lower bound of 68% confidence interval

Method of IPv6 connectivity

Google

- Based on the IPv6 address, we can infer how the user gets IPv6 access
  - Unfortunately, no good way of distinguishing native from tunnels based on the address alone
  - Vista with Teredo prefers IPv4 by default, so probably undercounted

Method	Global usage
6to4	67.9%
Native/other	29.1%
ISATAP	1.6%
Teredo	1.4%

- Some countries stand out
  - United States, Canada: 95% 6to4
  - **France: 95% native** (almost all free.fr)
  - China: 71% native, 25% ISATAP

Combined data, Aug–Oct 2008

Breakdowns by OS

Google

IPv6 penetration and connectivity type by operating system

Ranked by overall IPv6 penetration

Operating system	IPv6 penetration	Native/other proportion	6to4 proportion	Teredo/ISATAP proportion
Mac OS	**2.44%**	**9%**	**91%**	**0%**
Linux	0.93%	86%	13%	1%
Windows Vista	0.32%	55%	43%	2%
Windows Server 2003	0.07%	–	–	–
Windows XP	0.03%	50%	30%	20%
Windows 2000	<0.01%	–	–	–

52% of all IPv6 hits are from Macs with 6to4

97% of all Teredo users are on Windows (even undercounting Vista)

Combined data, Aug–Oct 2008

Summary

Brief analysis and conclusions

Overall trends

Google

- IPv6 prevalence is still low, but growing by the week
  - Large (and sometimes surprising) variations among individual countries
  - Still heavily influenced by single deployments (e.g., free.fr)
- It's not that broken
  - ~0.09% clients lost, ~150ms extra latency – don't believe the FUD
- The default policy matters – a lot
  - Vista: 10x IPv6 prevalence over XP (OS defaults to enabling IPv6)
  - Mac OS: 8x IPv6 prevalence over Vista
    (Airport Extreme with 6to4 as default)
- 6to4 is by far the most common transition mechanism
  (at least when you don't count Vista's not-preferred-by-default Teredo)
  - Probably in part due to the AirPort Extreme
  - Consider running your own 6to4 relay for return packets

Future work

Google

- Keep it running
  - Gather more data as time goes by
- Figure out *why* we lose users on the way
  - So we can fix it
- Run different experiments to get more accurate loss numbers
  - Paired data (i.e., two separate background requests) has been done before and is a possibility, but does not solve all problems
  - More client-side logic would help

Questions?