

LinkedIn Case Study: IPv6 at a Social Media Company



Stephanie Schuller

Implementation Manager
Infrastructure Architecture & Strategy



IPv6 at LinkedIn

Making the business case

3 Phases

Phase 1: Serving IPv6 externally

- What does “supports IPv6” mean?
- Application readiness checklist
- Challenges: data type & IP-based geo-location
- Launch sequence

Phase 2: Dual-stack datacenter

- Why build dual-stack?
- Challenges
 - IPv6 at scale
 - Operational tooling - What does “supports IPv6” mean?

Phase 3: IPv6-only datacenter . . . Our next goal!

Measurements: Traffic and speed

How do you make the business case when people don't even know what IPv6 is?



IPv6... Don't they have a vaccine for that?

The IPv6 Test

MAKING THE BUSINESS CASE

1. Turn off wifi
2. Go to **ipv6-test.com** on your browser

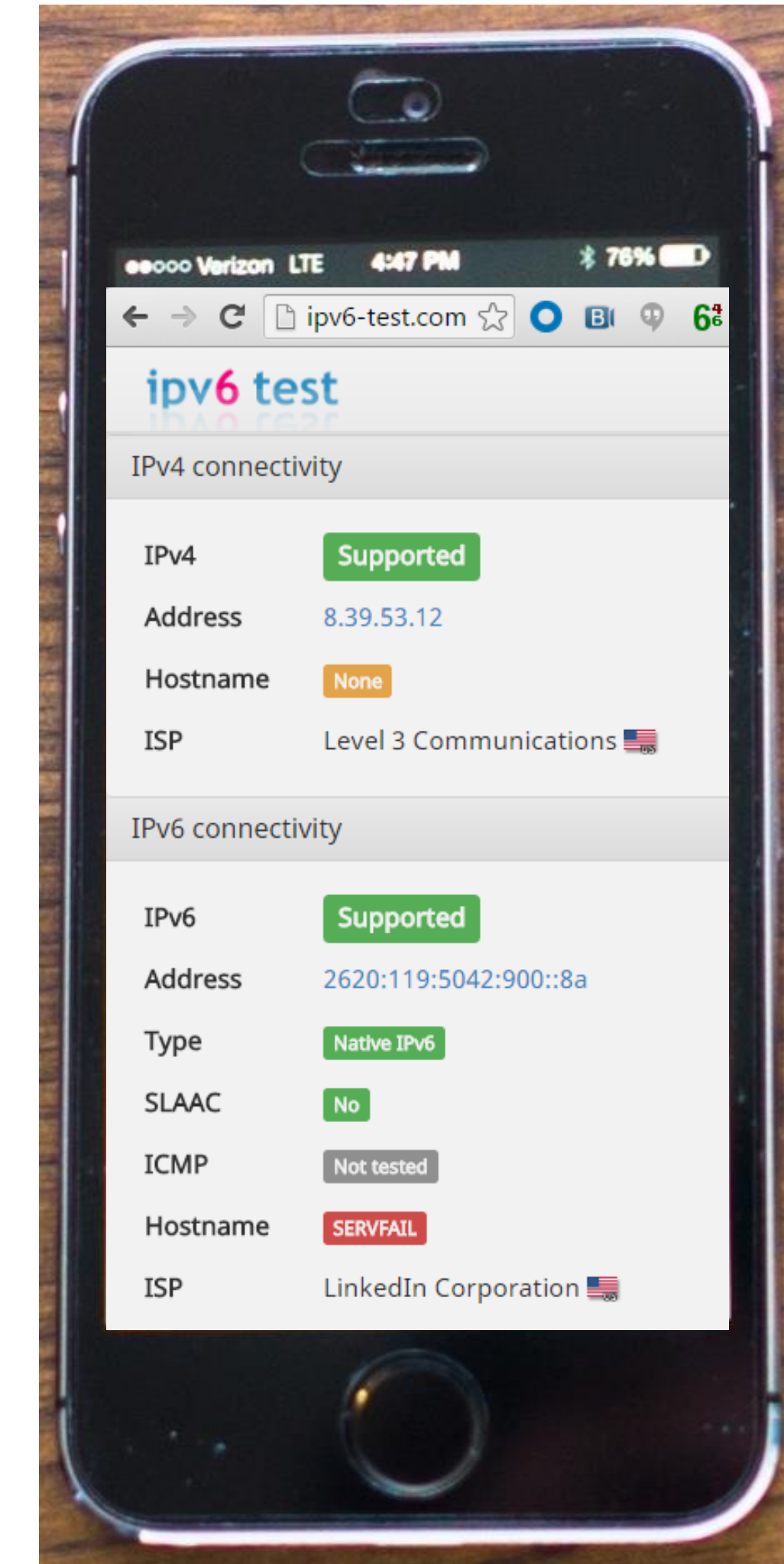
Your IPv4 address →

Old system designed in the 1970's
Too short – ran out of addresses

Your IPv6 address →

ISPs started moving to IPv6 years ago

**Our customers are already using IPv6.
We need to move to the modern internet.**



Two separate IPv6 adoption decisions

External, customer-facing	Easier to justify
Internal, employee-facing	Harder to justify

11 reasons to adopt IPv6

For enterprise businesses

1. Our entire business model relies on the internet. If we stay on outdated internet technology and let ourselves fall behind, what kind of business strategy is that?
2. Our customers are already using IPv6. ISPs started moving customers to IPv6 years ago. *(Demo IPv6-test.com on smartphone)*
3. Our industry peers are already serving IPv6.
4. IPv4 has a security hole. IPv4 uses CG-NAT, which voids key defenses.
(More: The world ran out of IPv4 addresses, so many people have to share one address. This lets malicious people hide.)
5. IPv6 is faster than IPv4 on mobile.

11 reasons to adopt IPv6

MAKING THE BUSINESS CASE

For enterprise businesses

6. Moving to IPv6 is inevitable. The only choice is whether to keep delaying.
7. We'll have to retrofit later if we keep building with legacy technology. Wastes people and money.
8. IPv4 was designed back in the 70's, when no one imagined the online world of today. We're spending so much effort trying to get that old technology to do what we need. That's wasted time and money.
9. Our industry changes fast. Keep up or be left behind.
10. (For acquisitions) Supporting IPv6 isn't optional – it's part of the transition to our infrastructure.
11. Our employees are already using IPv6. Devices have IPv6 turned on by default. We just aren't *managing* IPv6.

Reaching different audiences

Audience	Reasons for IPv6 that particularly resonated
Top executives	Catching industry trends, business strategy, keeping up with industry peers
Product managers	Catching industry trends, keeping up with industry peers, speed on mobile
Internal IT	Avoiding retrofits (wasted work), career development
Applications teams	Speed on mobile, keeping up with industry peers, career development
Production operations	Keeping up with technology, avoiding retrofits, speed on mobile, career development
Acquisitions	IPv6 support is not optional – part of transition to our infrastructure
Everyone	Security holes, embarrassment about using old technology

Evangelization methods & messages

MAKING THE BUSINESS CASE

1. Put IPvFoo on office Chrome browsers, IPvFox on Firefox

Message: Visibility

2. Highlight career development opportunity

Message: You'll want to work on this new technology project because it helps advance your career and skills

3. Frequent quick mention of IPv6 by upper management

Message: IPv6 is mandatory and strategically important

4. Teach people how to demo the IPv6 test (if their local provider offers IPv6)

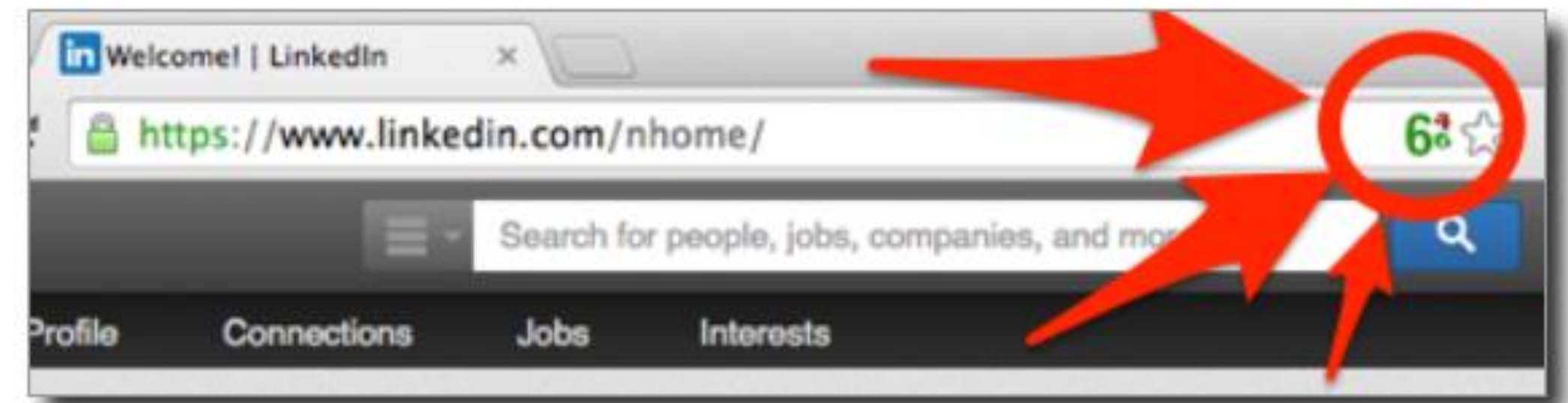
Message: Our customers are already using IPv6. IPv6 is already here.

5. Short videos – *nontechnical!* (Can show in gathering time before large meetings)

Message: Pride in the company's modern technology, strategic importance of v6 trend

6. (For internal IT) Build everything new with IPv6, of course. Put IPv6 requirement in your normal n-year refresh cycle for equipment and systems.

Message: IPv6 fits into your normal budget planning



PHASE 1: SERVING IPv6 EXTERNALLY

PHASE 2: DUAL-STACK DATACENTER

PHASE 3: IPv6-ONLY DATACENTER

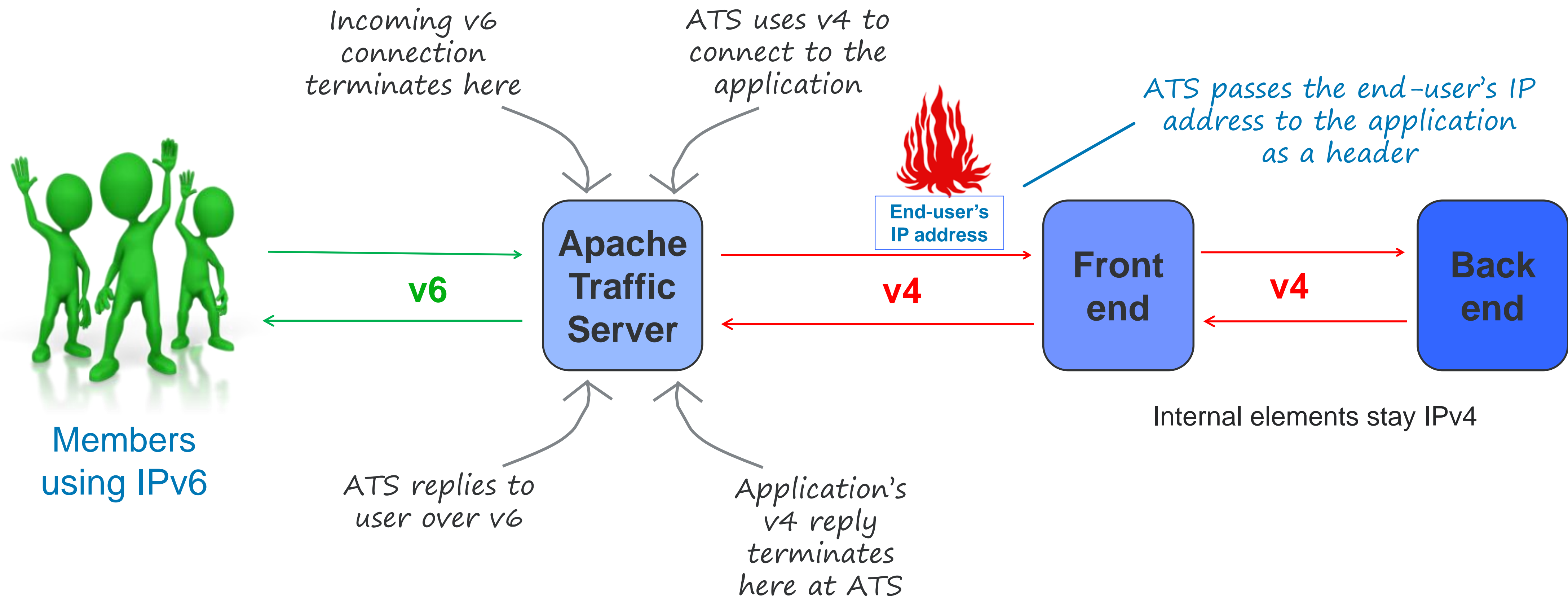
10
months

- Infrastructure upgrades
- LinkedIn Email feature
- LinkedIn applications can handle v6 addresses as data
- QA environments
- Offices
- Data systems & analysis
- Kafka – data pipeline
- IP-based geo-location
- Security
- Monitoring
- **Launch external Production: website, mobile, and API**

- Operational tooling
- QA & staging environments
- LinkedIn applications can run in dual-stack environment
- Internal applications can run in dual-stack environment
- LinkedIn applications and internal applications move to native v6 as ready
- Connect offices over v6
- **Launch dual-stack: Our new Oregon datacenter Production**

- LinkedIn applications run natively on v6
- Internal applications run natively on v6
- Operational tooling & monitoring run natively on v6
- **Launch IPv6-only: Our next new datacenter:**

PHASE 1: SERVING IPv6 EXTERNALLY



PHASE 1: SERVING IPv6 EXTERNALLY

What does “supports IPv6” mean?

This is not obvious to people!

- ✓ 1. Handles IPv6 address as a piece of data
- ✓ 2. Functionality that uses IP address works with IPv6 addresses
- 3. Communicates over IPv6

Providing IPv6 externally requires only #1 and #2

Application Readiness Checklist

PHASE 1: SERVING IPv6 EXTERNALLY

✓	Use case
	I store IP addresses in a database
	I store IP addresses in log files
	I do analysis that includes IP addresses
	I use IP-based geo-location
	My Production monitoring, trouble-shooting, or operational tools use IP address
	I use a third-party service for monitoring, performance testing, etc.
	My application uses a CDN (content distribution network used for caching and other functions)
	I identify a user by his IP address OR I differentiate users by their IP addresses
	I block/blacklist an abuser's IP address

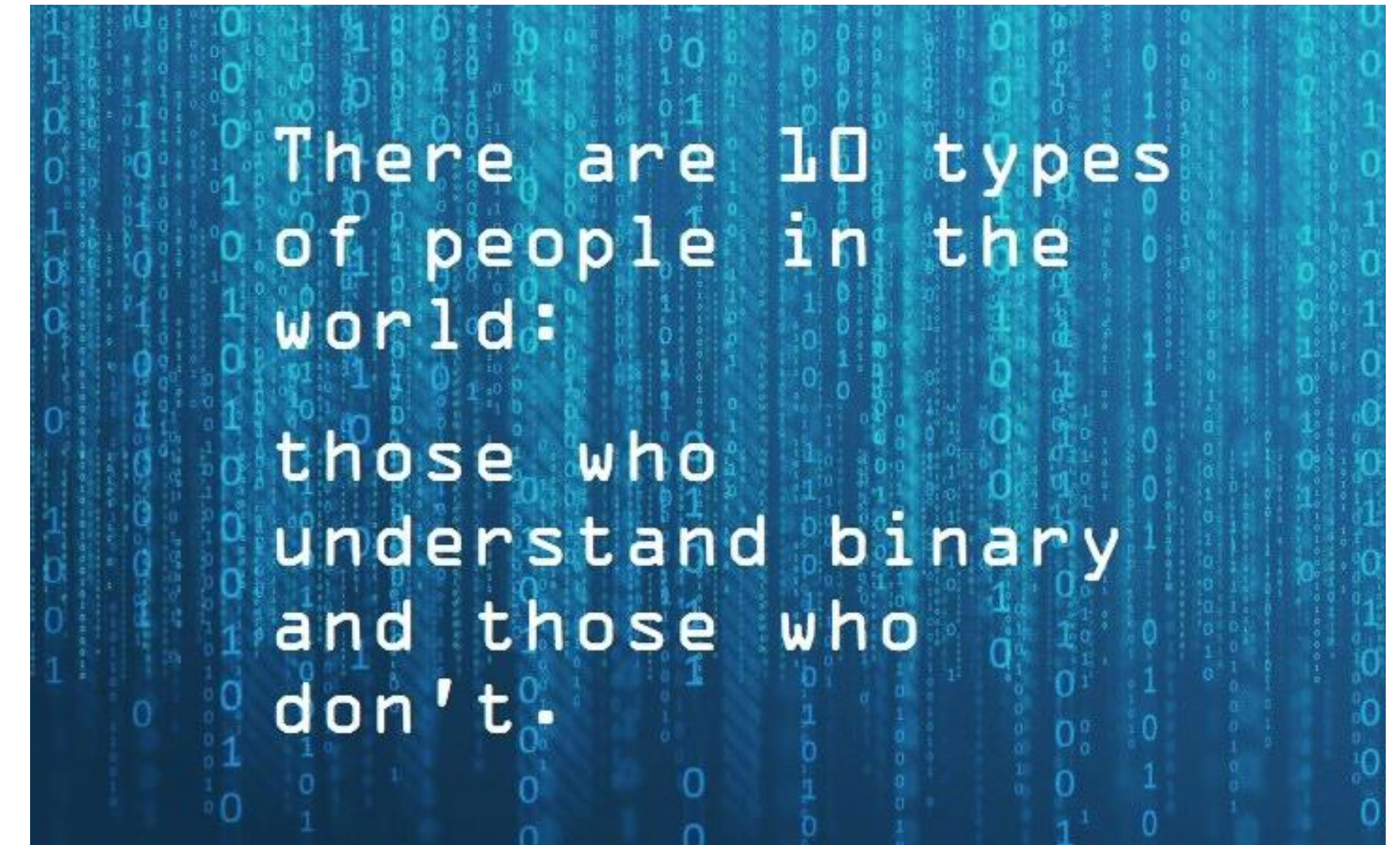
✓	Use case
	My application connects with an external partner
	I use 3rd-party utilities or open-source code in my application
	My application has a client or a mobile app installed on the end-user's device
	My users connect with mobile devices
	My code contains literal IP addresses
	My application uses broadcast
	My QA scripts use IP addresses
	My application allows users to directly connect to each other (peer-to-peer communication)
	I write IPv6 addresses in text
	I use a third-party cloud environment for QA or Production

Challenge 1: Changing datatype for IP address

With IPv6, you can't store IP address as string anymore.

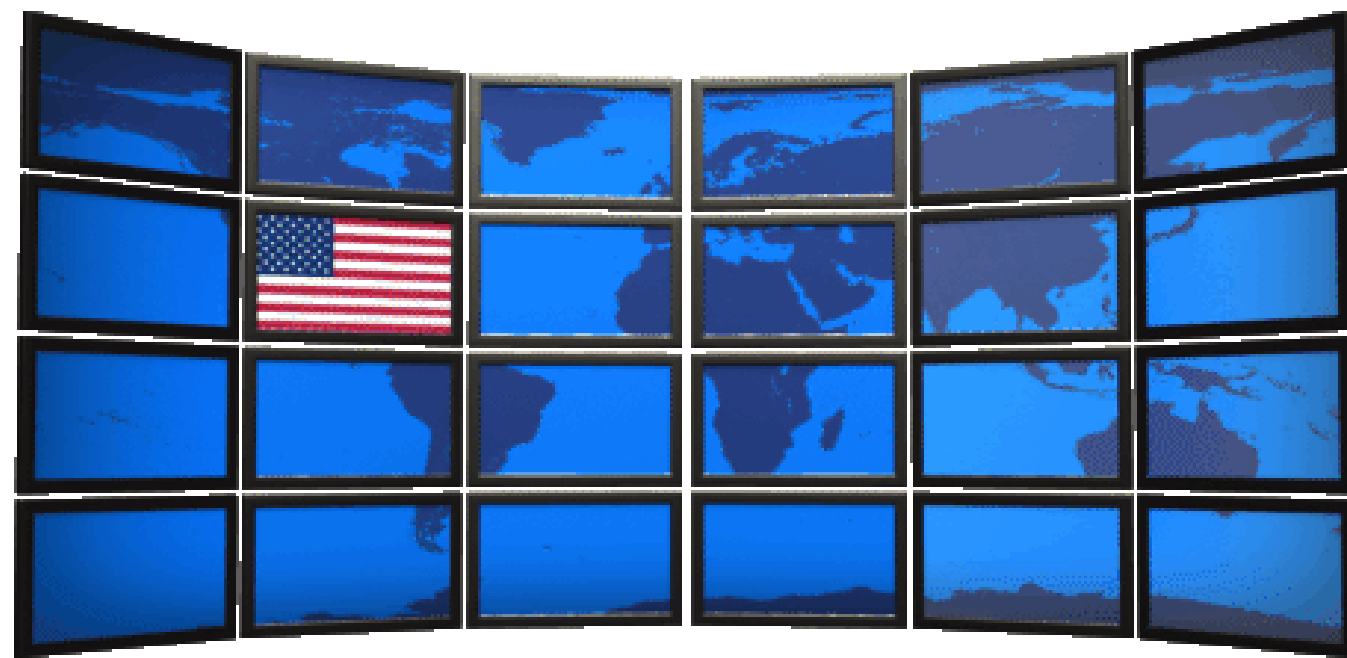
But how to make the change?

- Changed from string to byte array.
- Added a new IP field instead of modifying the existing field. Data jobs needed time to adjust to the new format without disruption.
- Stored v4 as binary also, so that the old field can be deprecated later
- Used v4-mapped v6 addressing in order to get a fixed length for both v4 & v6, which helps data analysis



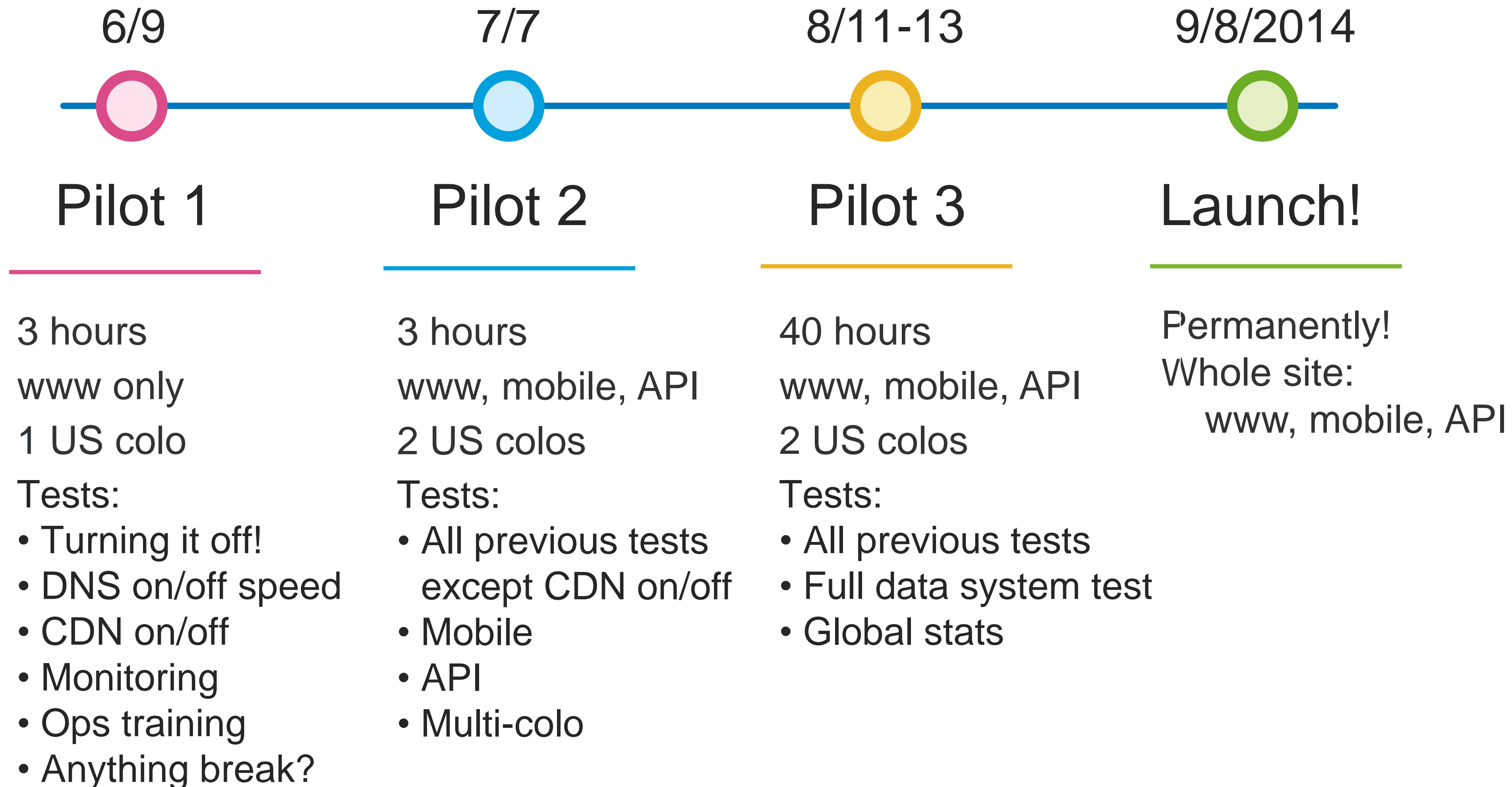
Challenge 2: IP-based geo-location

PHASE 1: SERVING IPv6 EXTERNALLY



- Poor quality of IPv6 geo-location data
 - Country level is good
 - ASN view is not good
- Used org in place of ASN
- Compared 2 geo vendors.
Coverage was roughly the same.

Launch sequence




Phase 1 accomplished!

PHASE 1: SERVING IPv6 EXTERNALLY

Win! →

← → ↻ <https://www.vyncke.org/ipv6status/allgreen.php> ☆ 🔍 📄 6° L B

IPv6 Deployment Status: the best IPv6 Sites

Name	Alexa	Web	Mail	DNS
 linkedin.com	17	www.linkedin.com 2016-04-13	mail-c.linkedin.com 2016-06-01	ns3.linkedin.com ns2.linkedin.com ns1.linkedin.com ns3.p43.dynect.net ns4.linkedin.com ns1.p43.dynect.net 6/8 2016-06-14
 yandex.ru	23	www.yandex.ru 2016-03-26	mx.yandex.ru 2016-03-26	ns2.yandex.ru ns1.yandex.ru 2/2 2016-06-14
 diply.com	52	www.diply.com 2016-06-07	alt1.aspmx.l.google.com 2016-03-26	a1-174.akam.net 1/6 2016-03-26
 xhamster.com	82	www.xhamster.com 2016-03-26	ALT1.ASPMX.L.GOOGLE.COM 2016-03-26	ns1.p29.dynect.net ns3.p29.dynect.net 2/4 2016-06-14
 C2C slideshare.net	169	www.slideshare.net 2016-04-13	aspmx3.googlemail.com 2016-03-26	ns2.linkedin.com ns1.p43.dynect.net ns3.linkedin.com ns3.p43.dynect.net ns4.linkedin.com ns1.linkedin.com 6/8 2016-06-14

PHASE 2: DUAL-STACK DATACENTER

1. Why build a dual-stack datacenter?
2. Challenges:
 - Scale
 - Operational tooling – what does “supports IPv6” mean?



Why build dual-stack?

PHASE 2: DUAL-STACK DATACENTER

Building a dual-stack datacenter before v6-only will let us:

1. Work out operational aspects
2. Train staff
3. Allow applications to run on v6 as soon as they're ready
4. Work with vendors to support v6
5. Build or obtain IPv6 operational tooling

Challenge 1: IPv6 at scale

PHASE 2: DUAL-STACK DATACENTER

Currently, it's harder for large company to implement IPv6 than a small one.

IPv6 for small org

A few servers with IPv6 that can be manually configured

VS

IPv6 for large org

Redundant and reliable web services

- Load balancer with IPv6
- Firewall with IPv6
- DDOS protection with IPv6
- Worldwide monitoring with IPv6

Challenge 2: IPv6-ready tooling

PHASE 2: DUAL-STACK DATACENTER

For tooling, what does “supports IPv6” mean?

1. Feature parity with IPv4
2. Runs natively on IPv6

We need to run all aspects of a device on IPv6:

- Provisioning
- Management
- Monitoring

Need IPv6-ready tooling

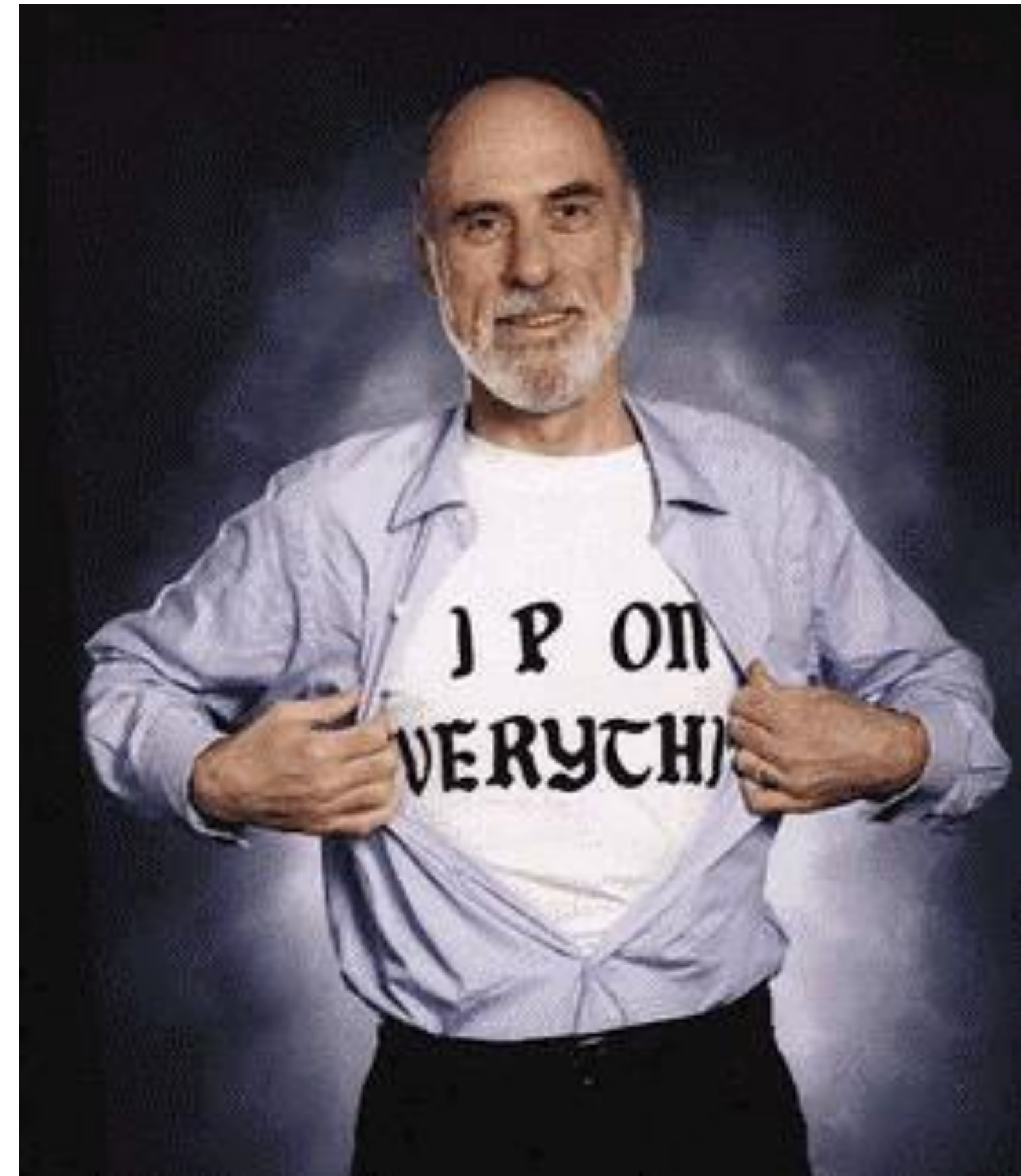
PHASE 2: DUAL-STACK DATACENTER



- Zero Configuration installation tools (PXE, TFTP,...)
- Automation tools (Puppet, Chef, CFE,...)
- Monitoring tools (Nagios, Cacti, IPMI,...)
- Management support tools (Term Console, PDU,..)
- Dynamic DNS (with SLAAC?)
- Hadoop
- Kafka
- Rest.li
- More ...

PHASE 3: IPv6-ONLY DATACENTER

Our goal!



LinkedIn IPv6 Measurements

Special thanks to:



Zaid Ali Kahn
Senior Director of Infrastructure Architecture and Strategy

Presented in March 2016
Networking Innovation and Research Symposium
Cisco – Ecole Polytechnique

More info: <https://www.linkedin.com/pulse/ipv6-measurements-zaid-ali-kahn?trk=prof-post>

Real User Monitoring (RUM)

- Synthetic monitoring doesn't work
 - Ex: Keynote, Gomez, Catchpoint, etc.
 - Limited coverage
 - Single user
- Real User Monitoring (RUM)
 - Beacon-based (Javascript/Boomerang lib)
 - Performance metrics from browsers of real LinkedIn users
 - Ex: TCP connect time, page download time

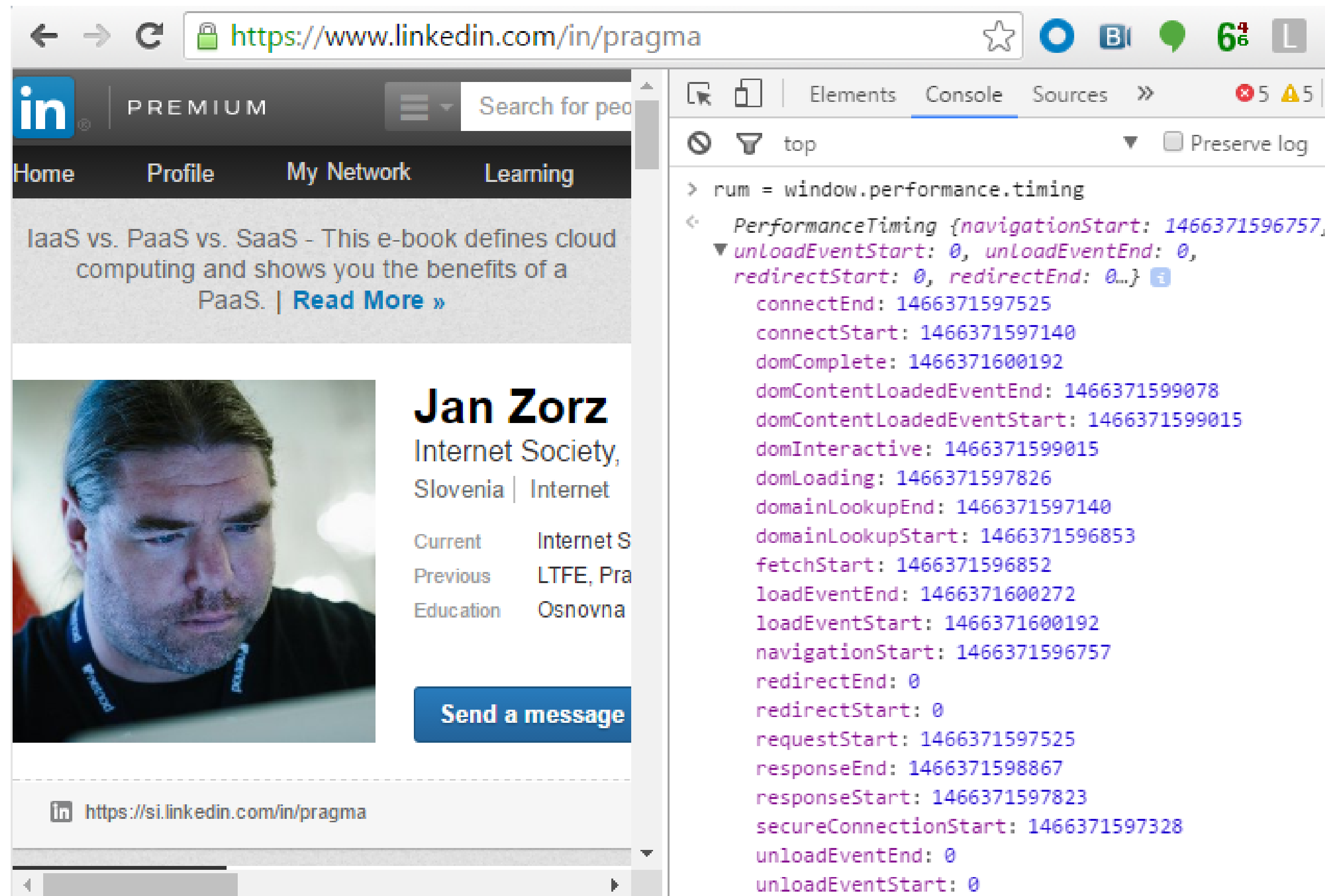
More info: Monitor and Improve Web Performance Using RUM Data Visualization

By David He, LinkedIn

<https://engineering.linkedin.com/performance/monitor-and-improve-web-performance-using-rum-data-visualization>

RUM – How

IPv6 MEASUREMENTS



Browse to www.linkedin.com on Chrome

Open “Console” on Chrome Dev Tools

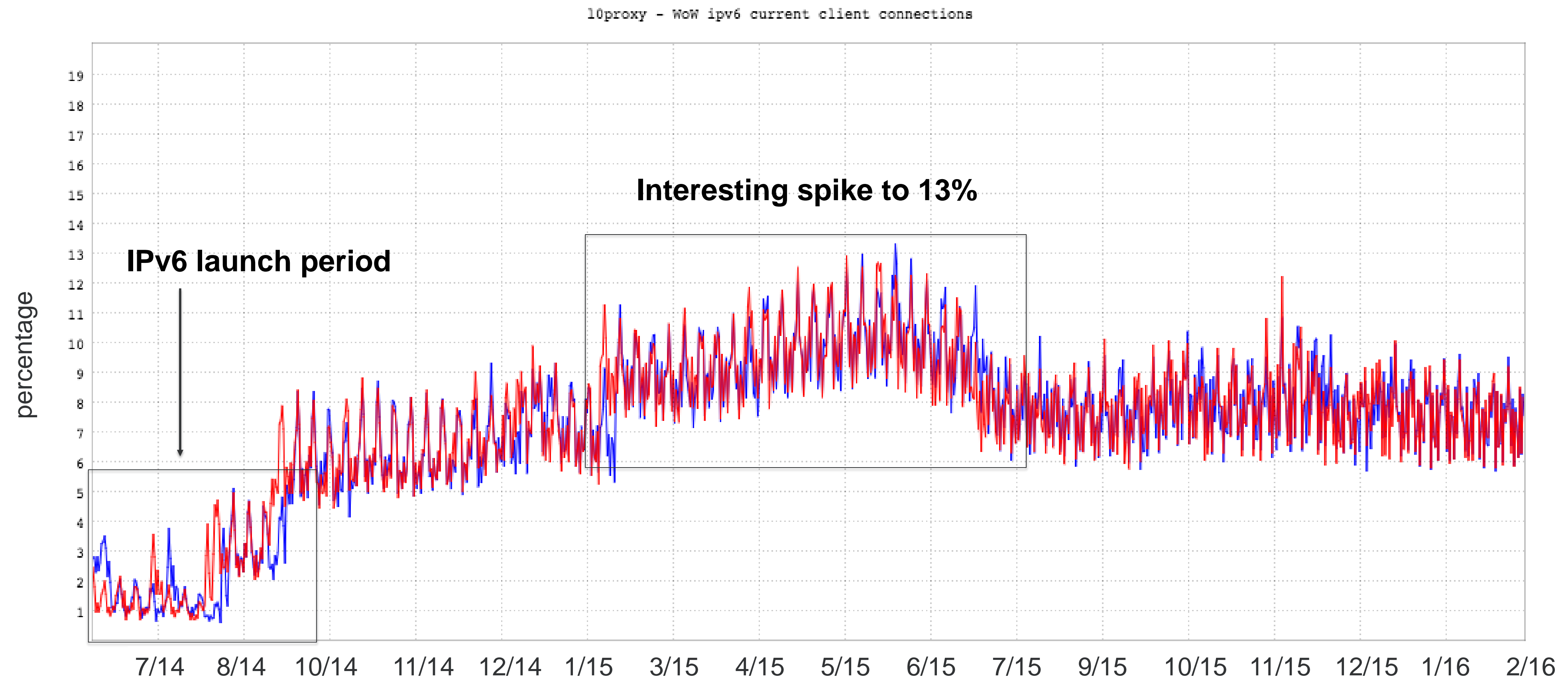
Mac : **Cmd + Opt + J**

Windows/Linux : **Ctrl + Shift + J**

```
rum = window.performance.timing
```


Global IPv6 growth at LinkedIn

IPv6 MEASUREMENTS

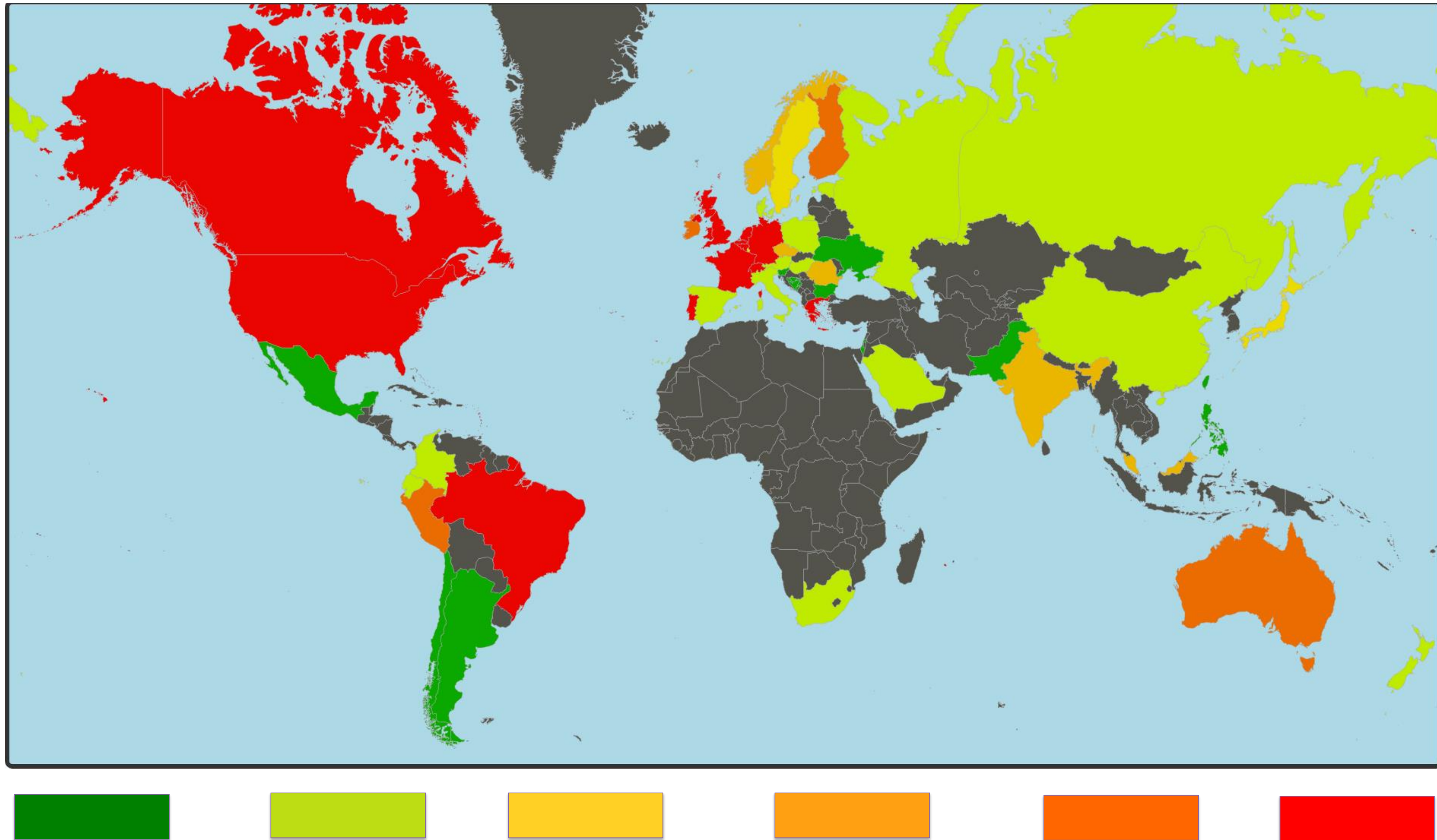


10% of LinkedIn traffic is IPv6

Similar to Google: <https://www.google.com/intl/en/ipv6/statistics.html>

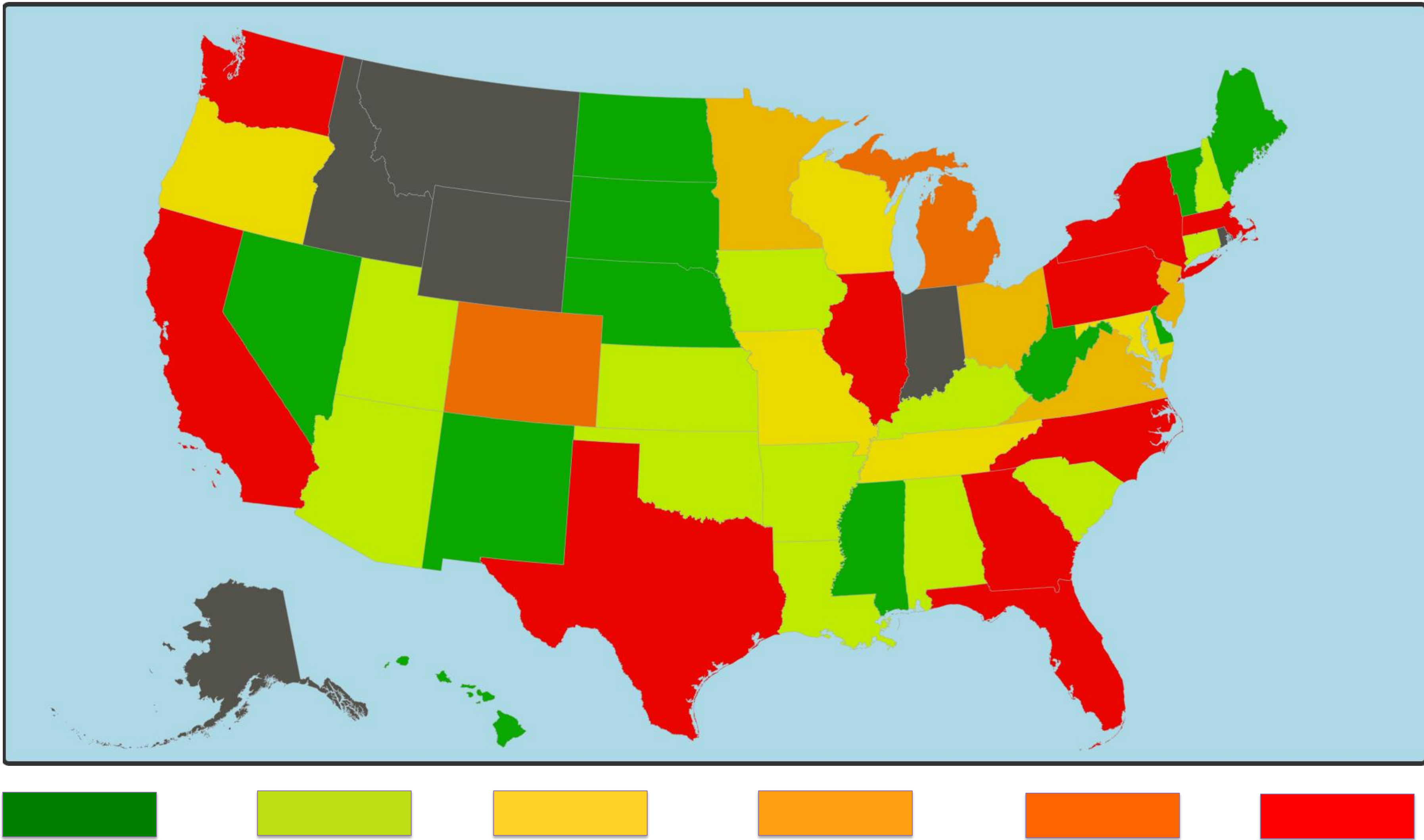
Global IPv6 heat map

IPv6 MEASUREMENTS



US IPv6 heat map

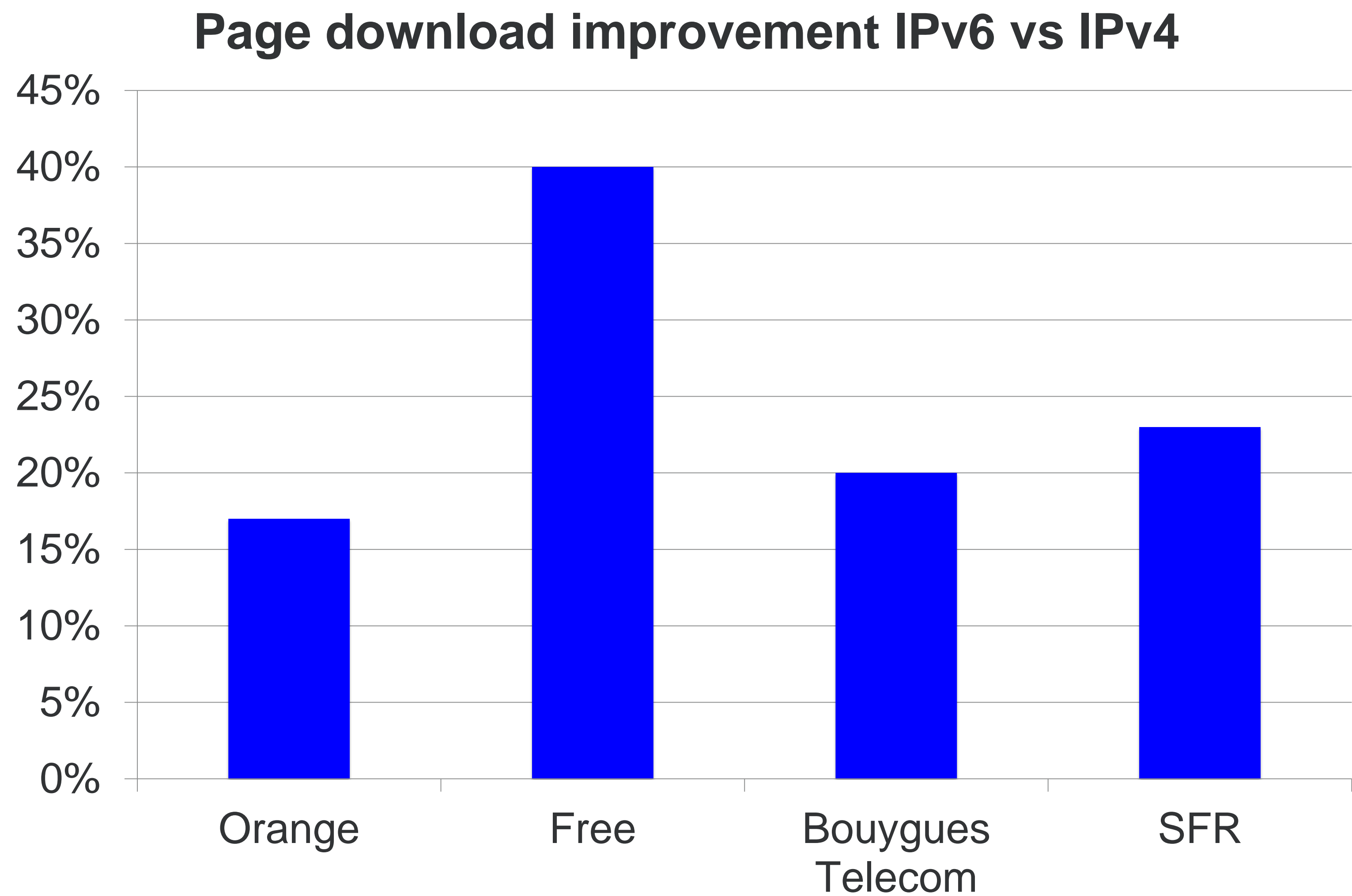
IPv6 MEASUREMENTS



Courtesy of Zaid Ali Kahn, LinkedIn

Europe mobile performance (France)

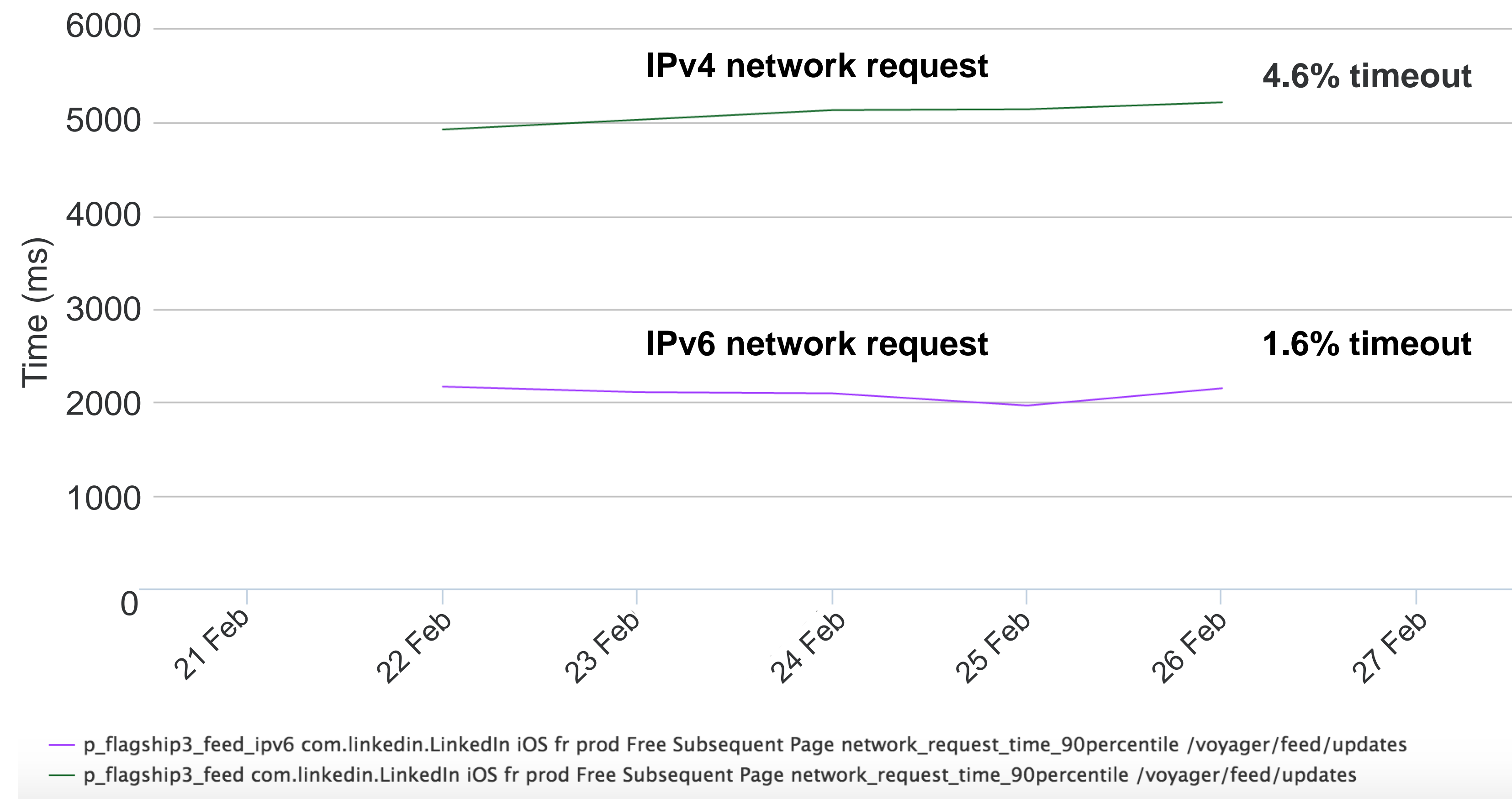
IPv6 MEASUREMENTS



Courtesy of Zaid Ali Kahn, LinkedIn

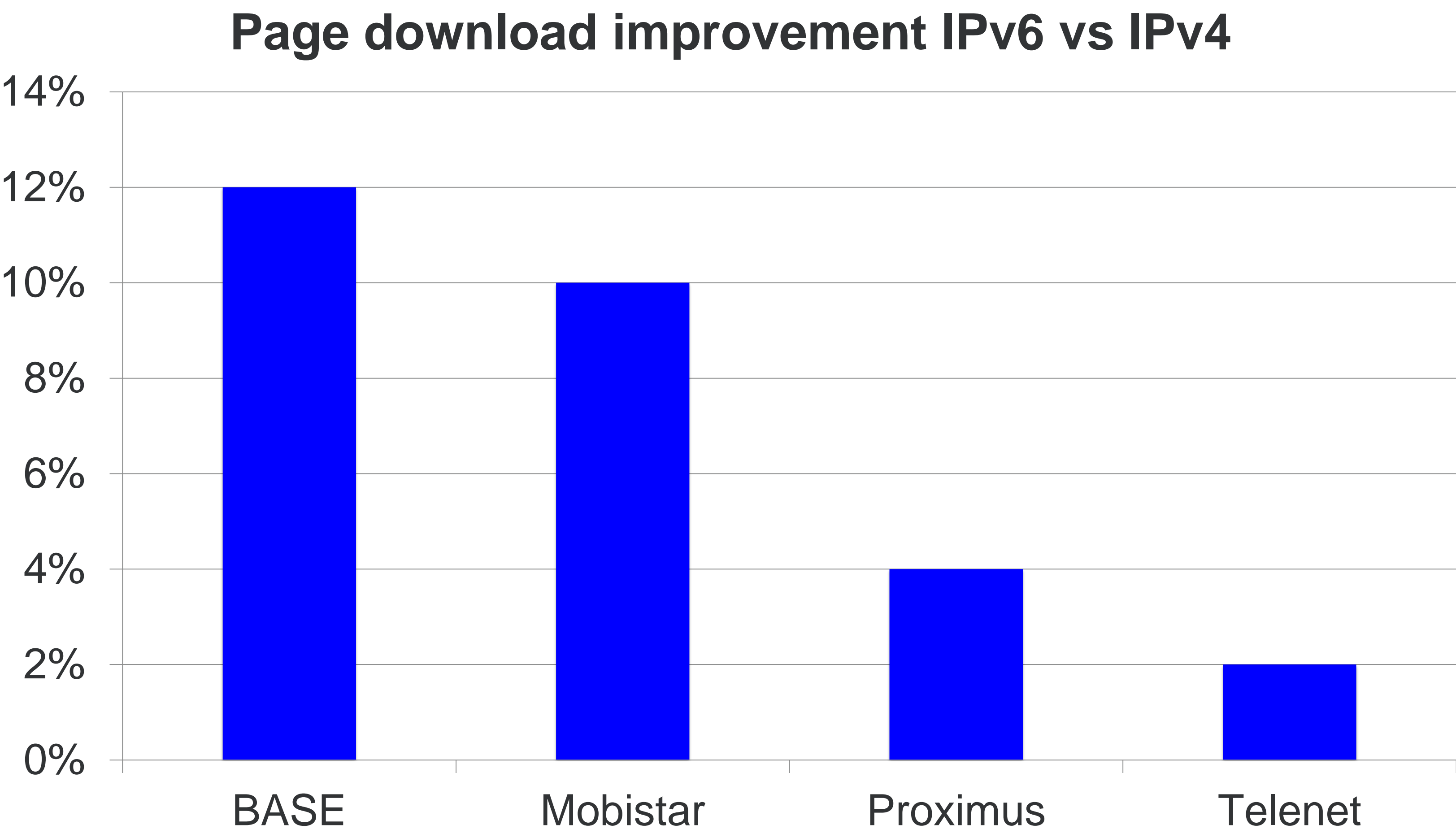
Why faster? Analysis of AS12322 (Free) **IPv6 MEASUREMENTS**

Higher network request time means higher TCP timeouts, yielding higher page download times.



Courtesy of Zaid Ali Kahn, LinkedIn

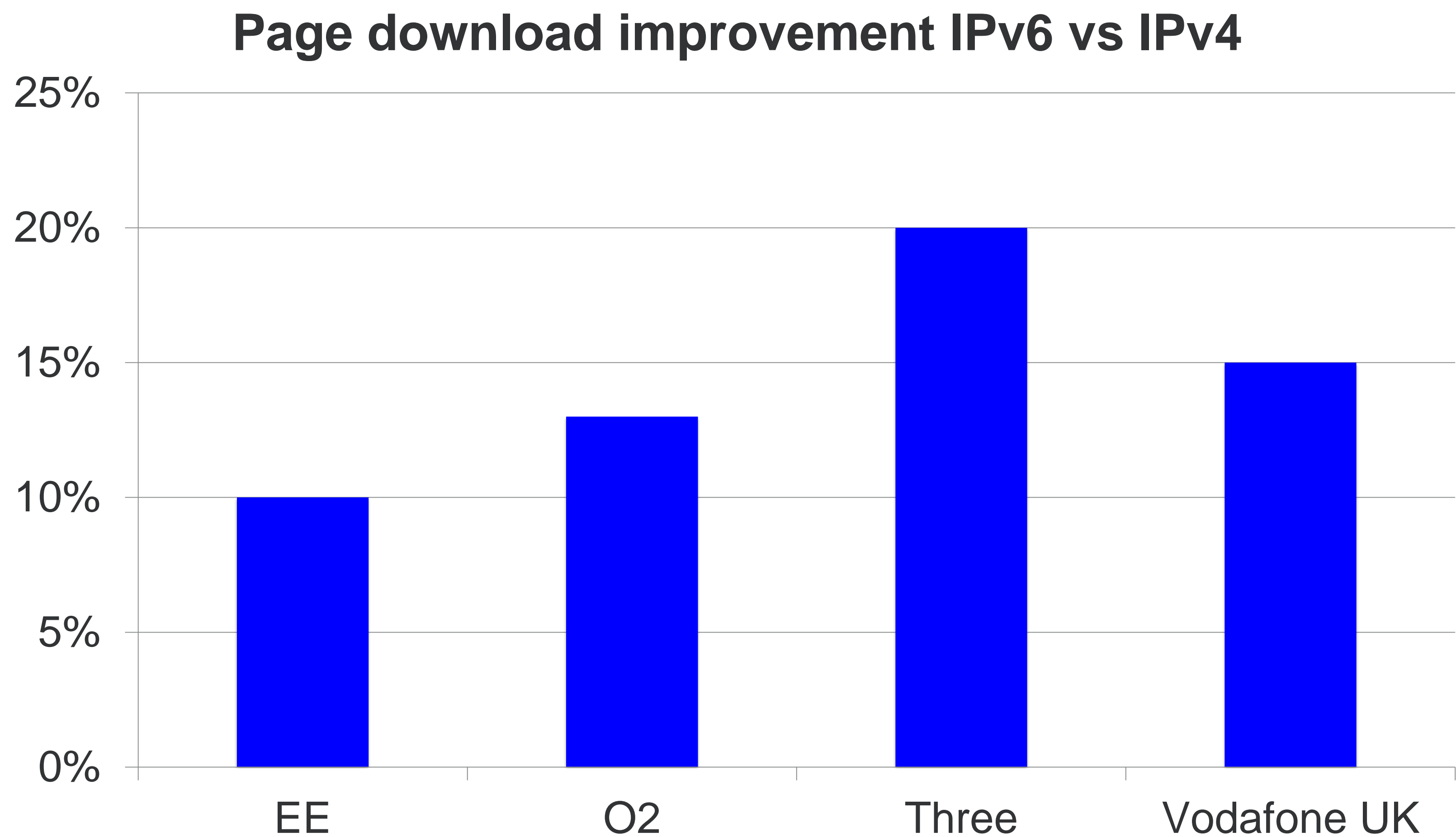
Europe mobile performance (Belgium) IPv6 MEASUREMENTS



Courtesy of Zaid Ali Kahn, LinkedIn

Europe mobile performance (UK)

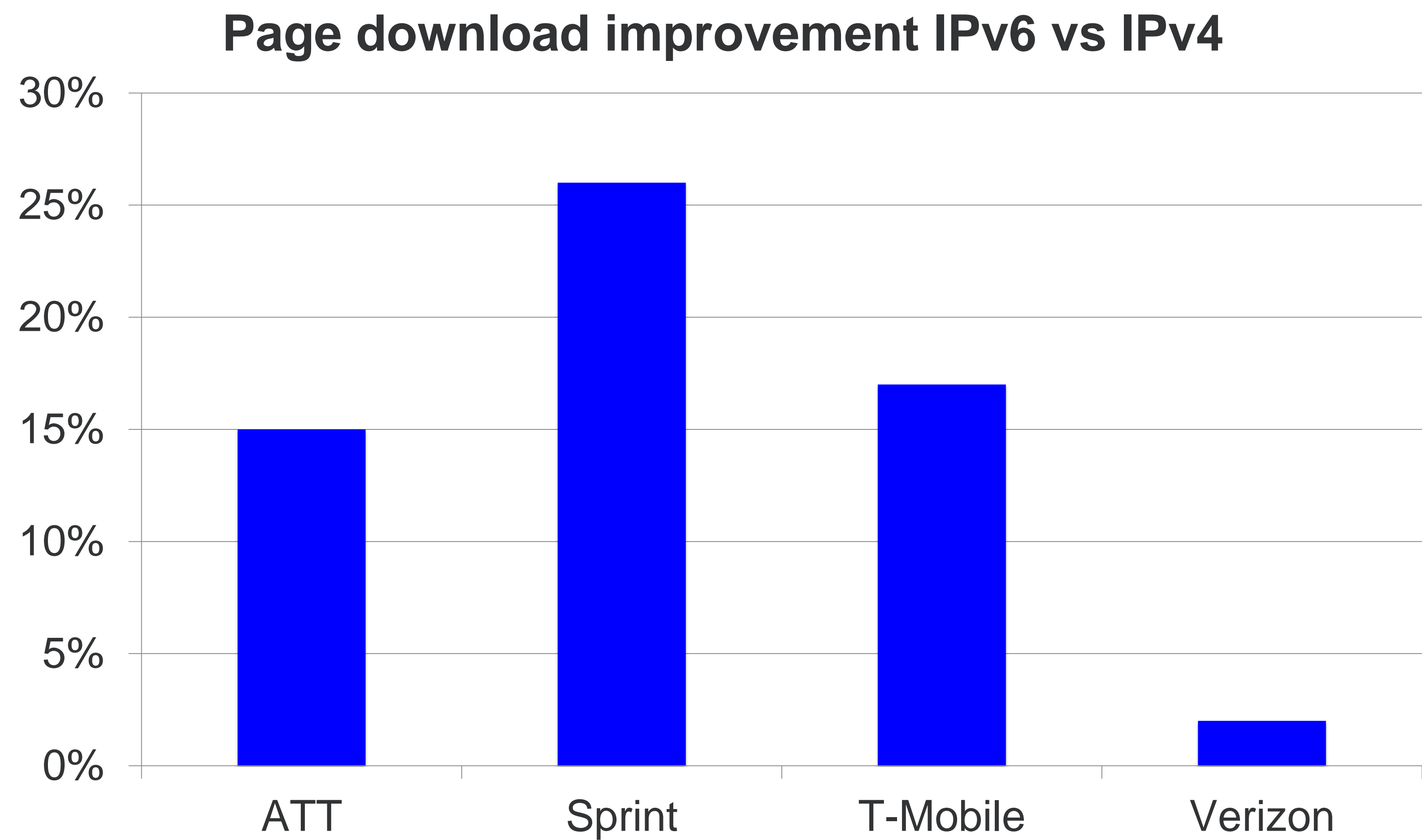
IPv6 MEASUREMENTS



Courtesy of Zaid Ali Kahn, LinkedIn

United States mobile performance

IPv6 MEASUREMENTS



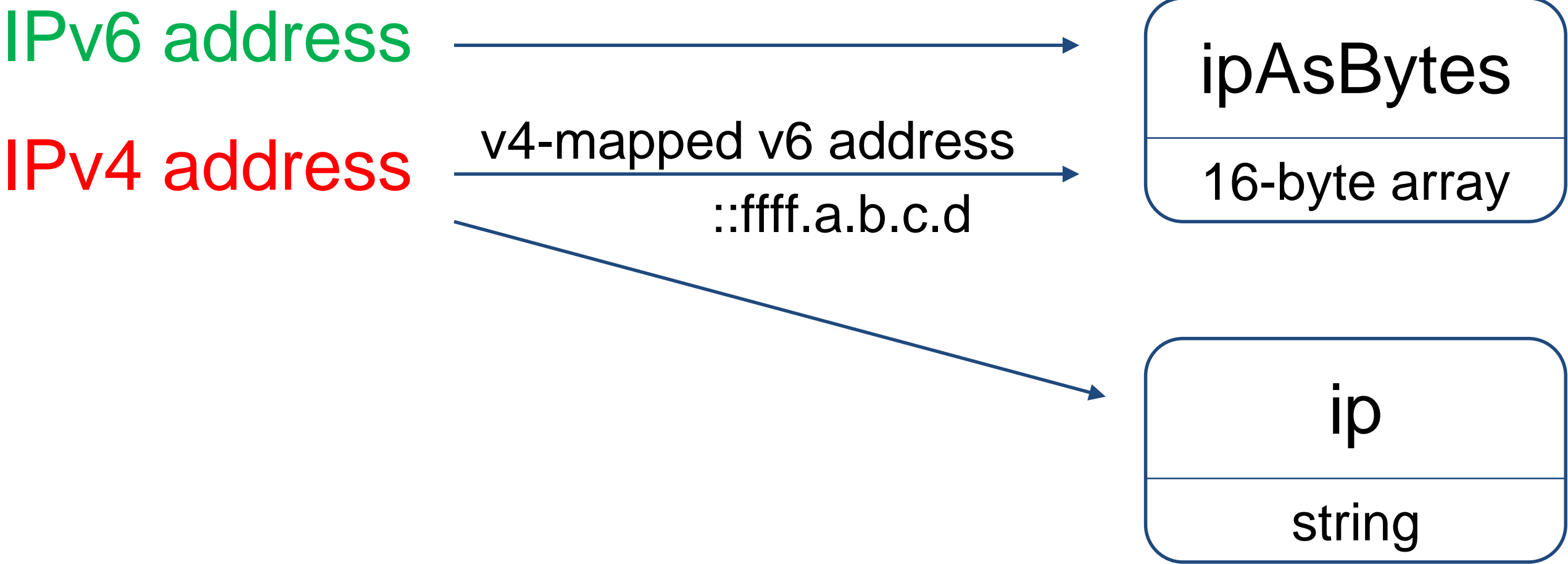
Courtesy of Zaid Ali Kahn, LinkedIn



String to binary transition

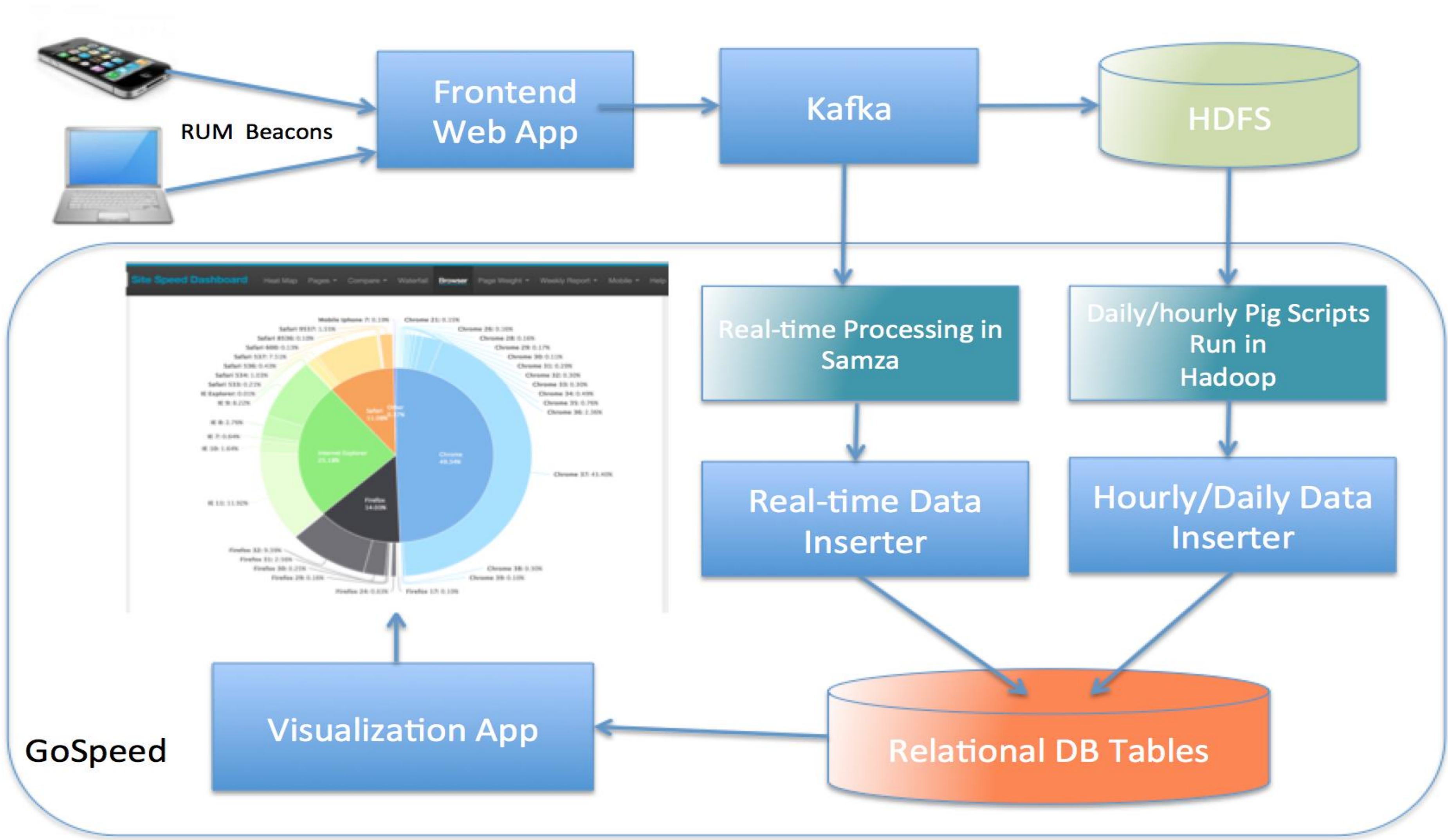
- 1. Add a new byte array field that will contain both IPv6 and IPv4 addresses
- 2. Retain existing string field containing IPv4 addresses (for backward compatibility)
- 3. Write IPv4 addresses to both fields
- 4. Write IPv6 addresses to new field only
- 5. Convert 2 years of historical IPv4 data to byte array (for continuity)
- 6. Data users modify their jobs to use new field. In the meantime, their jobs continue to use old field.
- 7. Deprecate old field when there's no more usage activity on it

Modify the library that generates Kafka events that include IP address, redeploy front ends



RUM telemetry architecture

IPv6 MEASUREMENTS



Courtesy of Zaid Ali Kahn, LinkedIn

RUM – How

IPv6 MEASUREMENTS

- Navigation timing API
 - Most browsers support it
 - Browsers record performance timers for each page load
- JavaScript timers
 - For everything else

RUM – How

```
rum = window.performance.timing
```

Connect Time:

```
rum.connectEnd - rum.connectStart
```

Download Time:

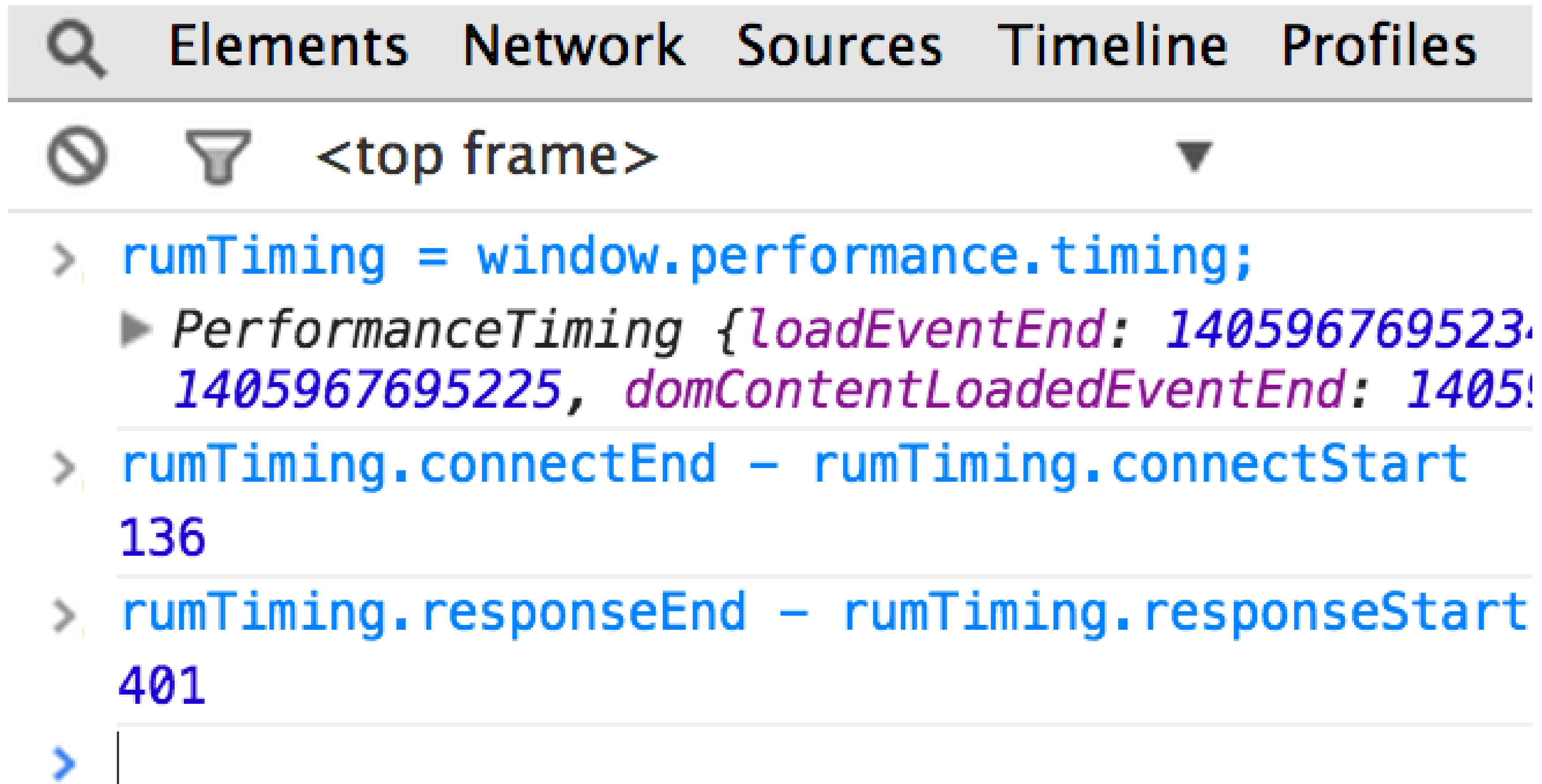
```
rum.responseEnd - rum.responseStart
```

Request to Response time:

```
rum.responseEnd - rum.requestStart
```

RUM – How

IPv6 MEASUREMENTS



The screenshot shows the Chrome DevTools Performance tab. The top navigation bar includes 'Elements', 'Network', 'Sources', 'Timeline', and 'Profiles'. Below this, the filter is set to '<top frame>'. The main area displays a list of performance metrics:

- `> rumTiming = window.performance.timing;`
 - `▶ PerformanceTiming {loadEventEnd: 140596769523.1405967695225, domContentLoadedEventEnd: 1405967695225, ...}`
- `> rumTiming.connectEnd - rumTiming.connectStart`
`136`
- `> rumTiming.responseEnd - rumTiming.responseStart`
`401`
- `>`

Simple IPv6 explanation

MAKING THE BUSINESS CASE

For a technical colleague

Irish Telecom Infographic: IPv4 vs IPv6 – Everything you need to know!

<http://www.networkworld.com/article/2692482/ipv6/infographic-ipv4-vs-ipv6.html>

For everyone

Internet Society: IPv6 and Why It Matters <http://www.internetsociety.org/what-we-do/internet-technology-matters/ipv6>

Simple IPv6 explanation

For Mom

Tested on a real mom!
(mine)

Simple IPv6 explanation

For Mom

Tested on a real mom!
(mine)

Simple IPv6 explanation

For Mom

Tested on a real mom!
(mine)

MAKING THE BUSINESS CASE

Simple IPv6 explanation

For Mom

Tested on a real mom!
(mine)

MAKING THE BUSINESS CASE

Simple IPv6 explanation

For Mom

Tested on a real mom!
(mine)

MAKING THE BUSINESS CASE

Simple IPv6 explanation

For Mom

Tested on a real mom!
(mine)

MAKING THE BUSINESS CASE

Simple IPv6 explanation

For Mom

Tested on a real mom!
(mine)

MAKING THE BUSINESS CASE

Simple IPv6 explanation

For Mom

Tested on a real mom!
(mine)

MAKING THE BUSINESS CASE

Simple IPv6 explanation

For Mom

Tested on a real mom!
(mine)

MAKING THE BUSINESS CASE

Simple IPv6 explanation

For Mom

Tested on a real mom!
(mine)

MAKING THE BUSINESS CASE

Simple IPv6 explanation

For Mom

Tested on a real mom!
(mine)

MAKING THE BUSINESS CASE