The Path to 100+ IXes

Kams Yeung, Senior Network Architect
kams@akamai.com

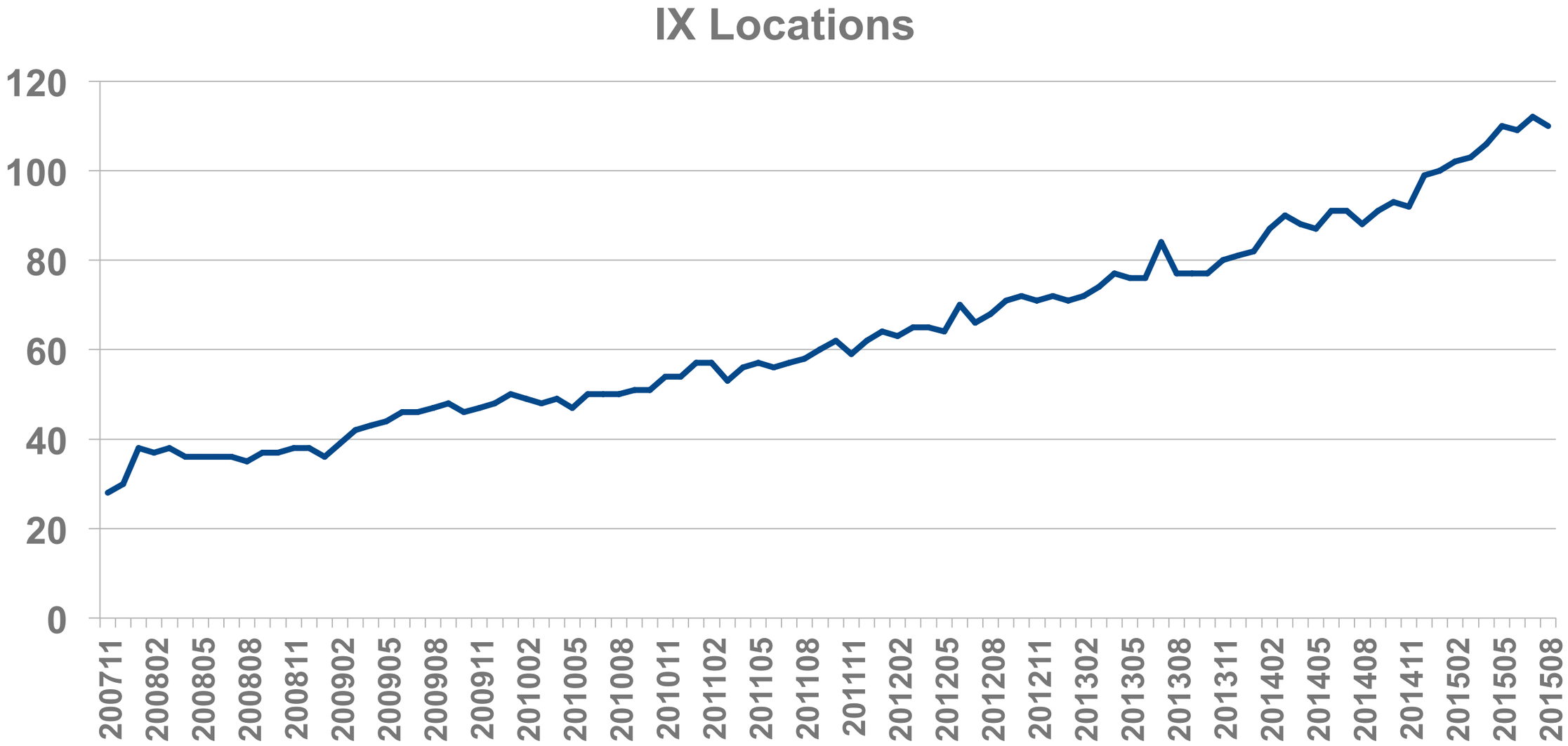
Introduction

- Akamai is the first CDN to pass traffic at over 100 IX locations.
  - This took 15 years!
- What/Why/How/When do we go to IXes?
- How does this fit into general trends in localization of traffic?
- What's next?

Where we started

- First IX in Americas: **PAIX** (~2000)
- First IXes in EMEA
  - UK: **LINX** (2001)
  - Continental Europe: **DECIX & AMSIX** (2001), **ESPANIX** (2002)
- First IX in APJ: **JPIX** (Early 2002)

By The Numbers: IX Locations

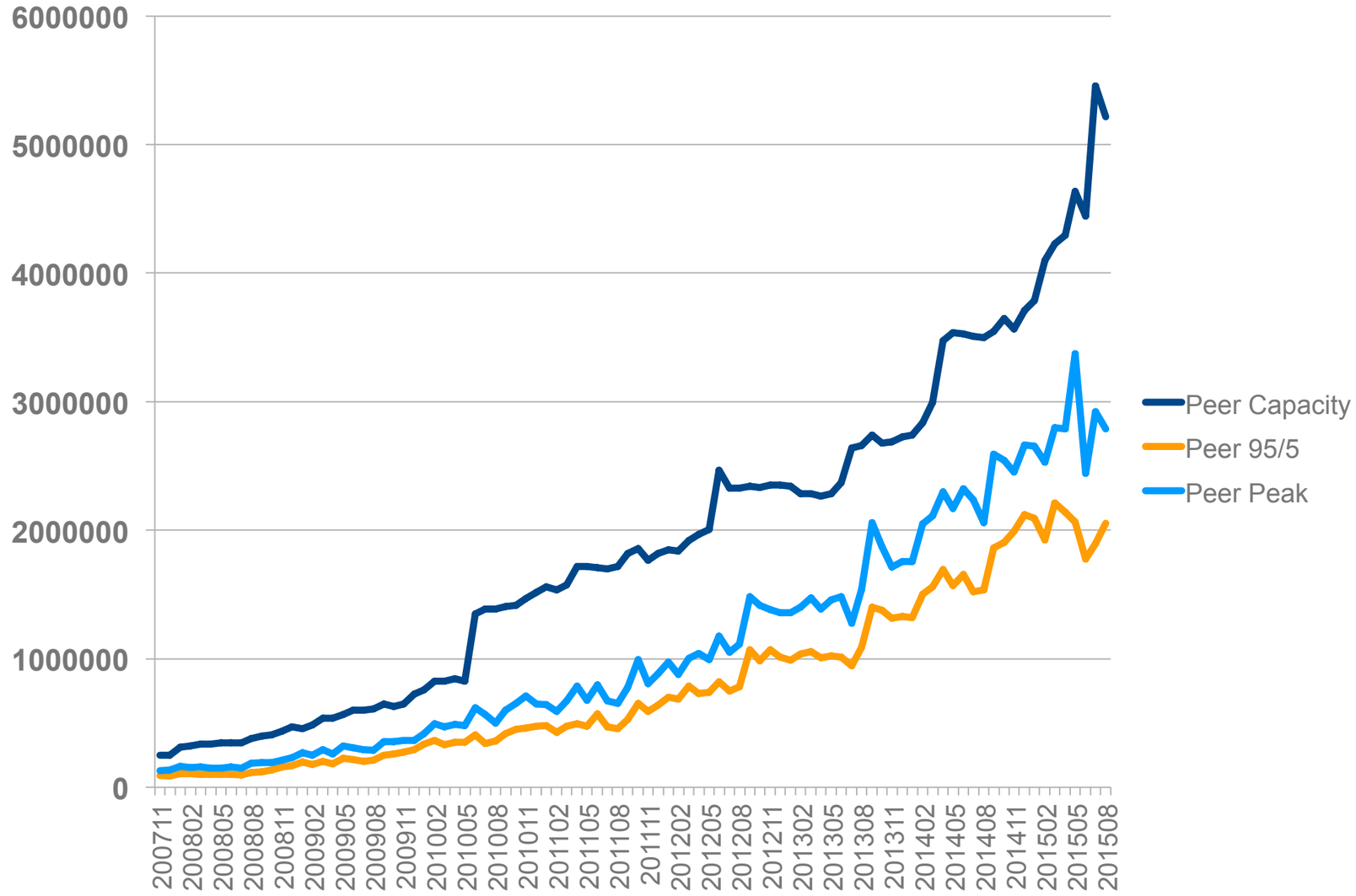


By The Numbers : Traffic

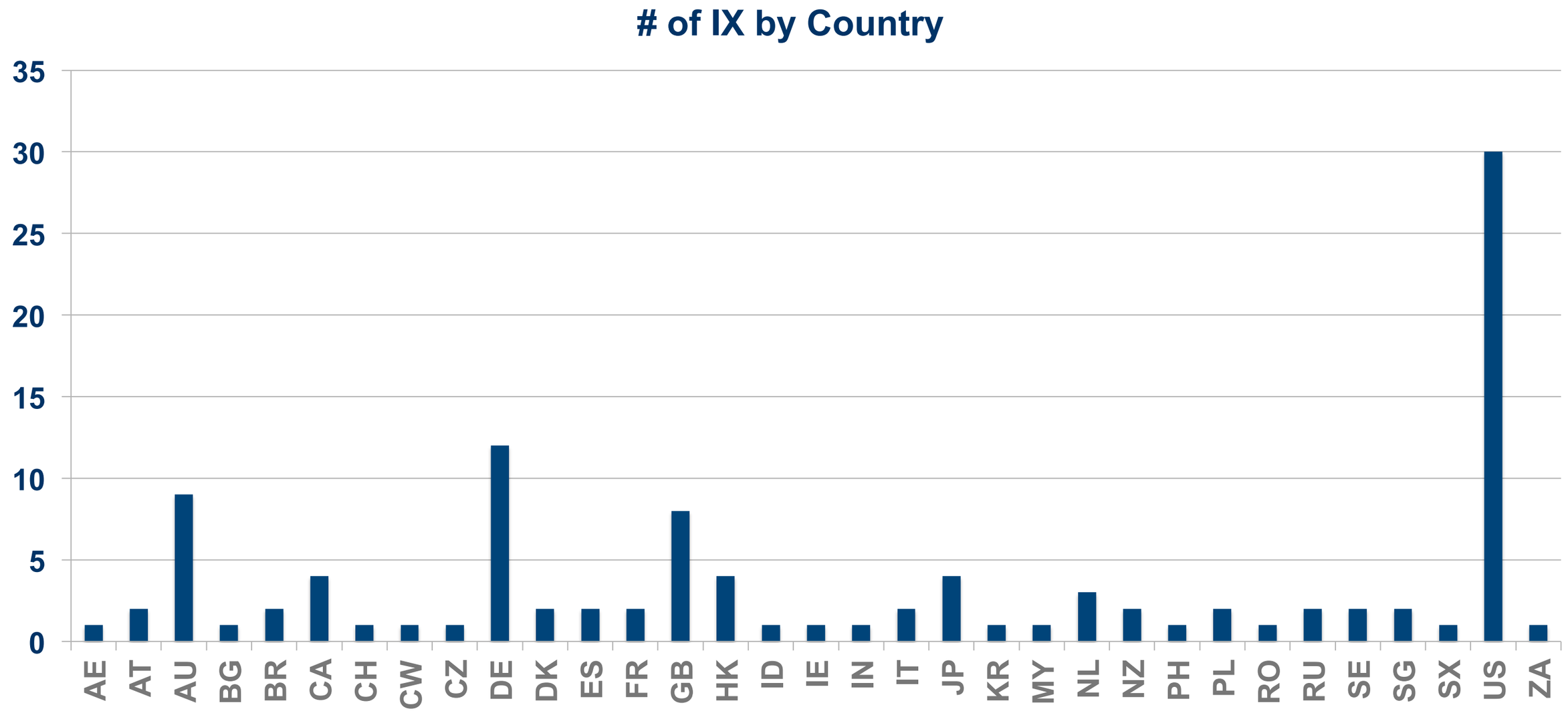
5.21 Tb of *public* peering capacity

110 locations

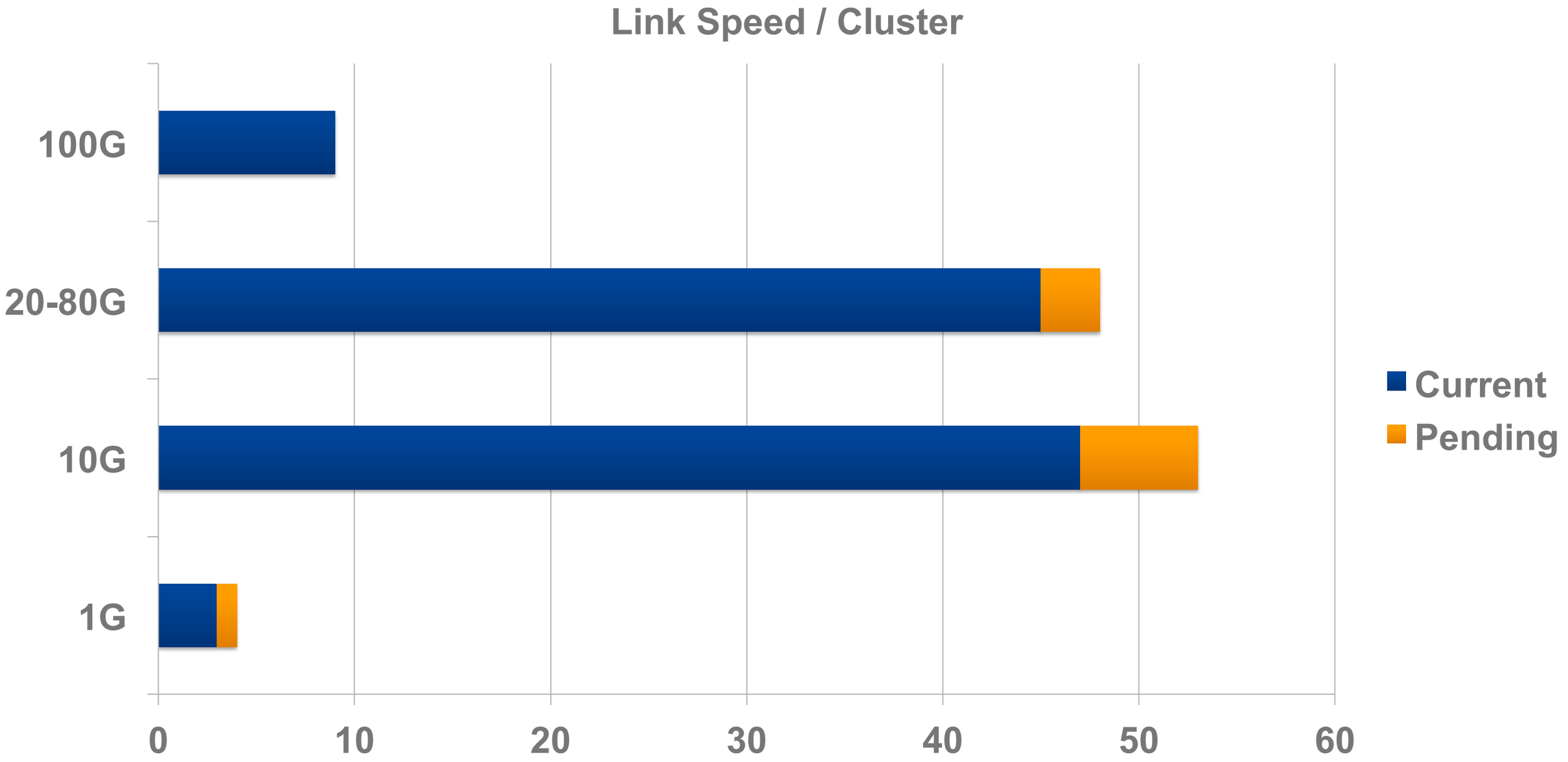
33 countries



By The Numbers: Clusters



By The Numbers: Technology



The Value Proposition for Peering

- Peering typically costs less than transit
  - The peering link is a predictable/fixed cost

- Lower distance = higher performance and reliability
  - High-bandwidth applications **require** low RTTs.
    - A stable 15 Mbps stream (typical of 4K video) **requires** < 34 ms RTT
      - And that's if there's no loss or retransmits.

Why do we join an IX?

- Reduces our costs
  - Less transit
  - Because we're a CDN, it typically means less transit for the IX members as well.
- Better overall performance for peers
  - Latency, topological diversity, overall throughput
- Geographical coverage/capacity
- Encourage the provider ecosystem in an area

Design Considerations

- One vs. Two LANs
  - Some IXs have two separate LANs with different members and capacities
    - e.g. LINX Juniper/Extreme,
    - In those cases, we will typically have separate clusters on each LAN.
    - Clusters may be different sizes, based on anticipated capacity.

- Akamai's Mapping allows serving traffic to a peer on all clusters at the same time

Design Considerations (cont'd)

- One vs. Multiple clusters in the same LAN
  - Multiple clusters instead of complex component redundancy
  - Same amount of peers on both clusters
  - Avoid issues with a single large cluster reallocating too much traffic
  - Always running active/active: we don't have idle spares

- Multiple locations instead of single locations
  - Put clusters in different physical locations to avoid physical fate-sharing, where possible.

Route Servers: pros and cons

- Route servers are a convenient way to jumpstart entry into an IX
  - Exchange routes with numerous IX members, with just 1 or 2 BGP sessions.
  - Preview how many routes are announced by members, for capacity planning.
  - *As a rule, we will peer with route servers when we enter an IX*

- Route servers also have drawbacks.
  - Fate Sharing for all BGP sessions
  - The setup and feature seat is different at each IX

In general, we will attempt to turn up sessions with other IX members as soon as we can.

IX vs. PNI: tradeoffs

- IX: shared bandwidth, low per-session costs
  - One-to-many relationship, turning up BGP sessions is cheaper & faster
  - **However,** IX fabric bandwidth is large, but members' links will be comparatively small
    - No control of the bandwidth or utilization of other peers until there is packet loss

- PNI: dedicated bandwidth, higher per-session costs
  - Reserved bandwidth/capacity, but cross connects and router ports cost money.
  - Limited resiliency: interruption of a single PNI is more likely during maintenance
    - A PNI only connects to a single cluster

At some point, it's cheaper to allocate a PNI than it is to take the traffic over the IX.

AANP-IX: An IX Operator Solution

- Dedicated set of Akamai servers, serving **only** the IX
  - Saves money on transit by serving traffic locally
  - Improved performance for IX members

- Attractive selling point for an IX to gain membership
  - More members = more traffic = more members

When do we help to “**build**” an IX?

- Multiple ways to “build” in IX.
  - Provide equipment/expertise for local talent that’s short on funds.
  - Create critical mass for an IX by being a “content anchor”.

- Growth potential of an area that would have lower costs with more cooperation
  - Adding in a content anchor gives them a reason to be in the same place.
  - Better peering = lower costs = more players and room for regional players to expand.
    - In addition, they get high-speed, low-latency access to content.

Akamai's IX Board Experience

Founding member:

Akamai board participation:

De-peering and leaving IX's

- We de-peered Networks
  - Because of broken route announcements

- Leaving IX's
  - Costs
  - Because they became dysfunctional (KleyerIX, FreeIX)

How Akamai Operates at an IX

- Inconsistent Routes
  - We only announce the local clusters
  - You don’t need all Akamai routes
- Best-effort delivery
  - No guarantees about what or how much traffic will be sent
- No SLA on any specific cluster
  - We can’t guarantee where any particular content will be sent from.
- High traffic outbound vs. inbound, because we’re a CDN

Issues

- Peers announcing the world
  - Usually gets noticed and handled quickly
  - Our mapping system can detect networks being far away
- IX route servers prepend their own AS (e.g. HKIX)
  - Makes route selection more difficult than it has to be.
- Bad BGP Traffic Engineering
  - Incomplete announcements, overlapping routes.
    - We might end up serving you over your transit.
  - **If you need to engineer traffic, *please talk to us*.** We have better tools than just BGP.

What does this all mean?

- 110 done, n to go.

- Getting traffic closer to the user is a more general trend
  - More local peering even in well-connected countries

- **If you're at an IXP with us, we will peer with you.**

peering@akamai.com

Special Thanks to…

- Matt Ringel and the rest of the Akamai Network Architecture Team for editing and feedback.

- Sean Butler, John Payne, and Patrick Gilmore for providing information from the beginning of Akamai.


Akamai
FASTER FORWARD