Routing optimisation without BGP

working with Akamai

to improve HTTP(S) traffic flows

Thomas Mangin – IXManchester 19th of April 2017
Thank you to Ronan Mullally from Akamai Technologies for his help

The players

Ronan Mullally, playing for team Akamai

A large CDN, "generator" of "desirable" content

And "Me", playing for team exa networks

An Akamai peer

The game

Better, Happier Peering

Exa Networks

Education & Business ISP

Eyeball Network

10 Gb National Core

- Manchester (first site), London, Leeds, Bradford, Sheffield
- More locations soon

100 Gb within POPs (using Arista 7280R series)

DSL/FTTC traffic

slightly over   2/3 comes from Equinix Manchester

slightly under 1/3 comes from Telehouse Docklands

Leased Lines

"even split" London, Manchester & Leeds

Dark Fiber

Bradford, Leeds, Sheffield, more to come

Exa Networks

Most traffic inbound
- **1/3 Akamai**
- **1/3 Google**
- **1/3 the rest** "the usual suspects"
  Apple, Microsoft, Netflix, Limelight, Amazon, Facebook
  And the *very* long tail

Our own 'in-house' **content filtering** solution
HTTP and HTTPS traffic - proxied & trans-proxied 'in the cloud'
Bringing back some traffic from London to Manchester

Very **open peering policy** (1,000+ IPv4 & IPv6 eBGP sessions)
We may need to reduce the number of sessions at some point

We provide Akamai with some **transit** for Akamai's IXLeeds cluster
(As it will be visible on our AS-Stats graphs)
Not selling AS30740 transit otherwise (but friends and family)

Exchange ▾ ASN	Speed RS Peer

Peering sessions with Akamai

IX Leeds	10G
30740	○
IXManchester	10G
30740	✓
LINX LON1 Main	10G
30740	✓
LONAP LON0	10G
30740	✓

Akamai traffic over a year
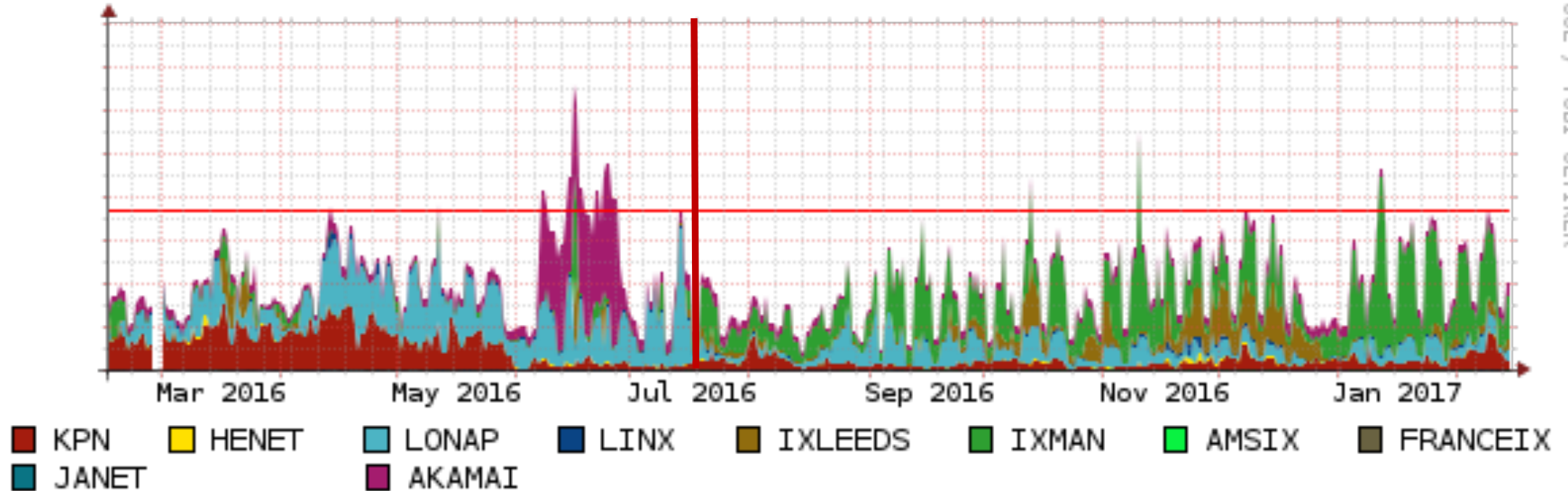
IXLeeds' Akamai cluster sending to Asia, using the transit link we provide them

AMS-IX via KPN

IXLeeds
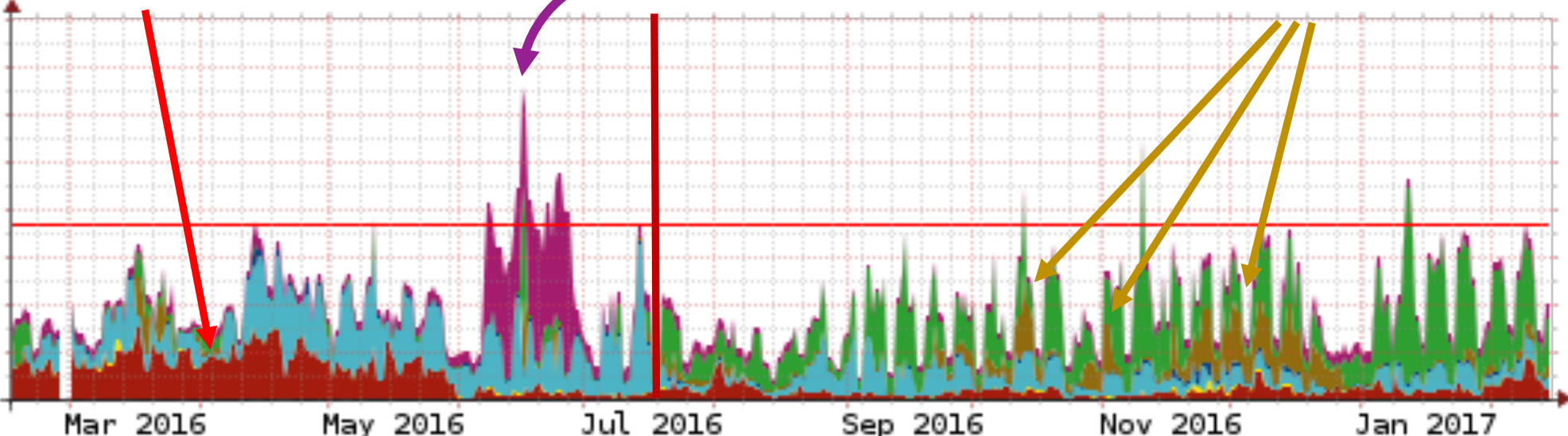
"Something" changed during the summer 2016 !

~~BGP~~ DNS Based Routing

Anycasted one recursive DNS servers,
per POP where we peer with Akamai:
- **London** (existed)
- **Manchester** (existed)
- **Leeds** (new)

Previously the setup was Active / Passive
with the active DNS server in Manchester.

A DNS server failure will cause
another POP DNS to be used.

ExaBGP

Eating our own dog food, ExaBGP is used
to anycast all our /32 service IPs,
including for DNS.

Detect DNS failure and stop announcing the service IP
should DNS fail to resolve.

Let Me Google That For You:
"exabgp healthcheck anycast DNS filetype:pdf"

PowerDNS

And we 😍 **DNSDist**.
You can do some **Fab** things using **Lua**.

Explanation: from the mouth of the network

(UK peering forum presentation)

How does it work?

The 'secret sauce' of a CDN is its ability to direct end-users to content

We call this Mapping

Proper Mapping makes the world a better place
- better performance
- lower cost
- happy networks
- Happy Users

Akamai maps using DNS lookups

Your client DNS
Hopefully yours, in your network

Akamai DNS

©2015 AKAMAI | *FASTER FORWARD™*

Explanation: from the mouth of the network

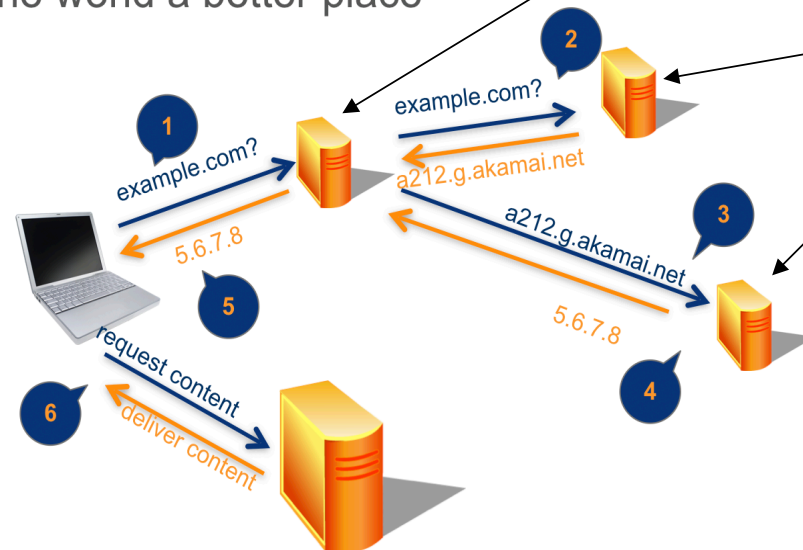(UK peering forum presentation)

How does it work?

The 'secret sauce' of a CDN is its ability to direct end-users to content

We call this Mapping

Proper Mapping makes the world a better place
• better performance
• lower cost
• happy networks
• Happy Users

Akamai maps using DNS lookups

©2015 AKAMAI | *FASTER FORWARD™*

Client Subnet in DNS Queries
https://tools.ietf.org/html/rfc7871

This information can be added by DNSDist
Did I say enough that we 😍DNSDist ?

Akamai monitor and maps your DNS servers
(path taken, latency, packet loss, …)
Packet loss/congestion on a link affect the decision
Reviewed every few hours

Akamai needs to know where your DNS are
(otherwise they will need to guess from latency)

Your user will be served based on its DNS resolver
(DNS in London or 8.8.8.8, traffic from London)

DNS traffic level matters
If a DNS is not generating enough traffic, it will be consolidated with others.

If you are using RFC 7871, let Akamai know, they may be able to take advantage of it.

"Auto-magical" changes

Not all Akamai clusters are born equal (content cached, size, etc.).

Here, Akamai decided to consider our Leeds DNS to be "logically" in Manchester (due to low DNS query volume).

Akamai's mapping is dynamic and changes to adapt to current conditions

- do not expect notifications
- the beast is "auto-magical"
- a bit like your peer "local-pref tuning" …

LINX Meetings are a great place to catch up with Akamai and discuss their current "policies".

16   Nov 2016   Jan 2017
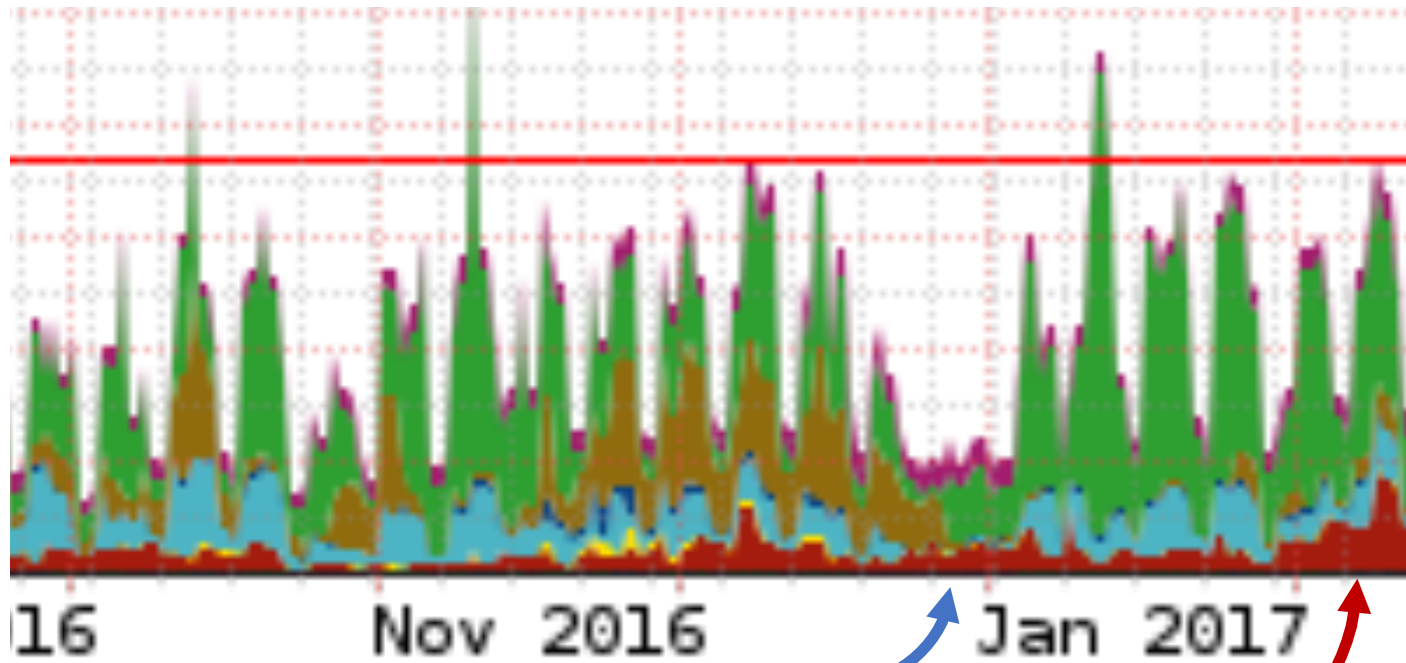
Another change

Most likely as we do not peer in AMS-IX with Akamai

~~DNS~~ $$$ Based Routing ?

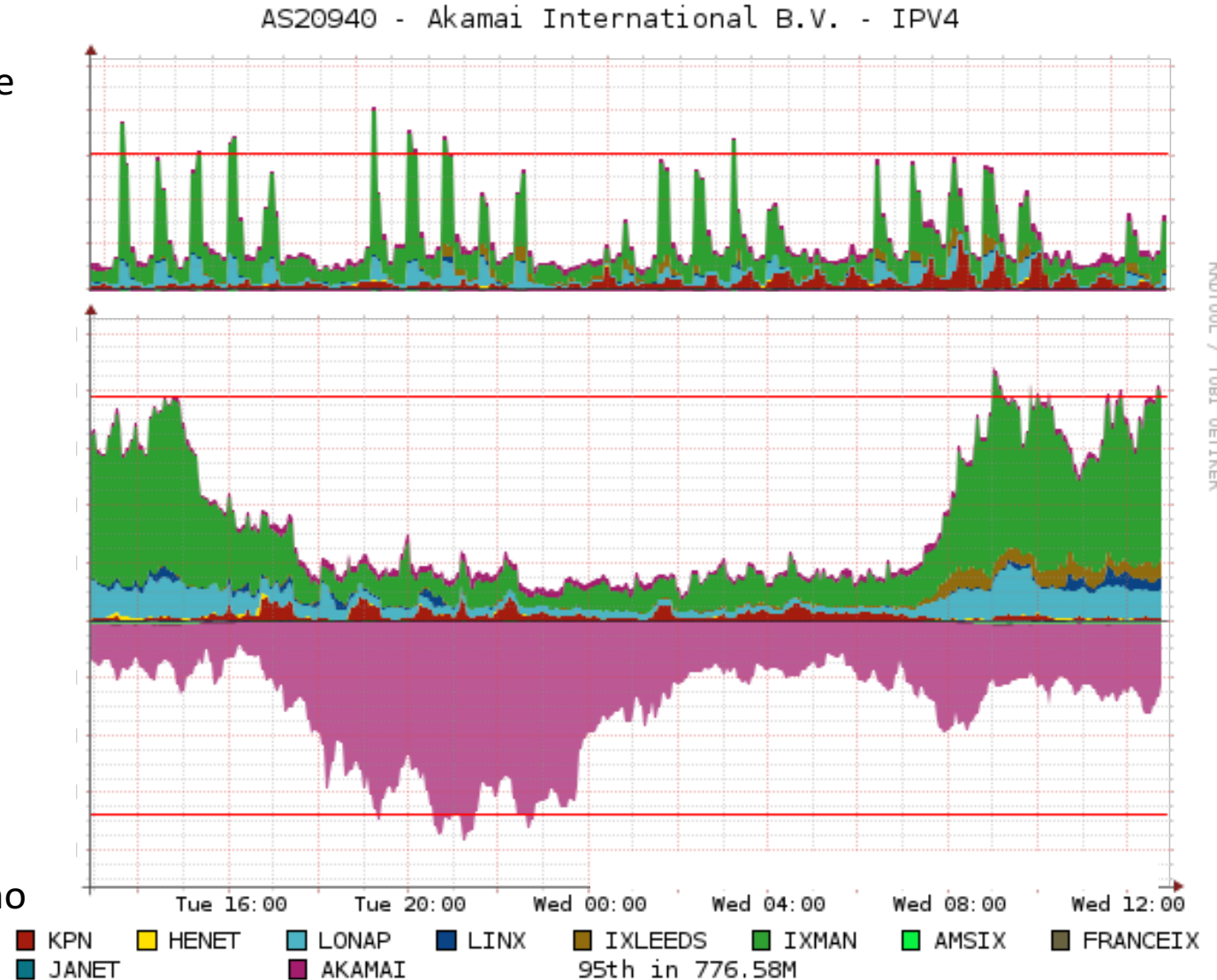Regional peering traffic level is **not** affected by the price of peering.

Traffic has been consistently coming from
    1 - IXManchester   (green)
    2 - LoNAP   (light blue)
    3 – KPN Transit in Manchester   (red)
    4 - IXLeeds   (brown'ish)
    5 - The rest

But during the data collection, pricing changed

    1 - (then) IXLeeds   1 - (now) IXLeeds
    2 - (then) LoNAP   2 - (now) IXManchester
    3 - (then) IXManchester   3 - (now) LoNAP

The price reductions at LoNAP and IXManchester had no impact on the routing

AS20940 - Akamai International B.V. - IPV4

RRDTOOL / TOBI OETIKER

KPN   HENET   LONAP   LINX   IXLEEDS   IXMAN   AMSIX   FRANCEIX
JANET   AKAMAI   95th in 776.58M

It's ancient data.. So now ?

Last presentation at LINX 96 in February
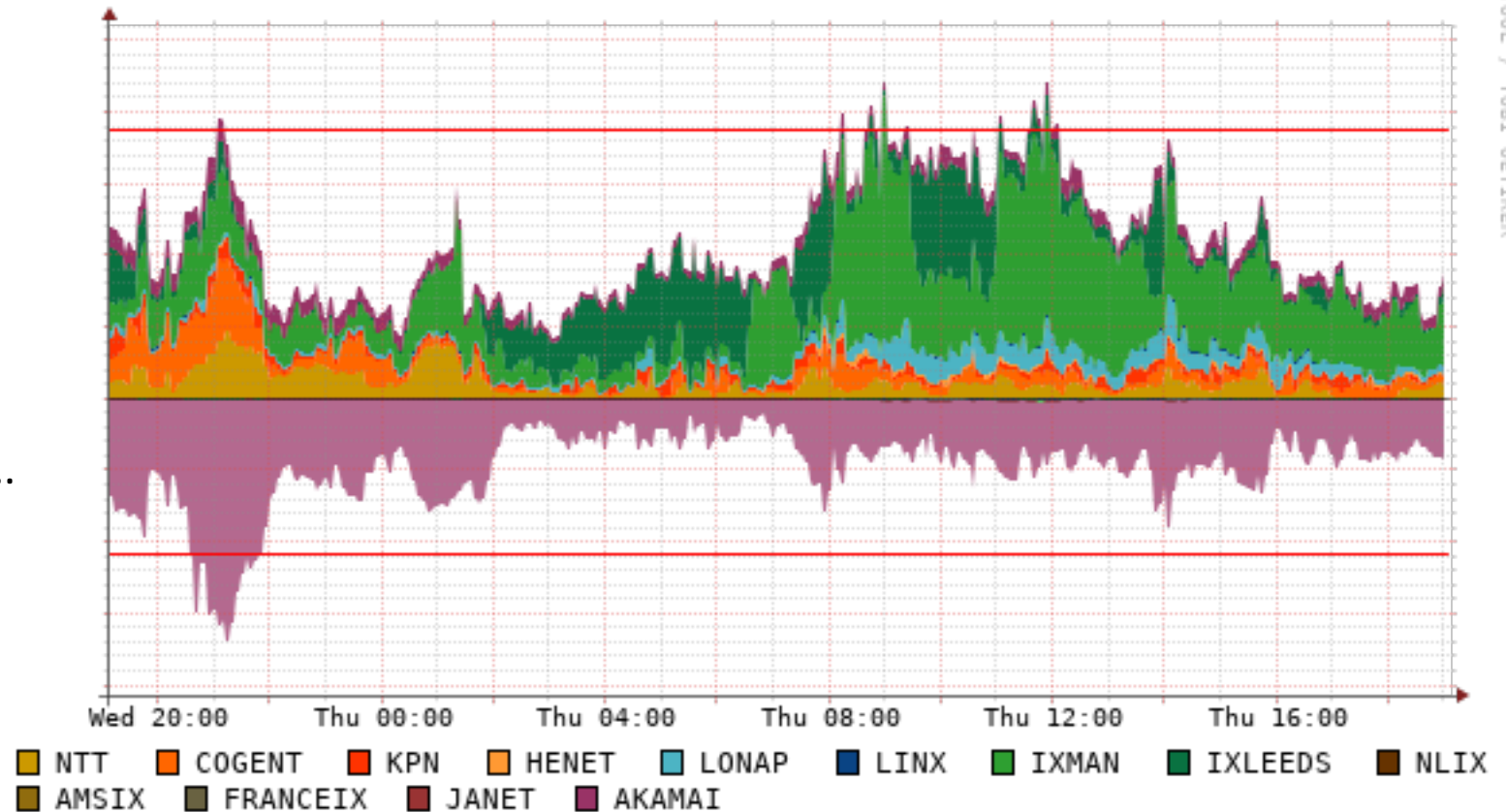
Both time the data was collected during schools holidays

Being an education ISP, it mean "quiet" time.

Added some transit providers since then but ..
  1 – IXManchester (most of the day)
  2 – Varied … It's auto-magical

IXManchester still our #1 peering location for Akamai.

AS20940 - Akamai International B.V.,US - IPV4

Legend: NTT COGENT KPN HENET LONAP LINX IXMAN IXLEEDS NLIX AMSIX FRANCEIX JANET AKAMAI

the colors are now nicer ☺

Conclusion

Installing more DNS resolvers …
improved cache locality, making for a better end-user experience

- Moved some flows from London to Manchester
  Reduced our London / Manchester core traffic

- Is only worth it if you have multiple peering points
  And eyeballs in different POPs

- If you want more traffic from regional exchanges
  Setup some regional DNS

- DNS failure can cause traffic re-routing
  Changing core link utilisation

Feel free to contact your usual Akamai contact, or their noc, should you have any questions I can not answer.
https://www.peeringdb.com/net/2