IPv6 Refresher

Tim Chown, Jisc (tim.chown@jisc.ac.uk)

STFC, RAL campus, 9-10 January 2023

Agenda

An IPv6 refresher

- **Today: 2-3pm**
  - Why deploying IPv6 is important
  - IPv6 protocol fundamentals
  - IPv6 addressing and configuration
- **Tomorrow: 2-3pm**
  - IPv6 thinking
  - Operational, deployment and security perspectives
  - Examples: GridPP IPv6 WG, some perfSONAR views
  - Measuring deployment progress

A quick Janet reminder

Jisc manages and operates Janet for the UK R&E community

The network has supported IPv6 natively for 20 years

All our external peerings are now dual stack (IPv4 and IPv6)

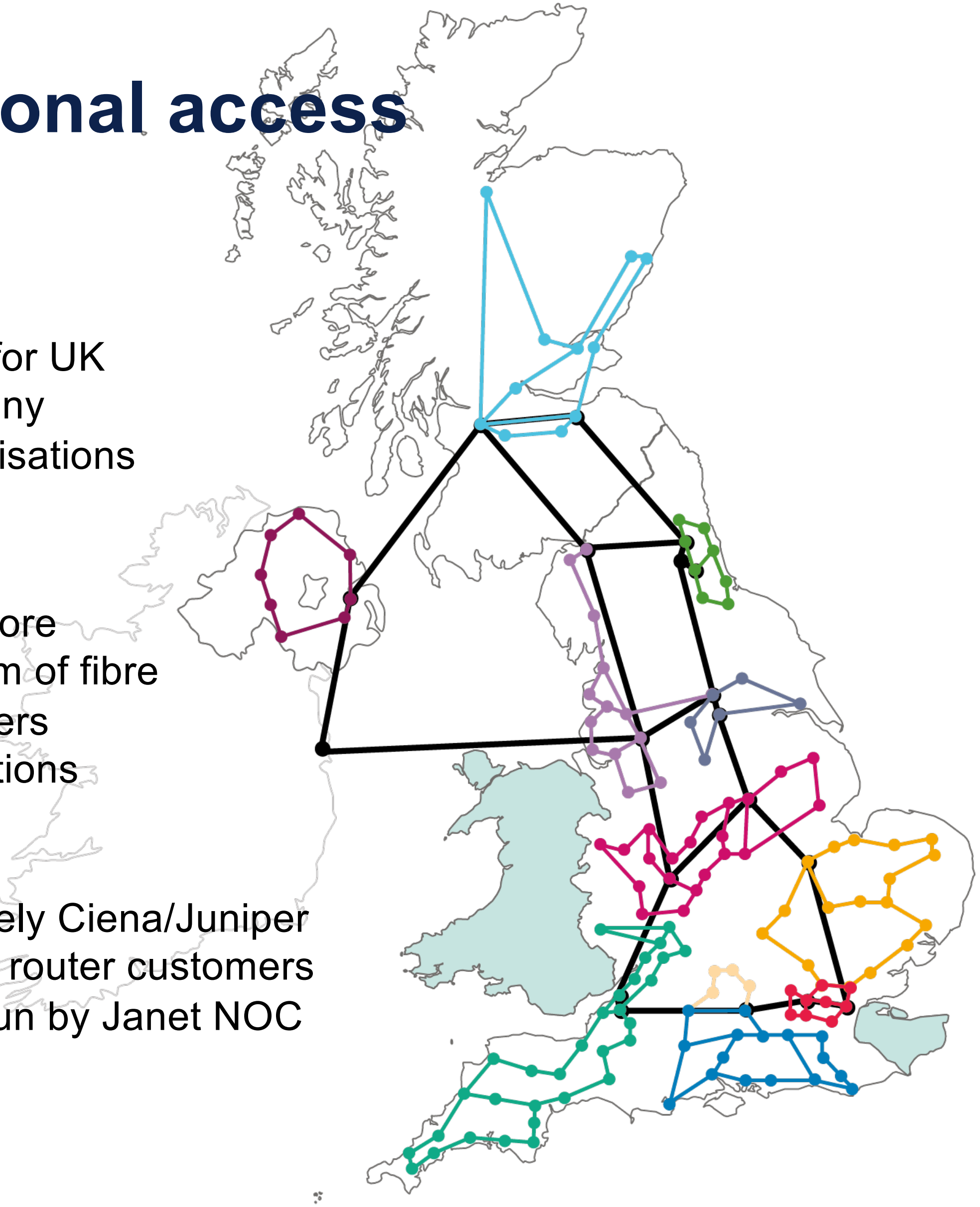
Janet backbone and regional access infrastructure

- Janet backbone
- Scotland
- North West
- Yorkshire
- Northern Ireland
- North East
- Midlands
- East
- South West
- Thames
- South
- London
- Public sector networks

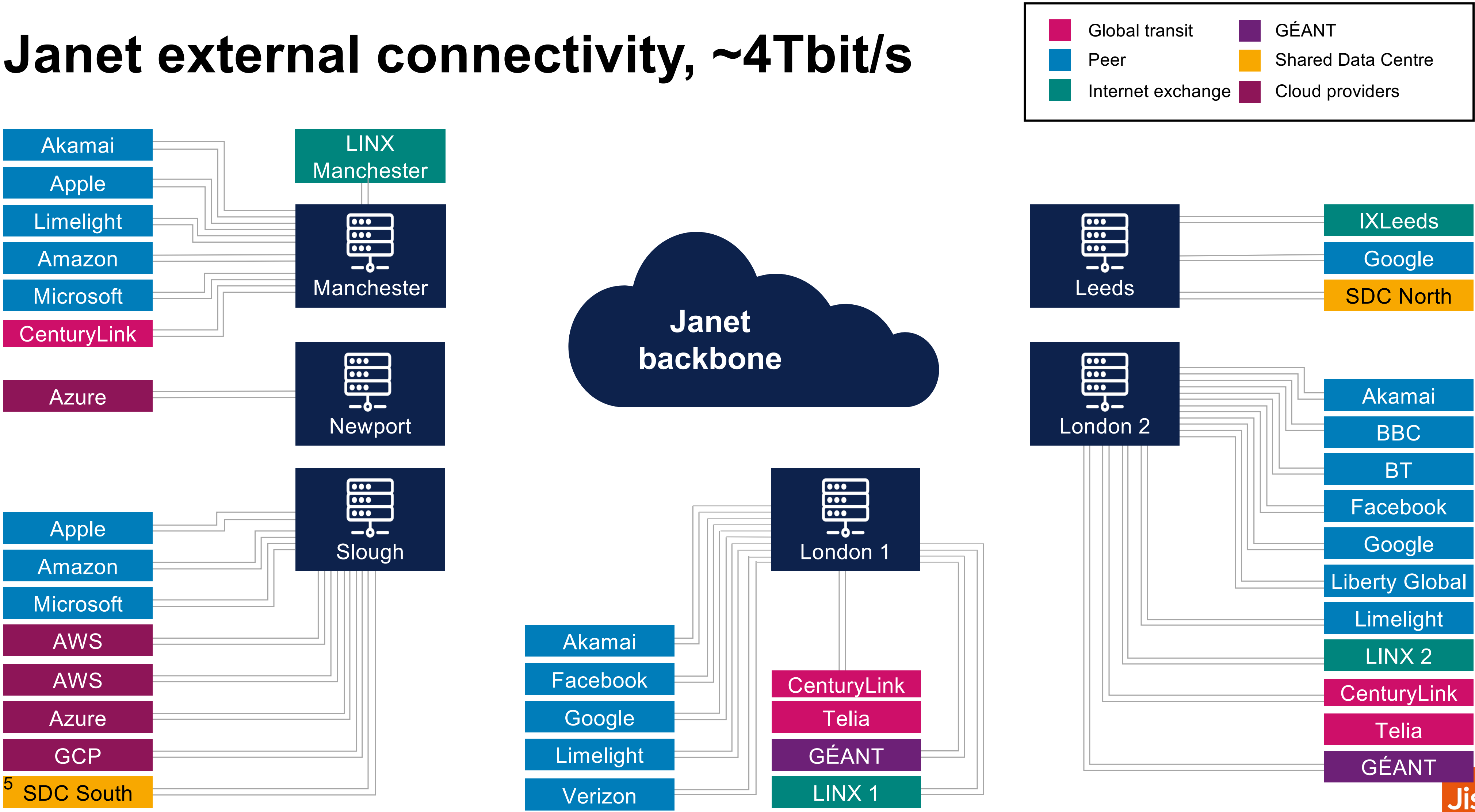
Jisc is the ISP for UK HE/FE, and many research organisations like STFC

800G in main core
Around 9,000km of fibre
~1,000 customers
~1,500 connections

Network is largely Ciena/Juniper
~430 managed router customers
~700 devices run by Janet NOC



Janet external connectivity, ~4Tbit/s



Jisc

The case for IPv6

Why deploying IPv6 is important

It's primarily about address space

- Sounds obvious, but worth emphasising
- The key difference is **128-bit addressing**
  - Enough globally unique address space to support future growth and innovation
- Ensures you can uniquely address all devices in your infrastructure
  - e.g., GridPP wants to directly address all worker nodes
- There's no new, unused IPv4 available; Jisc has no significant reserves
  - For more IPv4 addresses you need to go to the IPv4 broker market
  - Currently $40-$50 per address, see https://ipv4.global/
  - Or about $3M for a /16 of IPv4

Important recent developments

Many positive areas

- Reached point where over 40% of user traffic in the UK is IPv6
  - Linear growth implies IPv6 will be over 50% by the end of 2024

- US government issued new mandate to deploy IPv6-only by 2025
  - OMB-21-07 - important because vendors want the US government business
  - https://www.whitehouse.gov/wp-content/uploads/2020/11/M-21-07.pdf

- Much improved support in cloud and container platforms, e.g.:
  - Kubernetes support – see https://kubernetes.io/docs/concepts/services-networking/dual-stack/ (v1.20 onwards)
  - AWS – see https://www.ipv6.org.uk/2022/10/13/ipv6-council-annual-meeting-2022/

And other reasons still apply

Including…

- Supporting teaching and research
- Ensuring robust access to your public-facing services for IPv6-only client devices
- Avoiding a rushed deployment; building technical credit not debt
- Removing NAT and private addressing complexity from network operations and management
- Minimising dependency on an ever-more fragile IPv4 network (witness CGN, etc)
- Security in an ‘IPv4 only’ network; IPv6 is supported on common platforms and on by default
- Enabling innovation at the edge
- Scalability for campuses of the future; IoT is increasingly using IPv6 (e.g., Matter)
- Being ready for IPv6-only applications and communities (the WLCG direction of travel)
- .. Are there others? Presumably, your business case is already made?
- As Dave Kelsey pointed out, new capabilities in IPv6 such as enhanced packet marking

Why not deploy?

Do you have specific concerns?

- We can discuss later, or feel free to ask now ☺
- It's quite useful to look at other concerns, whether real or FUD, at https://ipv6bingo.com/

IPv6 Excuse Bingo

It's too complicated	Hex is hard	There's no certification track	Our vendor doesn't support it
Android doesn't support DHCPv6	IPv6 is a security risk	No one else has deployed it	Our Dynamic DNS doesn't support it
Can't we just buy more IPv4 addresses?	Larger headers are less efficient	IPv6 isn't an Internet Standard yet	IPv6 is just a fad
Those stupid Privacy Extension addresses keep changing	We forgot to include IPv6 in our last RFP	IPv6 just isn't a priority	We don't need that many addresses

Made with excuses from ipv6excuses.com
Suggest a new excuse: Tweet to @ipv6excuses

Jisc

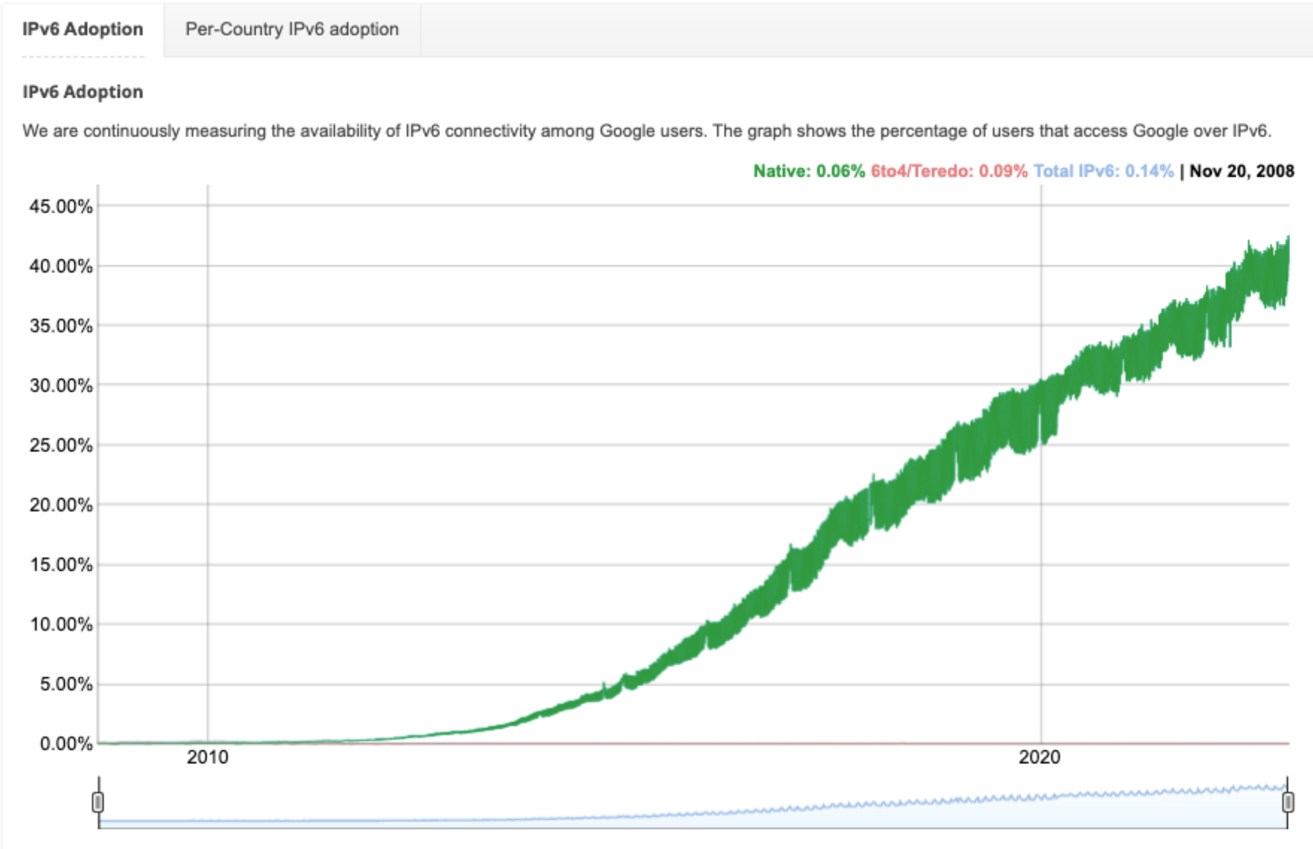
Deployment status and measurement

How much of total traffic is IPv6?

- There are various measurements out there for IPv6 deployment, including summaries at
  - http://www.worldipv6launch.org/measurements/
  - https://labs.ripe.net/Members/mirjam/content-ipv6-measurement-compilation
- Overall, Internet user traffic is around 40% IPv6, with the UK similar
- IPv6 adoption varies by sector: residential, mobile, enterprise
- Janet IPv6 traffic sits at under 10%
- R&E is well behind commercial adoption
  - Probably due to no perceived **NEED** for IPv6, today

Jisc

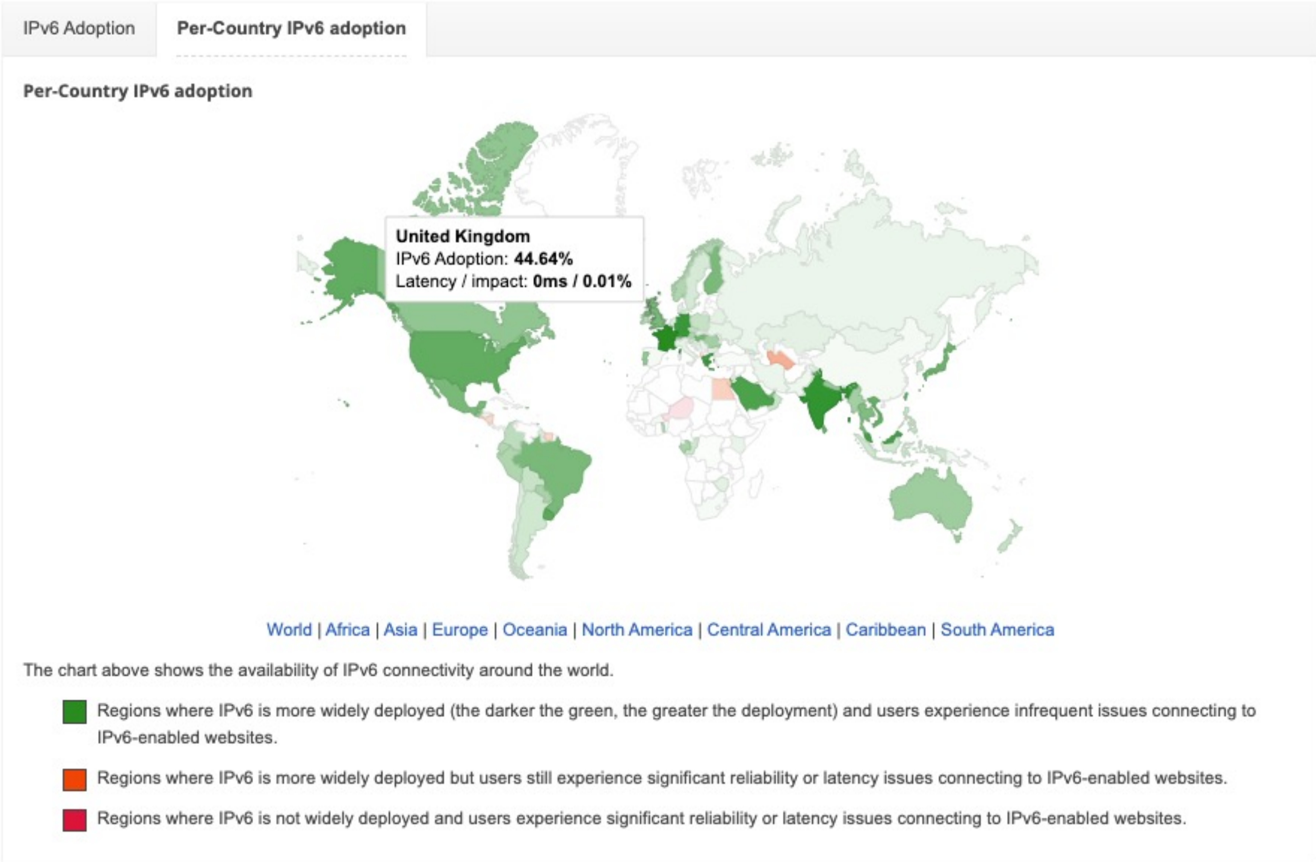
World adoption, as seen by Google



https://www.google.com/intl/en/ipv6/statistics.html

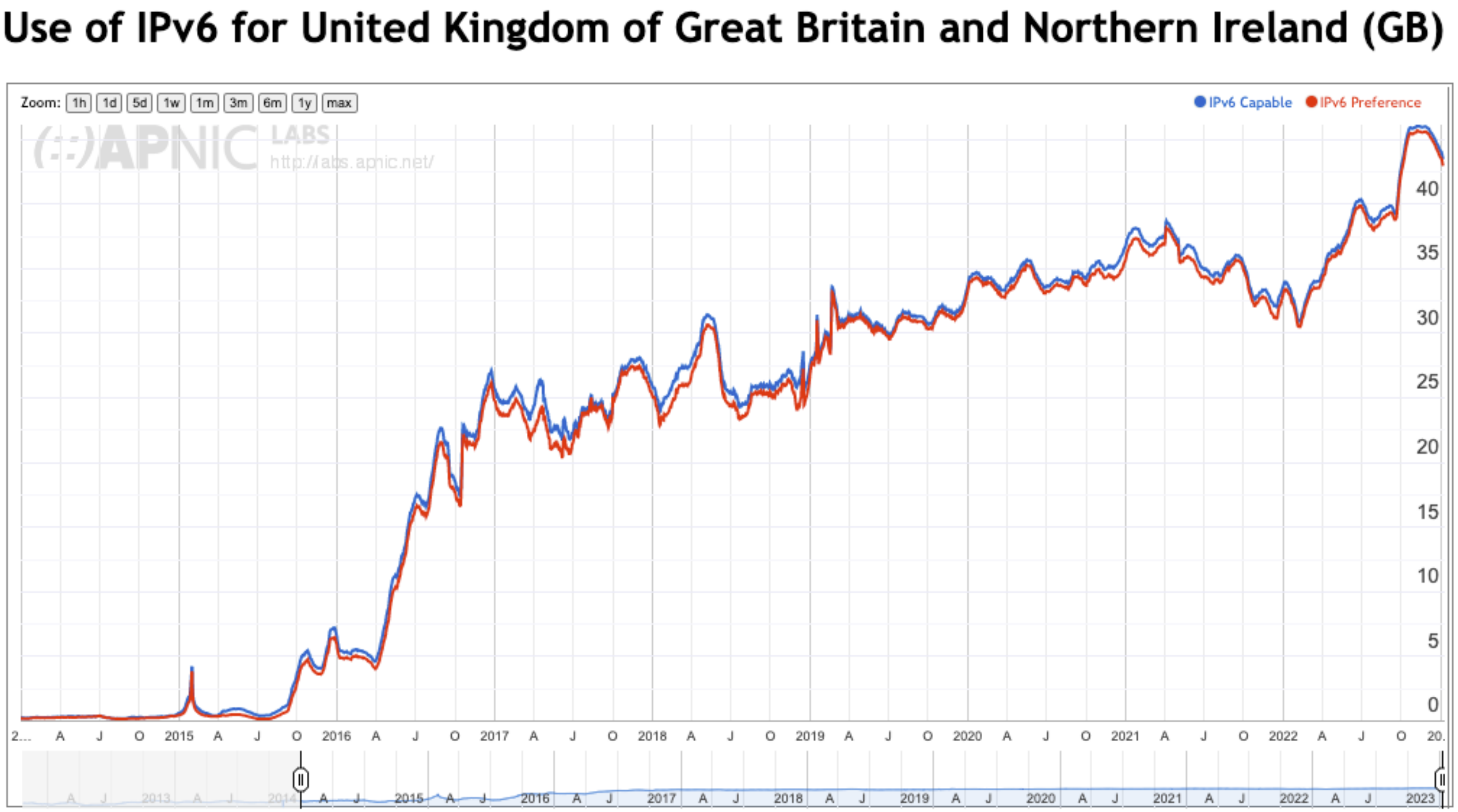
Jisc

UK, as seen by Google



https://www.google.com/intl/en/ipv6/statistics.html#tab=per-country-ipv6-adoption

UK, as seen by APNIC stats (via Google ads)



https://stats.labs.apnic.net/ipv6/GB

IPv6 adoption at Janet sites?

Not a great story, yet

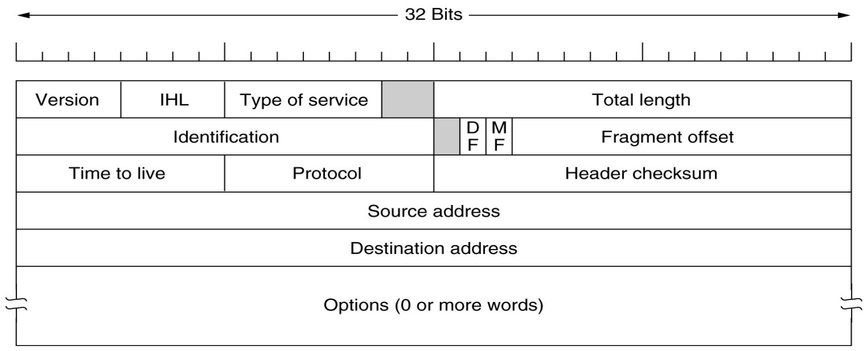
- Around 160 universities in the UK; just over 100 have an IPv6 assignment
  - But only 28 have IPv6 address space that has been "seen"
- Of the Times Top 20 universities, all have an assignment, 14 have traffic seen, 10 have IPv6 DNS, 5 have IPv6-enabled their web presence but only 3 mention IPv6 on their CS syllabus (source: Graeme Bragg @ Networkshop 50, 2022)
- On our Janet peering with GÉANT (international R&E traffic):
  - 17% of exported traffic is IPv6, and 8% of imported traffic is IPv6
- On general commodity external traffic:
  - 2.5% of exported traffic is IPv6 and 1.6% of imported traffic
- The APNIC stats show Janet at around "**3.5% capable**" for IPv6 user traffic
  - We need to change this!

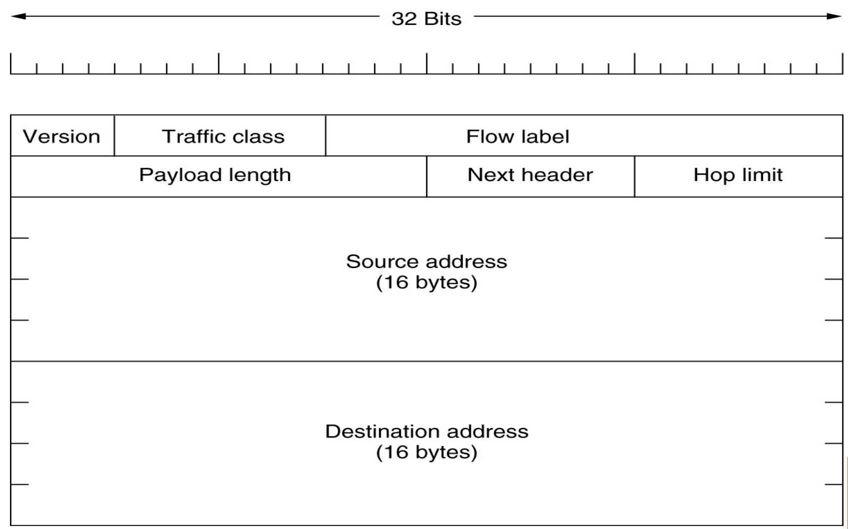
IPv6 fundamentals

IPv6 packet format

A streamlined header, with optional Extension Headers (EHs)

- Fixed length main header
- Chain of **optional** EHs, e.g.:
  - Fragmentation header
  - IPsec headers
- IPv4 and IPv6 **NOT** compatible
  - No direct IPv6-only to IPv4-only
  - Need some form of translation to talk
  - But can run both, as dual-stack
- **EHs** can be problematic (RFC 7872)
  - Most traffic has no EH present





IPv6 Addressing

What changes with IPv6?

- IPv6 addresses are 128 bits
- The notation used to write addresses is different to IPv4
  - e.g., **2001:db8:e380:d0:920a:2380:b230:1106**
- IPv6 also has different scopes of addresses
  - Link-local (under fe80::/10), Unique Local Addresses (ULAs, under fc00::/7), and globally unique addresses (GUAs)
  - **Multi-addressing is the norm** – more later
- IPv6 address allocation and assignment policies are determined by the Regional Internet Registries (RIRs), and their members, which for us is RIPE
  - See https://www.ripe.net/publications/docs/ripe-738 (updated many times)

Writing IPv6 addresses

And abbreviating them

- As we just saw, IPv6 addresses are written as eight sets of four hexadecimal characters, e.g.
  - **2001:0db8:0000:0000:baad:cafe:1234:5678**
- To abbreviate, you can omit any leading zeros in any set of four
  - **2001:db8:0:0:baad:cafe:1234:5678**
- And you can replace **one** series of :0: fields with ::
  - The above example can be 2001:db8::baad:cafe:1234:5678
  - … but why only one?
    - Consider **2001:db8::baad::1234:5678**

Jisc as your ISP

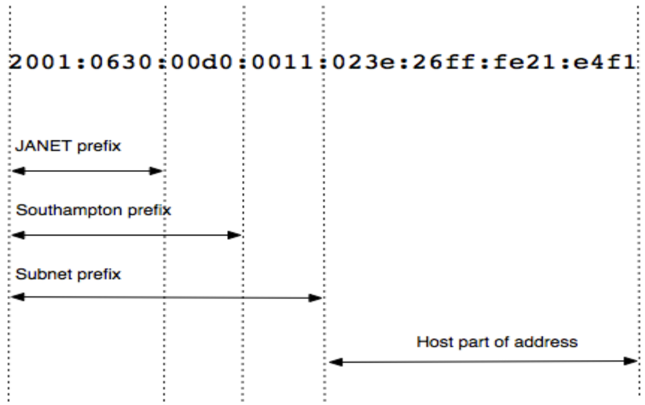
Connecting to Janet with IPv6

- The Janet backbone is dual-stack throughput
- IPv6 is available natively as part of the Janet IP Connection service
  - Presumably, you have already had the Jisc Service Desk turn IPv6 on
  - https://www.jisc.ac.uk/janet-ip-connection
- Jisc as an ISP is also a Local Internet Registry (LIR)
- As an LIR, Jisc obtained the prefix 2001:630::/32 from the RIPE NCC in 1999 (was a /35 back then, now the default is /32 with a /29 reserved)
  - Jisc can provide you an IPv6 prefix for your organisation on request
  - **Default** prefix assignment to a Janet-connected site is a **/48**

Address assignments

From LIRs to end sites

- A typical prefix breakdown for a university site (or RAL) might be:



- NB: the ISP and site prefix lengths here are just defaults
- A Janet site may thus have up to 2^16 host subnets
- We'll see why host subnets are 64 bits soon

What if I want more than a /48?

Getting bigger prefixes

- You can work with Jisc to document any rationale or need for a larger address space from our allocation, and Jisc can provide advice on that (we have to justify the use to the RIPE NCC)
- Or you can apply for LIR status yourself, as a small number of universities have done (QMUL for example) and get a /32 that way. The cost is low thousands of Euros each year.
- There has been some discussion on this between STFC and Jisc

- **Jisc will route your traffic whichever approach you choose**

Jisc

Configuring addresses

New: stateless autoconfiguration (RFC 4862)

- IPv6 protocols are defined in the Internet Engineering Task Force (IETF)
  - Base IPv6 protocol – RFC 8200 (an Internet Standard)
- There is the choice to use the “comfy slippers” of DHCPv6
- But IPv6 introduces **stateless address autoconfiguration (SLAAC)**
  - SLAAC implicitly means all host subnets are /64 in size
  - It’s widely used in the campuses who have already deployed
  - Only a handful are using DHCPv6, so far
- Or you can configure addresses manually or through automation
- What do you currently do for IPv4?

Aside: IPv6 Neighbour Discovery

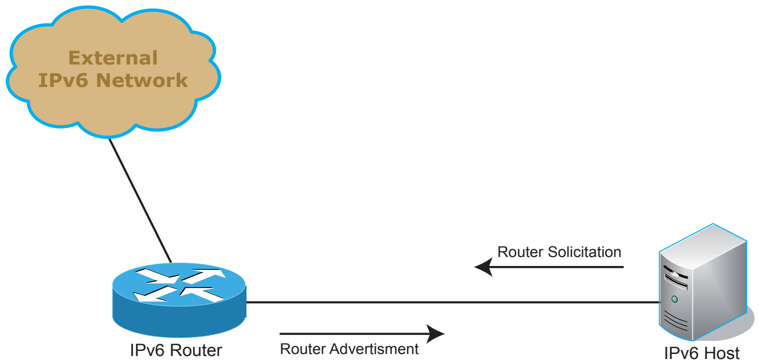
A suite of five link-local messages between hosts and routers

- Uses multicast (under ff00::/8) ICMPv6 – see RFC 4861
  - The messages are layer 3 rather than layer 2 (like ARP broadcasts)
  - IPv6 also does not have IP subnet broadcast addresses
- **Neighbour Solicitation (NS)** and **Neighbour Advertisement (NA)**
  - Provides the functionality of ARP for IPv4; find MAC address for an IP
  - Includes Duplicate Address Detection (DAD) for tentative addresses
- **Router Solicitation (RS)** and **Router Advertisement (RA)**
  - RAs pass hosts information about routers on their link
  - RAs can include an ‘M’ bit to say ‘use DHCPv6’
- **Router Redirect**

SLAAC

IPv6 address autoconfiguration

- Router periodically sends an RA, or hosts can solicit one with an RS
  - Information in the RA has an associated preferred / valid lifetime
- Host configures its address using the 64-bit network prefix in the RA
  - The host part of the address (last 64 bits) should be set by RFC 7217 to be a unique ‘random’ value per prefix attached to
- RAs also allow hosts in the same subnet to configure default router, DNS (with RFC 8106), MTU, and on-link prefixes
- SLAAC means IPv6 has 64-bit host subnets
  - A de-facto default, even with DHCPv6
  - See also RFC 7421 on “Why /64?”



IPv6 temporary / privacy addresses

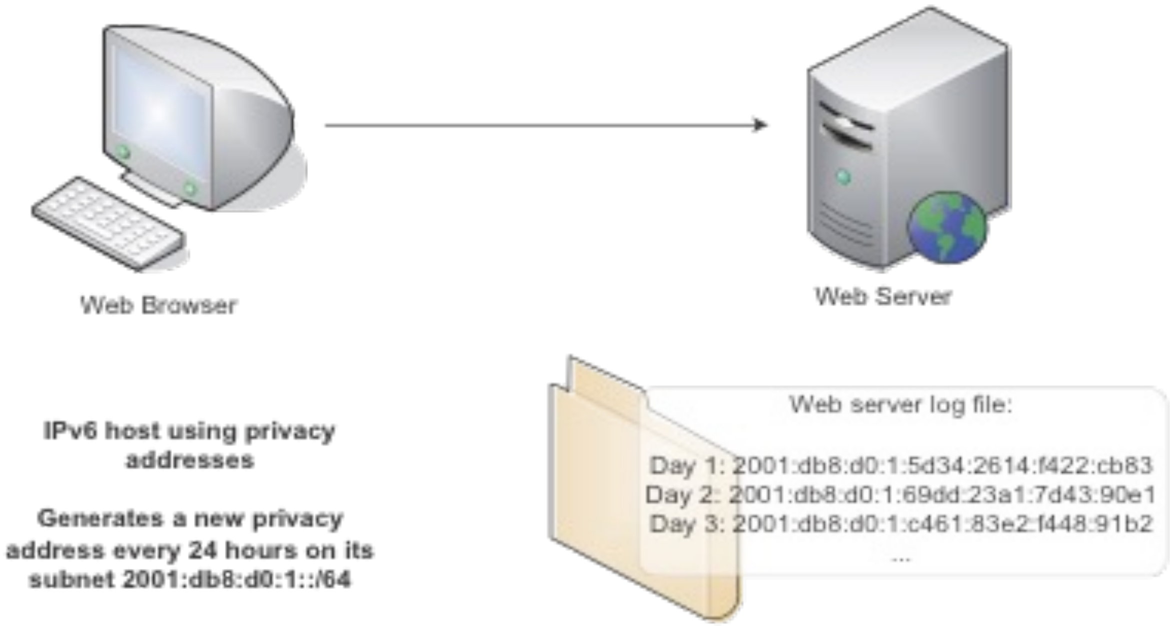
Designed to improve user privacy

- A host can additionally configure a temporary address (RFC 8981, was 4941)
  - Used as source address when sending
- Originally designed for cases when the host part of the address embedded the device's 48-bit MAC address (with 16 bits of 'fffe' padding in the middle)
  - Facilitated tracking when connecting to different networks
  - Arguably less important since the introduction of RFC 7217
- Privacy addressing is generally turned on by default
  - May also generate new privacy addresses over time for immobile devices
  - Typically want to disable this for servers, and on managed systems
  - No way to disable for BYOD
- Good for users, but adds challenges for management

The privacy address challenge

Are those addresses different hosts, or one host with many addresses?

•Consider a client and a web server



Example 1

Staff Mac laptop

```
$ifconfig en7
en7: flags=8863<UP,BROADCAST,SMART,RUNNING,SIMPLEX,MULTICAST> mtu 1500
   options=6467<RXCSUM,TXCSUM,VLAN_MTU,TSO4,TSO6,CHANNEL_IO,PARTIAL_CSUM,ZEROINVERT_CSUM>
   ether 9c:eb:e8:c1:49:13
   inet6 fe80::105a:2384:34ee:23f9%en7 prefixlen 64 secured scopeid 0xe
   inet6 2001:db8:d510:1101:184a:98b6:aaf4:3e52 prefixlen 64 autoconf secured
   inet6 2001:db8:d510:1101:ac8c:fe54:7923:4fc9 prefixlen 64 autoconf temporary
   inet 192.168.0.12 netmask 0xffffff00 broadcast 192.168.0.255
   nd6 options=201<PERFORMNUD,DAD>
   media: autoselect (1000baseT <full-duplex>)
   status: active
```

Example 2

CentOS server

- Config example for *ps-slough-10g.ja.net*
- Manually configured IP, no temporary addresses

```
$ ifconfig em1
em1: flags=4163<UP,BROADCAST,RUNNING,MULTICAST>  mtu 1500
        inet 194.81.18.229  netmask 255.255.255.254  broadcast 255.255.255.255
        inet6 2001:630:3c:f803::a  prefixlen 126  scopeid 0x0<global>
        inet6 fe80::d294:66ff:fe4c:37bf  prefixlen 64  scopeid 0x20<link>
        ether d0:94:66:4c:37:bf  txqueuelen 10000  (Ethernet)
```

DHCPv6

The more familiar configuration method

- Stateful DHCPv6
  - For full device configuration, as per RFC 8415 (was 3315)
  - Various options for DHCPv6 Unique Identifiers (DUIDs), including RFC 6939 to key addresses to MAC addresses and use a relay
- Stateless DHCPv6
  - For ‘other configuration information’ – NTP, DNS, …
- Still need RAs for default router and on-link prefixes
- No DHCPv6 for Android – an issue for IPv6-only and DHCPv6
  - https://issuetracker.google.com/issues/36949085

Address planning

How you use your IPv6 assignment

- A similar process to IPv4
- Subnetting will typically follow geographic or administrative boundaries
- Lots of ‘clever’ things you can do because the bits are there
  - Nice guide written by SURFnet over 10 years ago
  - https://www.ripe.net/support/training/material/IPv6-for-LIRs-Training-Course/Preparing-an-IPv6-Addressing-Plan.pdf
- As mentioned before, all IPv6 host subnets are /64 due to SLAAC, whether you use SLAAC or DHCPv6 (have you decided?)
- You will probably have congruent IPv4 and IPv6 prefixes
  - The rationale for subnetting is the same either way
  - Your IPv4 subnets may vary in size over time (/24, /27, /22,…)

Address planning (2)

Other considerations

- How big do you wish to make your layer 2 subnets?
  - If you go very big (esp. a flat layer 2 campus!) then there will be a lot of 'chatter' traffic from IPv4 ARP, IPv6 ND, etc
  - At the other extreme, RFC 8273 describes using a prefix per host for IPv6
- Consider whether all devices in a subnet will be configured the same way
  - RAs provide the same information to all hosts
- Will you use ULAs?
  - It may be tempting for systems you know do not communicate off site
  - But I don't know of any Janet sites doing so (not that ULAs would be seen…)
- Use /127 for point-to-point as per RFC 6164 (but you can reserve a /64)

Routing protocols – a quick note

There are new IPv6-compatible versions

- These are well-established
  - Multiprotocol BGP (RFC 2545)
  - IS-IS (RFC 5308)
  - OSPFv3 (RFC 5340)
  - RIPng (RFC 2080)
- Most campuses/enterprises probably run OSPFv2 or IS-IS
  - Can run OSPFv2 alongside OSPFv3
  - Check support for multi-AF (address family) OSPFv3 if looking to use OSPFv3 for both protocols

Questions?

Anything you'd like to talk about so far?

- Does what we've covered match what you might have read before?
- Anything need a little more discussion?
- Have you got an address plan for RAL?

- We'll move on to the more operational topics in part 2
  - "IPv6 thinking"
  - Security, transition tools, dual-stack considerations, etc

IPv6 thinking and operational considerations

How do you need to adjust your thinking?

What makes IPv6 different?

- The larger addresses and their format
  - Need to consider all the places literal addresses appear – see RFC 5952
- Multi-addressing is normal
  - IPv6 LL, IPv6 GUA, IPv4 (private or global), IPv6 privacy addresses
  - Need to consider address selection – defined in RFC 6724 (prefer IPv6, prefer matching scope, longest prefix match, …)
- Fragmentation is only done by end hosts, not along the path
  - Thus need PMTUD to work; RFC 4890 advises on ICMPv6 filtering
- The minimum MTU for IPv6 is bigger, at 1280 bytes

Jisc

A summary

	IPv4	IPv6
Address length	32 bits	128 bits
Prefix length	Varies, typically /24	Always /64 in host subnets
Address configuration	DHCPv4	Stateless Autoconfiguration DHCPv6
Addresses used	Private **or** Global	Link-local **and** Global
Address resolution	ARP	Neighbour Solicitation / Advertisement
Host Path MTU Discovery	Optional	Required
Fragmentation	By hosts or routers	Only by hosts
Private addressing	RFC 1918	Unique Local Addresses (ULAs) (not designed for use with NAT)

Operational considerations

Examples

- What level of host/user accountability do you need?
  - Will you need new tools to handle the impact of multi-addressing?
- What about NAT?
  - Do you use it for IPv4? If so, for what reason?
  - It’s possible in IPv6 (at least as NPTv6) but not recommended
- Do you want to run 9000MTU?
  - Janet supports it, and there is a clear performance advantage
  - If so, it’s needed end-to-end, and you’ll need to ensure PMTUD works

Dual-stack considerations (2)

Running both protocols

- The common deployment model for campuses is dual stack
- Can then use IPv4 to IPv4-only systems, and IPv6 to IPv6-only
- Application can pick either protocol for a dual stack destination
- There's "Happy Eyeballs" for browsers (RFC 8305, was 6555)
  - Try IPv4 and IPv6 in parallel, pick the one that connects first
- Other applications should failover quickly if one protocol fails
- Application protocol selection has proven a challenge for GridPP
  - i.e. finding out **why IPv4 is used when IPv6 is available**
  - e.g., a Java preference was not set to IPv6 for dCache

What about DNS?

It's the same, but different

- Delegation happens as per IPv4 for forward and reverse DNS
  - Use AAAA records for IPv6 where you use A records for IPv4
  - Reverse DNS for IPv6 is nibble-based under ip6.arpa
    - So would be under 8.b.d.0.1.0.0.2.ip6.arpa for 2001:db8::/32
- All common IPAM systems should support IPv6
  - Infoblox is quite a popular solution used in R&E sites
- Don't add a host's AAAA entry to DNS until all services on it support IPv6 else there will be nothing listening for the connection

- Jisc offers a primary and secondary DNS service, and a network resolver service with RPZ built-in for malware protection

Jisc

What about transition tools

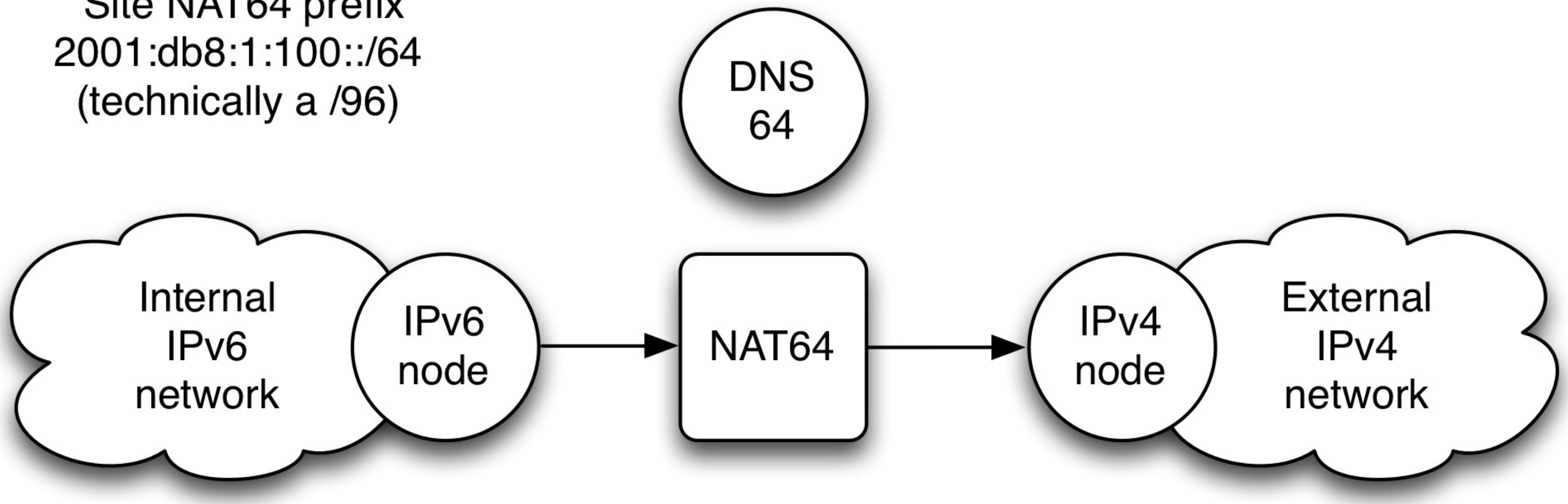
The focus of ‘transition’ has shifted

- Originally multiple tunnelling methods were defined to support connecting between IPv6 ‘islands’, for hosts and routers
- With native IPv6 now pervasive on the R&E backbones, this need has largely gone away
  - Tunnel brokers can still be handy, e.g., *tunnelbroker.net*
  - A dual-stack VPN is a form of tunnel broker - do you run one?
- With a growing interest in running IPv6-only, the focus has shifted to translation, specifically for IPv6-only to IPv4-only content
  - The usual approach is NAT64/DNS64 and 464XLAT (for literals)
  - Also some use of MAP-T (RFC 7599)

NAT64 / DNS64 example

The ‘trick’ is used of DNS64, which returns a ‘fake’ IPv6 address formed by appending the IPv4 address to the site’s chosen NAT64 prefix, and the IPv6-only client sends to that address, with the NAT64 box doing the address and protocol conversion (out and back)

Site NAT64 prefix
2001:db8:1:100::/64
(technically a /96)



Src: IPv6 address
Dst: <IPv6-prefix>:<IPv4-address>

2001:db8:1:100::984e:4001

152.78.64.1

Security

Need to implement your policy for both IP versions

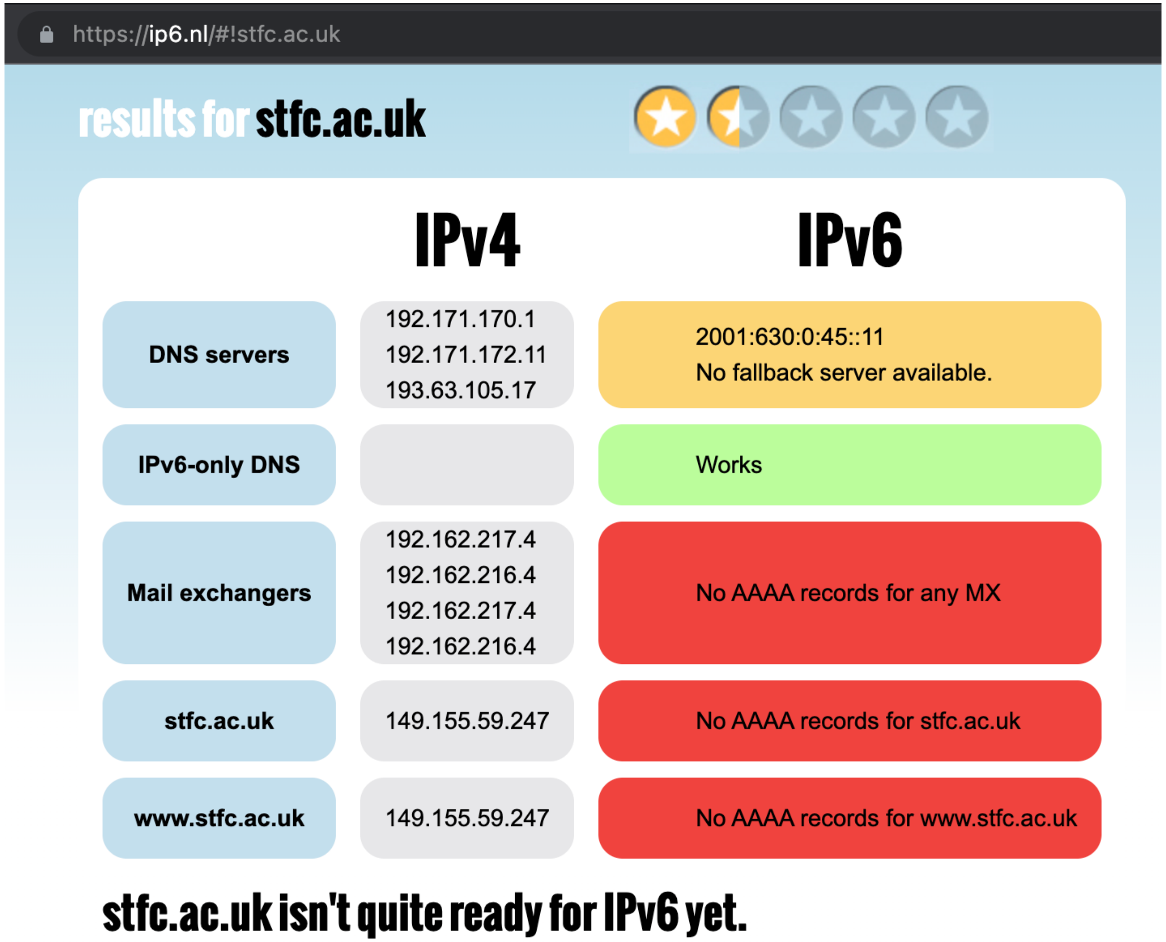
- Leaving IPv6 ‘open’ is not a great idea
- Dual-stack means you need to support / manage both protocols
  - Look to use tools that manage objects consistently, e.g. by using dual-stack firewall objects not separate IPv4 and IPv6 objects
- Focus where possible on feature equivalence
  - RA Guard is as important as DHCPv4 Guard, ARP vs DAD spoofing
- But recognise new threats
  - e.g., abuse of IPv6 extension headers
- And differences
  - Classic IP scanning is less feasible (see RFC7707)
- David Holder @ UKNOF - https://www.youtube.com/watch?v=SbgbExbu1kk

(Unix) Troubleshooting

Not that different, for example:

- Tools like ping and traceroute work the same way
- Netstat can provide information for either protocol
- Wireshark and tcpdump work as expected
  - Useful for looking at ND traffic for example, like RAs
- You can check an ND cache like an ARP cache
- Bad ICMPv6 filtering might impact connectivity / PMTUD
- There are some other handy web-based tools
  - e.g., https://test-ipv6.com/, https://ipv6-test.com/, https://ip6.nl/

The site checker…



Considerations for deploying IPv6

Most guidance distils down to a phased approach

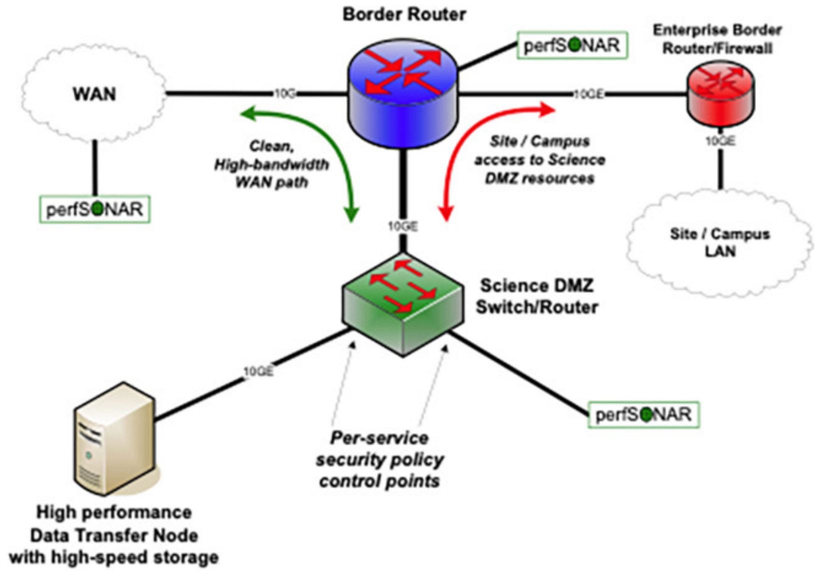
- Planning
  - Acquiring address space, forming an address plan, training, platform review, security policies
  - Ensuring IPv6 requirements set in procurements (even if not turning IPv6 on yet)
  - Choose how you will do host configuration
- Piloting / testing
  - Form a test plan, deploy a testbed, perhaps connect testbed directly to your edge router
- Production
  - Choose a deployment target (public services, WiFi, etc.), confirm dual-stack approach
  - Build support from the core first: edge connectivity, firewalling, backbone, DNS
  - Ensure your monitoring and management tools work with IPv6
  - Then turn on IPv6 in the selected edge network (router advertisements, etc.)
- Ongoing build-out and enhancement
  - Business as usual

Deployment – no Big Bang needed

You don’t need to do everything on day one

- R&E deployments to date have initially focused in one of the following areas:
  - Public facing services, particularly web, but also email and DNS
  - Campus WiFi (eduroam)
  - Computer science department, labs, etc. (teaching and research)
  - Computing service
  - Science DMZ (e.g., GridPP site)
- You still need to get IPv6 running in your core and edge network, but a focused initial deployment can be a more manageable project
- You can also learn from an initial smaller deployment to enhance a future wider deployment

Aside: Science DMZ



Handling science and business traffic

- ESnet documented “**Science DMZ**” principles ~10 years ago
  - https://fasterdata.es.net/science-dmz/
- Key design elements:
  - Local network architecture to differentiate large science flows
  - Well-tuned data transfer nodes (DTNs)
  - Performant data transfer tools (**FTS**, Globus, etc)
  - Persistent monitoring of network characteristics (**perfSONAR**)
- Avoid the large flows traversing the main campus firewall
  - Apply security policy efficiently, save costs on the stateful DPI firewall capacity
- UK GridPP has evolved the same principles from experience

Jisc

Procurement

Be sure to procure IPv6-capable products

- As a general guide, tenders should require feature parity, or at least a statement / roadmap from a vendor on IPv6 capability
  - Even if you don't plan to turn IPv6 on yet on the product

- There is RIPE guidance available for different types of equipment
  - RIPE772: https://www.ripe.net/publications/docs/ripe-772

- There's also the IPv6 Forum's IPv6 Ready Logo programme
  - https://www.ipv6ready.org/
  - (but not all vendors have put equipment forward for this)

When can I remove IPv4?

Simplifying your IPv6 deployment longer term

- The prudent initial deployment approach is currently dual-stack
- But the question soon becomes “where can I remove IPv4?”
  - Internal management network? (Facebook is 100% IPv6 internally)
  - On your site WiFi?
  - For research communities? The WLCG is heading this way
- You will likely need to deploy tools that support IPv6-only devices accessing IPv4-only content, such as NAT64/DNS64/464XLAT
  - But will all your applications work when using such tools?
  - The more you use IPv6 for applications, the less you need translation

Some IPv6 examples

IPv6 at Jisc

We are having a renewed push!

- Janet has been dual-stack for around 20 years
- Most of our network and security-related services support IPv6
  - NTP, DNS, eduroam peerings, etc
- We are having a renewed push internally to ensure IPv6 is supported in our broader service portfolio
  - Setting IPv6 in procurement via our ITQ template
  - Including IPv6 in our PLM service transition gate
  - Together in principle all new services will support IPv6
- Is this an approach you might use or are using for STFC?

GridPP IPv6

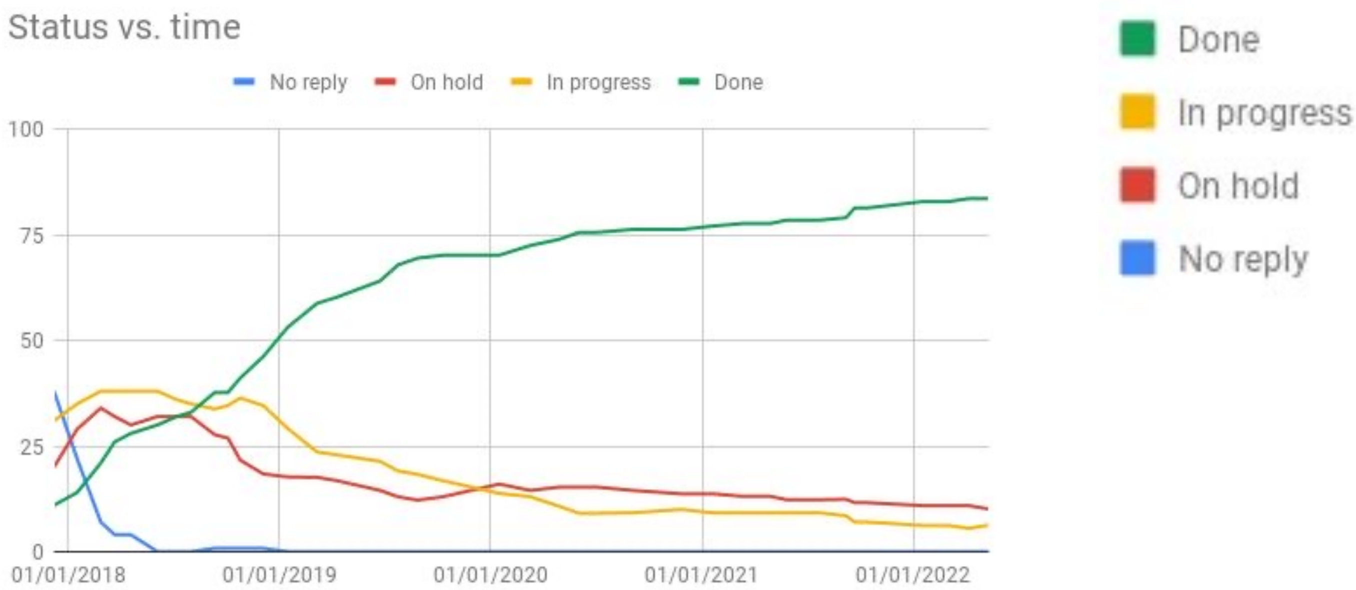
Probably the most significant IPv6 use case on Janet today

- Several years of work by the HEPiX IPv6 WG
  - https://hepix-ipv6.web.cern.ch/
- As a result, IPv6 is enabled at 15 of 19 UK GridPP sites
  - https://twiki.cern.ch/twiki/bin/view/LCG/WlcgIpv6
  - With two outliers awaiting new network equipment
- Great example of 100G of CERN traffic over IPv6 to Imperial
  - https://shapingthefutureofjanet.jiscinvolve.org/wp/uncategorized/100gbps-of-cern-data-over-ipv6-on-the-janet-network/
- QMUL has just gone live at 100G
  - Netsight showed 95Gbps+, but not yet checked the IPv4/IPv6 ratio

Some WLCG/GridPP IPv6 adoption slides from Dave Kelsey’s Networkshop talk in June 2022

Deployment of IPv6 on WLCG Tier-2s

Versus date



STATUS now	% of total
Done	84%
In progress	6%
On hold	10%
No reply	0%

IPv6 deployment – WLCG Tier2s

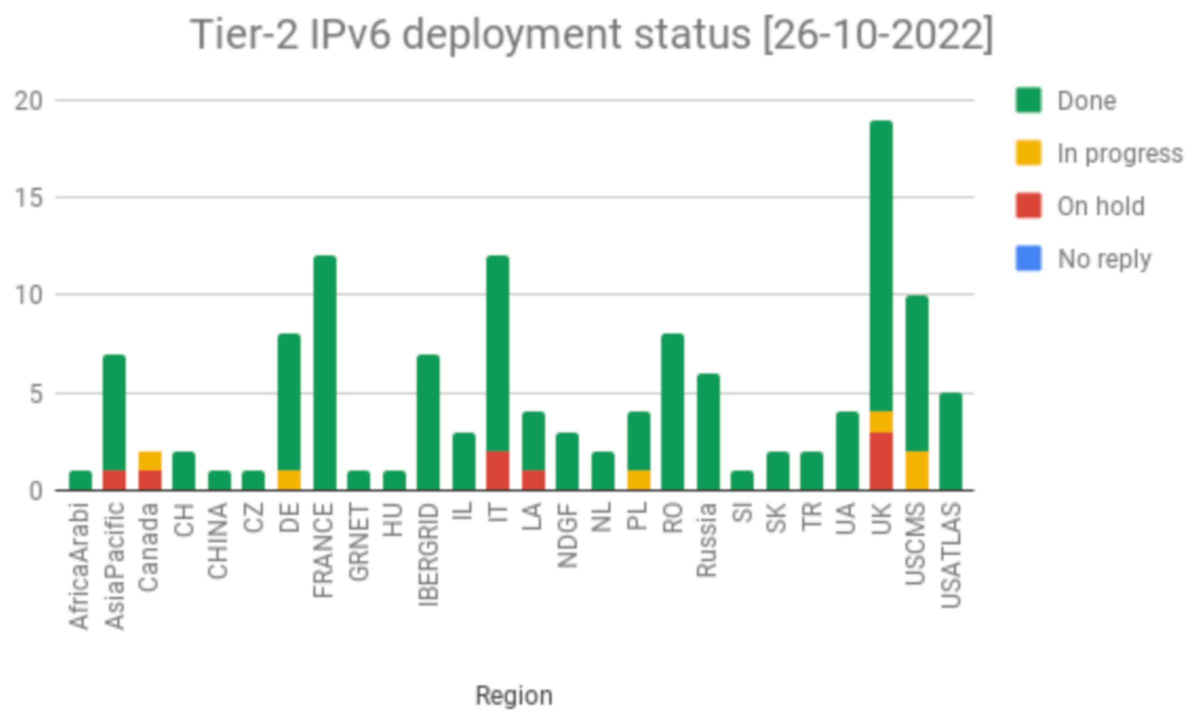
By country/region

UK GridPP Universities:

Done: Bristol, Brunel, Cambridge, Durham, Edinburgh, Imperial*2, Lancaster, Manchester, Oxford, QMUL, RAL, Sheffield, Sussex, UCL

In Progress: Glasgow

On Hold: Birmingham, Liverpool, RHUL

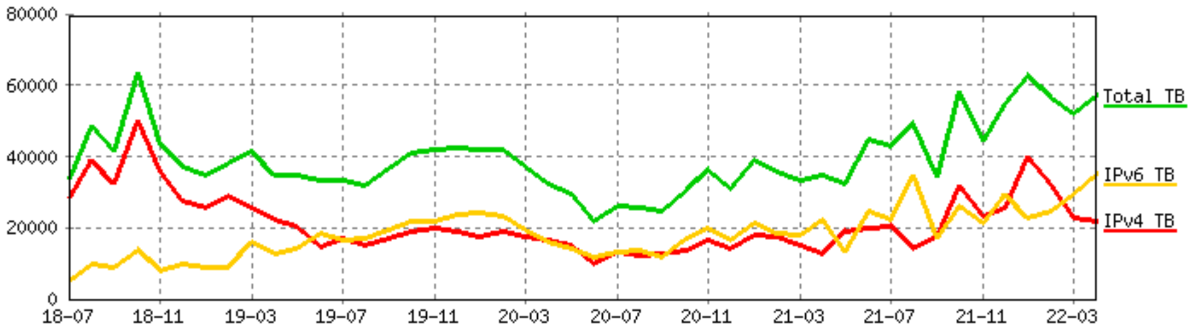


Reasons for delay: either awaiting University network upgrades or awaiting IPv6 DNS

IPv6 traffic on LHCOPN/LHCONE at CERN

LHCOPN and LHCONE IPv4 and IPv6 traffic volumes seen at CERN Tier0

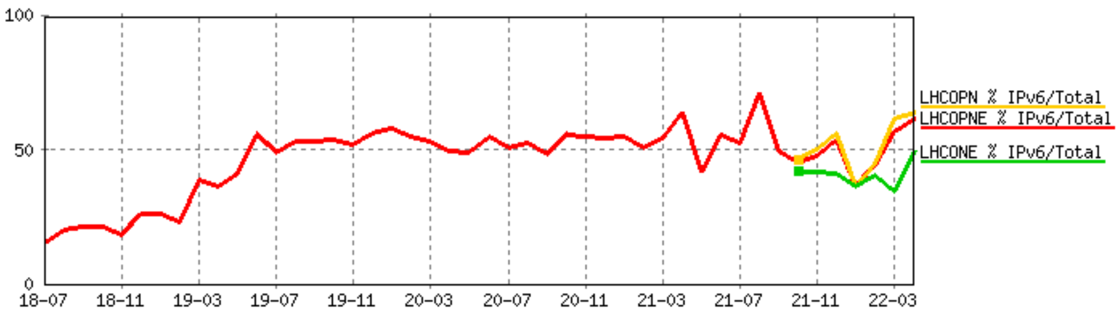
LHCOPN+LHCONE IPv4 and IPv6 traffic volumes month by month



IPv6 traffic on LHCOPN/ONE as seen at CERN

- ~40-70% of all traffic is IPv6
- from June 2019 onwards

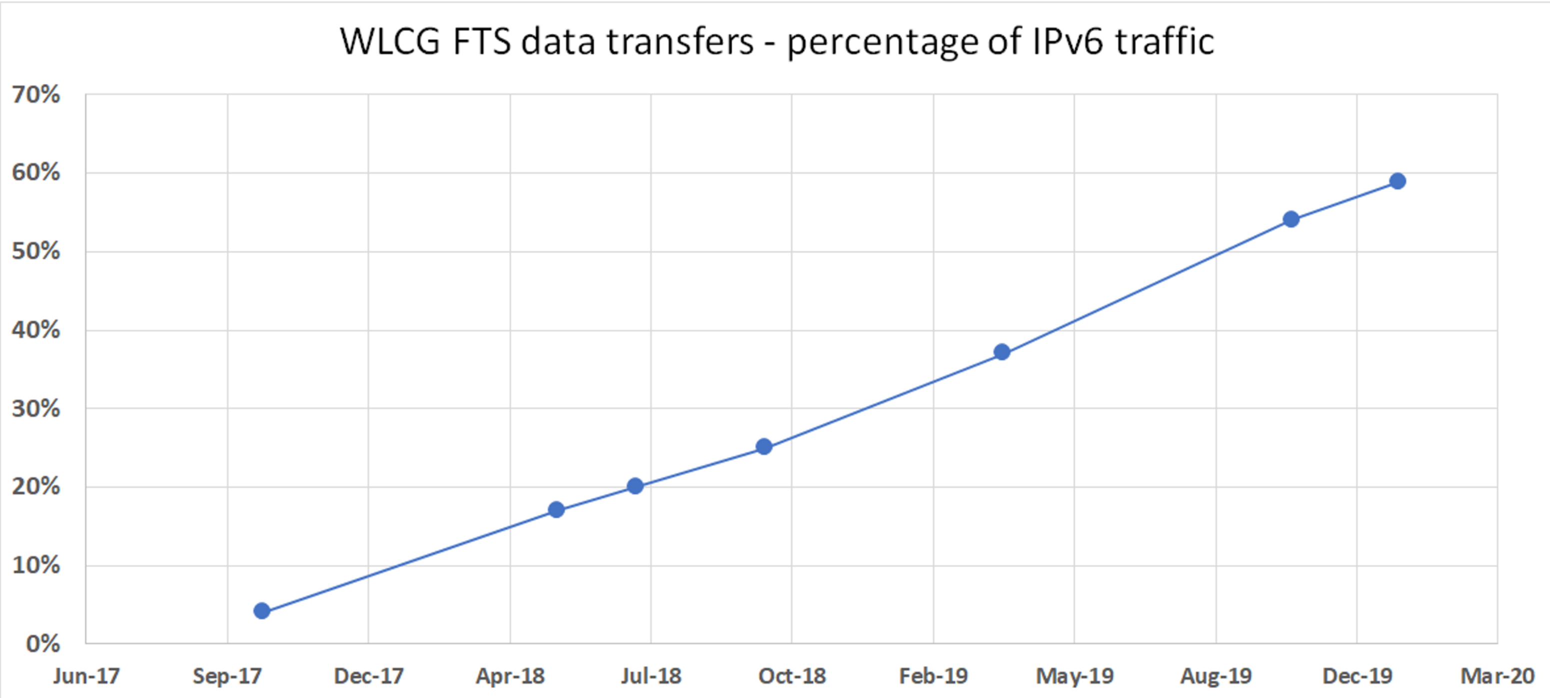
Percentage of IPv6 traffic over the total



Jisc

% of WLCG Data Transfers over IPv6

2017-2020   all experiments – File Transfer Service (FTS)



IPv6 works!

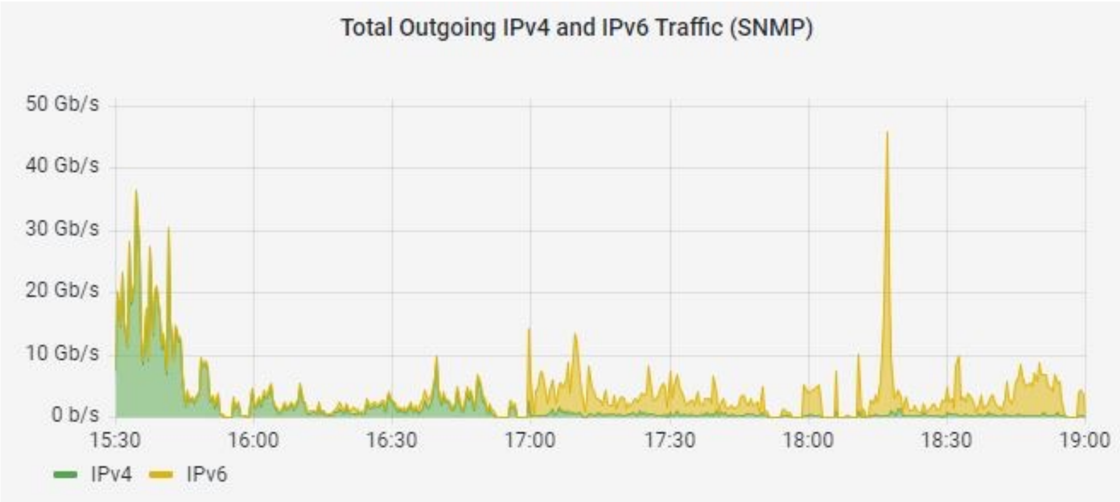
Experiments and physicists are happy

and unaware of the protocol used!

Jisc

IPv4/6 choice for dCache/WebDAV transfers

java.net.preferIPv6Addresses (default: false) **- Now set to "true"**



Green: IPv4; Yellow: IPv6

Default behaviour changed to prefer IPv6 at 17:00 local time on 14 Feb 2022

The fix works!

Now chasing all sites to change the configuration

perfSONAR

Measurement of network characteristics

- Free, open source – https://www.perfsonar.net
- Easy to download and install on CentOS7 (and Debian)
- **Very useful to have persistent testing**: collect history of network characteristics – throughput, loss, latency, path
  - Can test IPv4 and/or IPv6
- You can test against our 10Gbps node in the Jisc Slough data centre
  - https://ps-slough-10g.ja.net/, and *ps-slough-1g.ja.net* for latency
- We now also have a 100Gbps capable perfSONAR node in London
  - Available on *ps-london-bw.ja.net* and *ps-london-lat.ja.net*

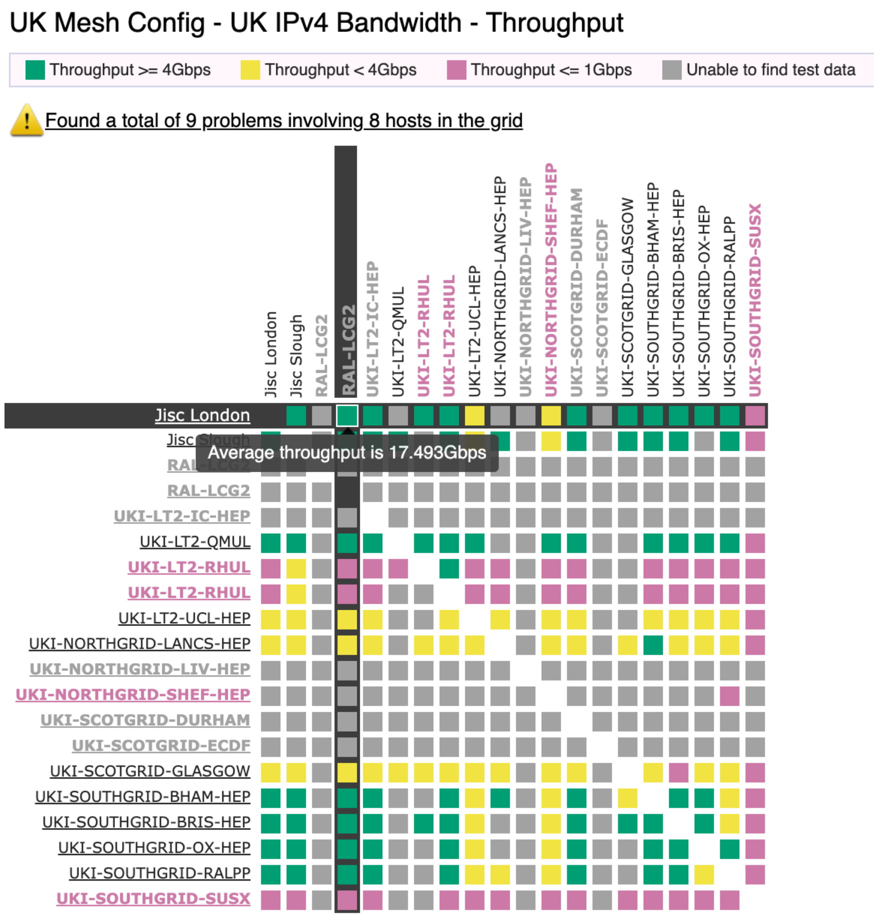
UK GridPP perfSONAR monitoring

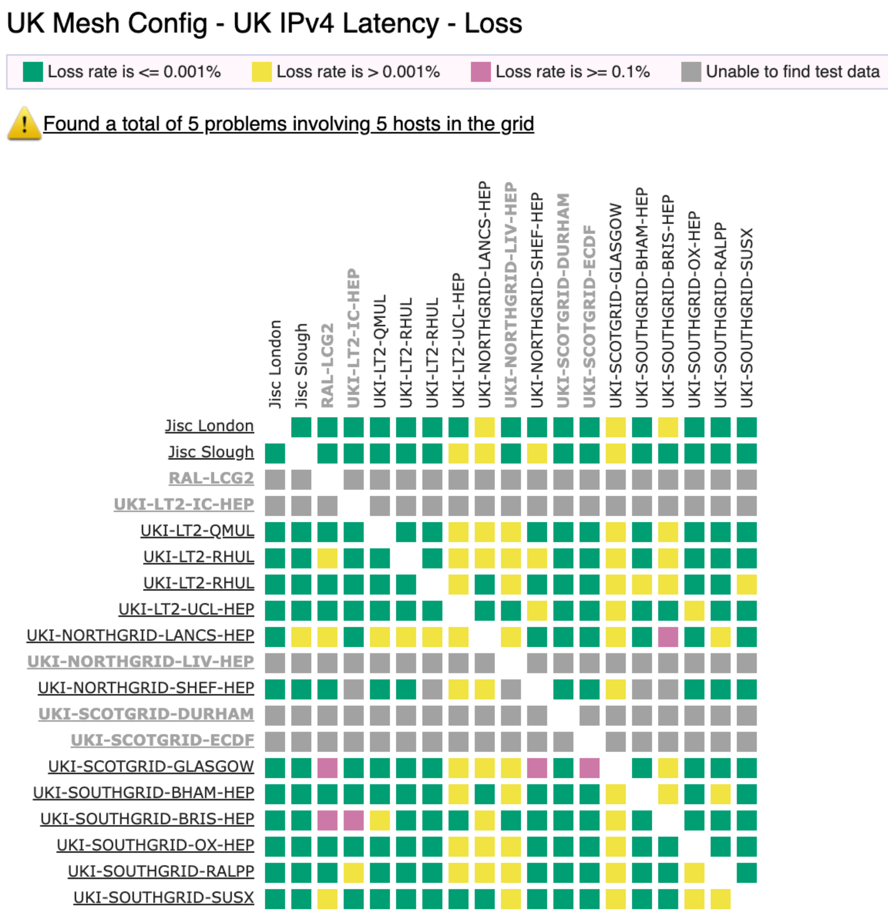
Network performance monitoring for IPv6 and IPv4

- UK GridPP mesh
  - https://psmad.opensciencegrid.org/maddash-webui/index.cgi?dashboard=UK%20Mesh%20Config
- Runs persistent loss/latency/path tests and runs throughput tests roughly every 6 hours between all perfSONAR servers in the mesh
- Includes an interesting multi-interface example
  - i.e., you should use one interface for bandwidth and a separate interface for latency/loss
  - See https://docs.perfsonar.net/manage_dual_xface.html

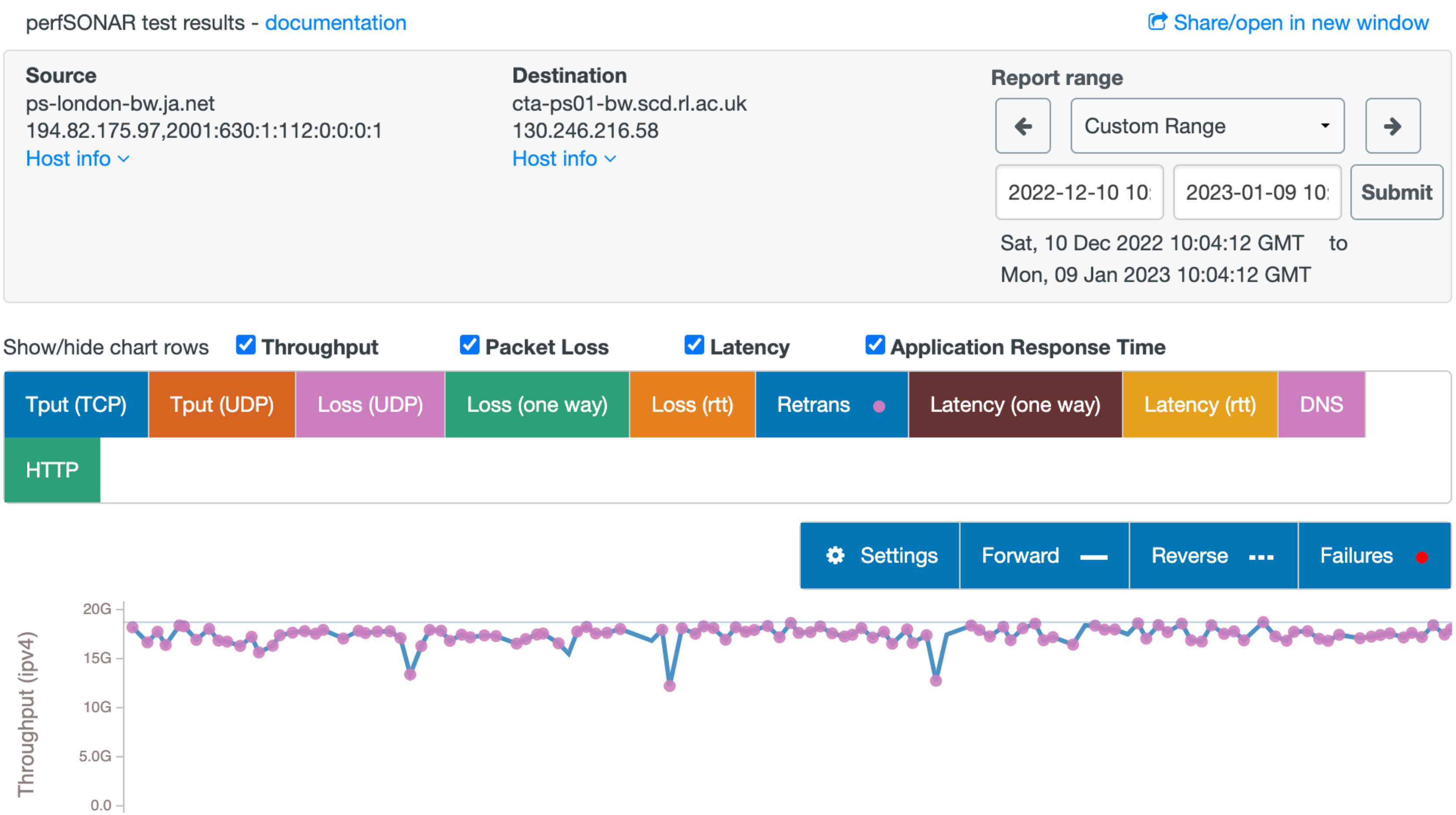
Jisc

UK GridPP mesh example





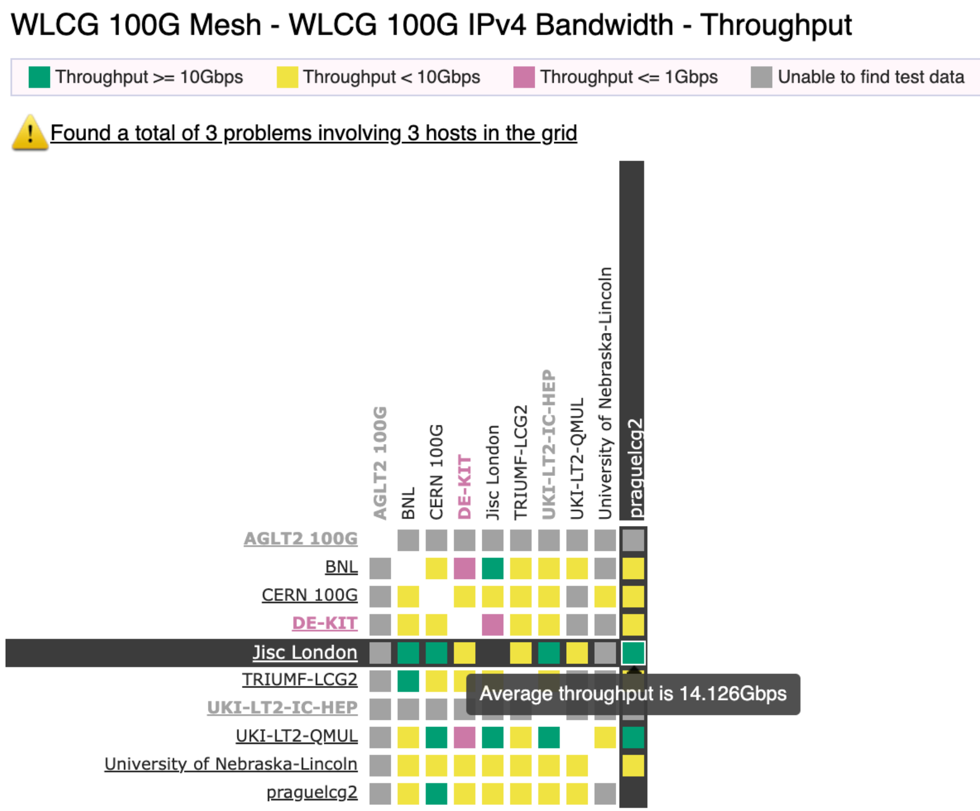
Drilling down… good performance, IPv4 only?



WLCG 100G perfSONAR mesh

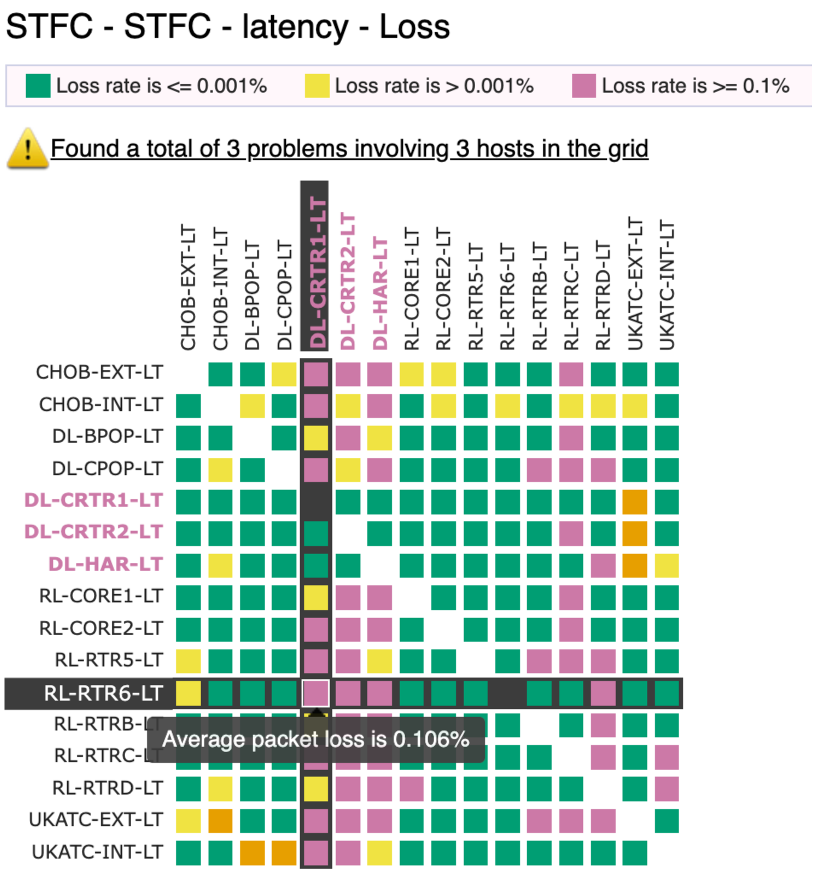
Testing 100G configurations

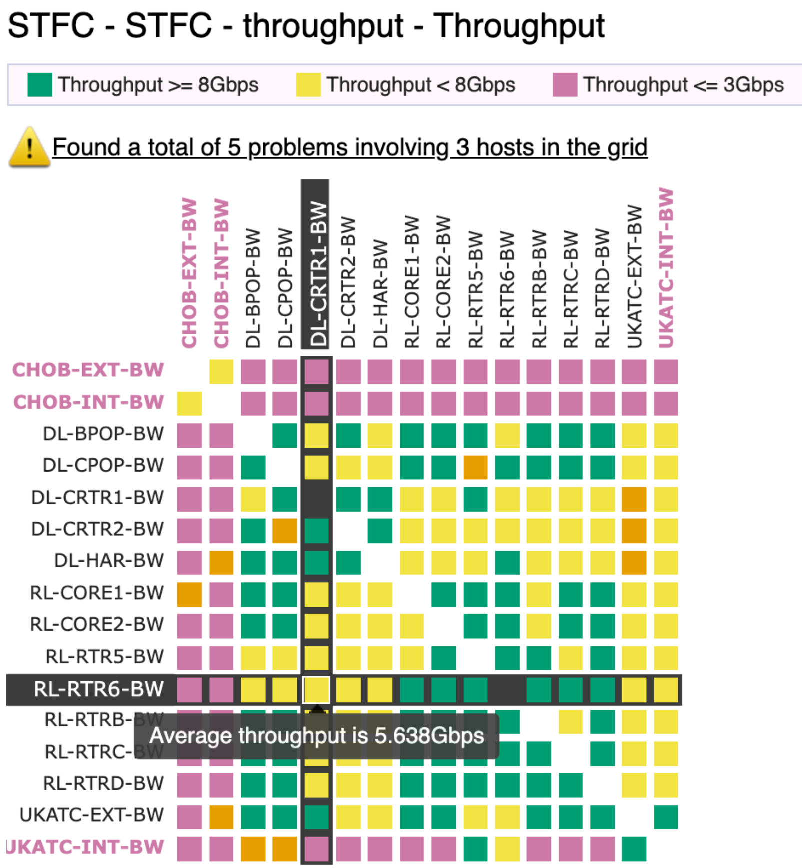
- For single stream iperf throughput tests with perfSONAR, 15-25Gbps is a good result.
- 100G single stream isn't realistic
- See
  - https://psmad.opensciencegrid.org/maddash-webui/index.cgi?dashboard=WLCG%20100G%20Mesh



Drilling down, here with IPv4 and IPv6

STFC perfSONAR – new(ish!)





https://psmaster.rl.ac.uk/maddash-webui/

Again drilling down (a RAL to DL example, IPv4)

https://psmaster.rl.ac.uk/maddash-webui/details.cgi?uri=/maddash/grids/STFC+-+STFC+-+throughput+-+Throughput/RL-RTR6-BW/DL-CRTR1-BW/Throughput

Measuring success

How can we measure RAL IPv6 adoption?

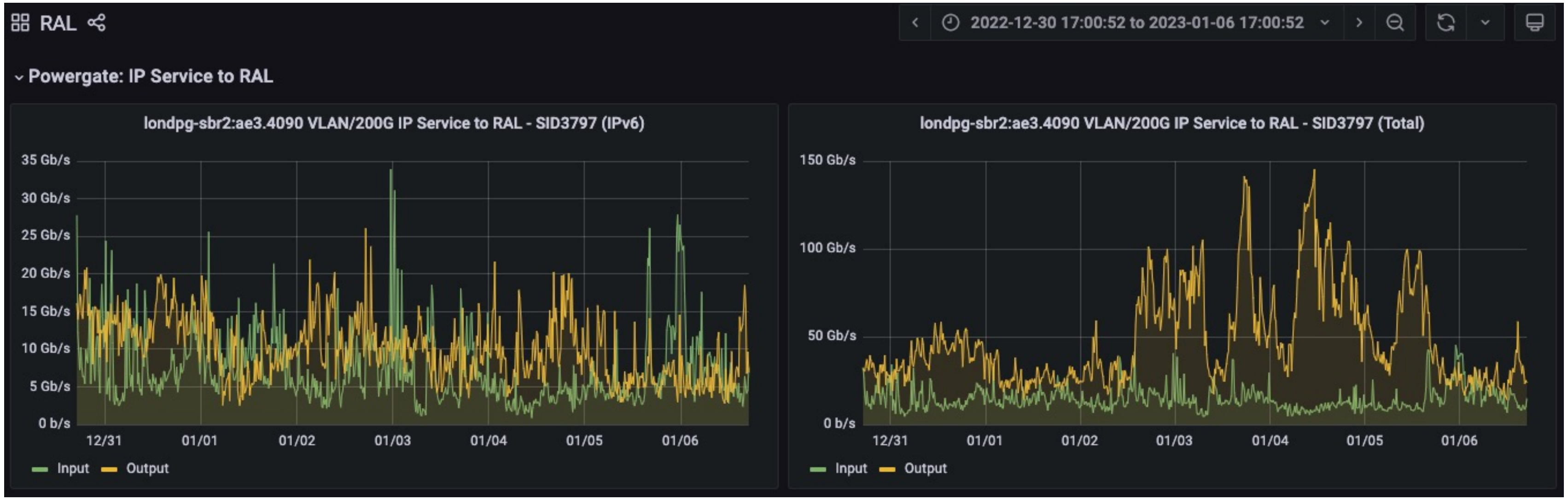
What insight do we have now? What more is needed?

- GridPP view – FTS logging, for example
- Look at the other logging for WLCG tools and experiments
  - As regularly presented at HEPiX IPv6 WG meetings
    - e.g. https://indico.cern.ch/event/1185115/
- Jisc Netsight is blind to IP version, but data is there to be collected
  - Jisc is awaiting new OSS tooling thus not enhancing Netsight yet
  - What about data from routers at RAL and STFC sites?
- A potential Jisc-STFC collaboration

IP(v6) traffic to/from RAL

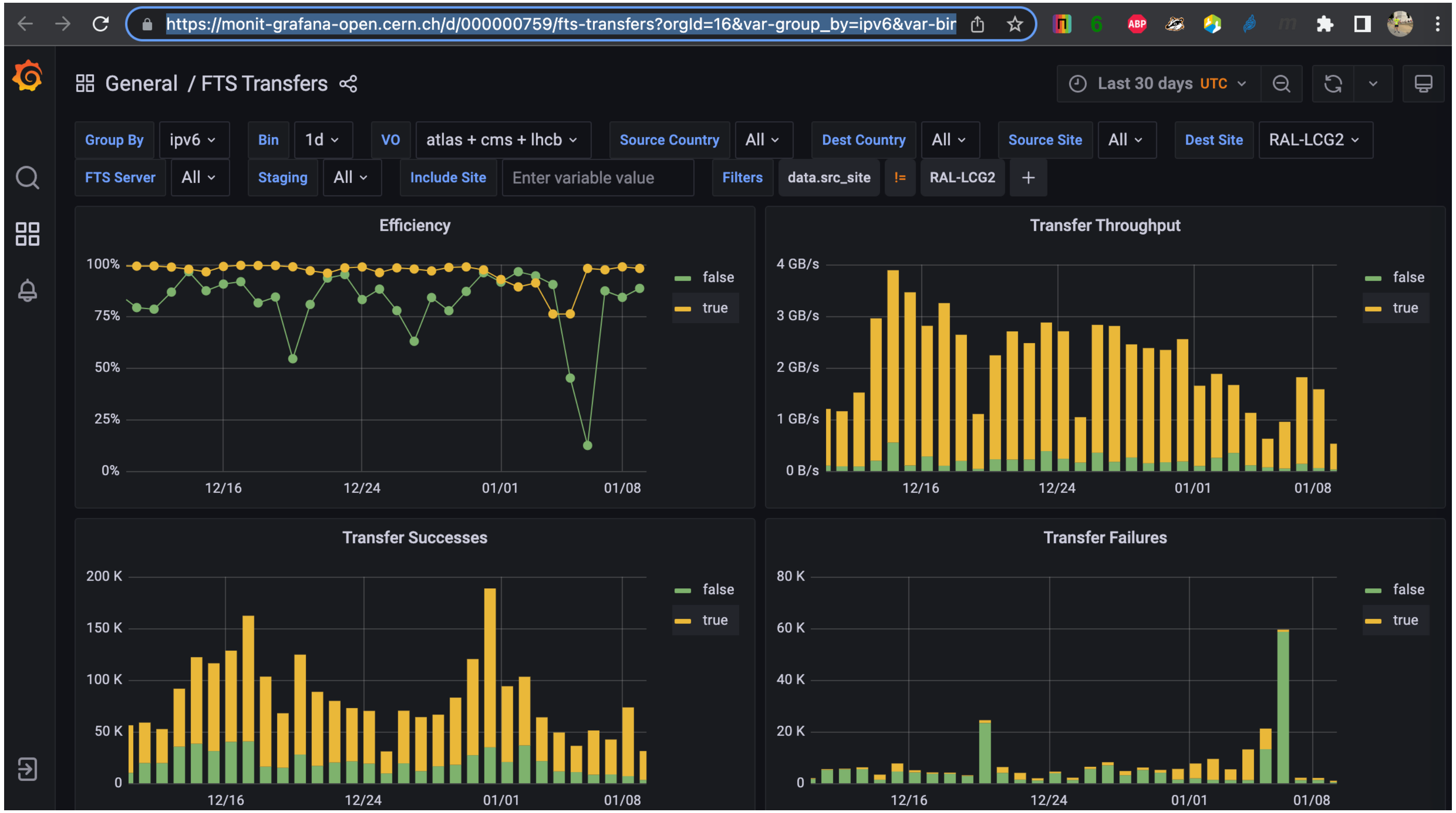
As seen by Jisc

- Rob Evans kindly produced a week of (probably atypical) RAL IP data
  - Shows a peak of 25-30Gbps IPv6 against a max of nearly 150Gbps
  - https://snapshots.raintank.io/dashboard/snapshot/4BnJ7qWDD1c2OyKTJfew85asYhYd9FKF?orgId=2



Jisc

IPv6 traffic – FTS view to RAL-LCG2



https://monit-grafana-open.cern.ch/d/000000759/fts-transfers?orgId=16&var-group_by=ipv6&var-bin=1d&var-vo=atlas&var-vo=cms&var-vo=lhcb&var-src_country=All&var-dst_country=All&var-src_site=All&var-dst_site=RAL-LCG2&var-fts_server=All&var-staging=All&var-include=&var-filters=data.src_site|!%3D|RAL-LCG2

(with thanks to Duncan)

Finding more information

IPv6 resources

Some useful links

- Jisc material:
  - IPv6 service page: https://www.jisc.ac.uk/ipv6
  - Advice and guidance page: https://www.jisc.ac.uk/guides/how-to-begin-an-ipv6-deployment
  - Janet IPv6 Technical Guide: https://repository.jisc.ac.uk/8349/1/janet-ipv6-technical-guide.pdf
  - Training: https://www.jisc.ac.uk/training/ipv6-fundamentals
  - And we're happy to spend time chatting with you, just ask
  - JiscMail ipv6-users@jiscmail.ac.uk list, subscribe at https://jiscmail.ac.uk/IPv6-users

- Other sources:
  - UK IPv6 Council: https://www.ipv6.org.uk/ - recent talks on GridPP and AWS support
  - UKNOF: https://www.uknof.org.uk/ - many talks on IPv6
  - Procurement : https://www.ripe.net/publications/docs/ripe-772 (updated - was RIPE 554)

Discussion and next steps

Discussion

Next steps?

- Are you now more comfortable progressing IPv6 deployment at RAL?
- Are there things that you need to know more about?
- Are there specific concerns or current problems we didn't touch on?
- Are the next steps clear?
- Can we measure the use (and non-use) of IPv6?

- I'm very happy to work with you as needed via Jonathan and others

Contact:

Tim Chown

tim.chown@jisc.ac.uk

4 Portwall Lane,
Bristol, BS1 6NB

customerservices@jisc.ac.uk

jisc.ac.uk