The HEPiX IPv6 working group (Update)

David Kelsey (STFC UKRI)

HEPiX Online Workshop, 26 April 2022

On behalf of all co-authors in the HEPiX IPv6 working group

Active in HEPiX IPv6 Working Group – last 12 months

- M Babik (CERN), M Bly (RAL), N Buraglio (ESnet), T Chown (Jisc), D Christidis (U Texas/ATLAS), J Chudoba (Prague), C Condurache (EGI.eu), J Flix (PIC), C Grigoras (CERN/ALICE), B Hoeft (KIT), H Ito (BNL), D P Kelsey (RAL), E Martelli (CERN), S McKee (U Michigan), R Nandakumar (RAL/LHCb), K Ohrenberg (DESY), F Prelz (INFN), D Rand (Imperial), A Sciabà (CERN/CMS), E Simmonds (FNAL), T Skirvin (FNAL)
- Many more in the past, and others join from time to time
- *and thanks also to WLCG operations, WLCG sites, LHC experiments, networking teams, monitoring groups, storage developers…*

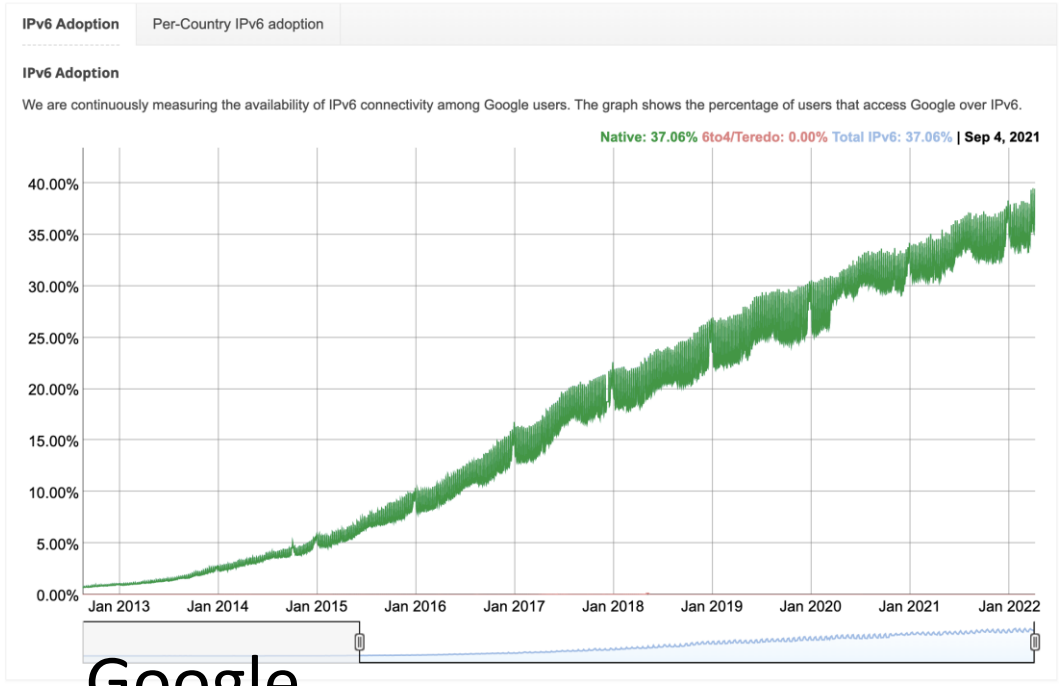
Overview

- IPv6 WG meetings since October 2021 HEPiX workshop
  - 2 December 2021, 18/19 January 2022, 3 March, 7 April, (3 * 1 hour & 1 * 2-halfdays – all virtual)
- Agenda of 18/19 Jan 2022 meeting included: (I will include some of these today)
  - News from sites and experiments
  - Dual-stack storage status (Tier1), Tier2
  - News from RNTWG & packet marking subgroup
  - IPv6 in ESnet
  - IPv6-only testing
  - Encourage deployment of dual-stack WNs

https://indico.cern.ch/event/1115437/

- Why do data transfers use IPv4 between two dual-stack endpoints?
- Plans for IPv6-only
- Reminders for WLCG sites
- Summary

General IPv6 traffic continues to grow (Google and Facebook)



Google

Facebook

IPv6 use cases

- Main driver was and still is: “support IPv6-only CPU”
  - Ongoing concern of lack of IPv4 addresses (at some sites)
- Other (more recent) drivers
  - Support packet marking in TCP for monitoring by RNTWG
    - Research Networking Technical Working Group
      - Packet marking includes use of IPv6 headers
  - US Federal Government - 80% of services to be IPv6-only by 2025-26
    - We hear regular updates from Phil DeMar & Tim Skirvin (Fermilab), Nick Buraglio (ESnet) and Hiro Ito (BNL)

IPv6 use case (packet marking) Scientific Network Tags (Jan 2022)

(slide from Marian Babik) - www.scitags.org

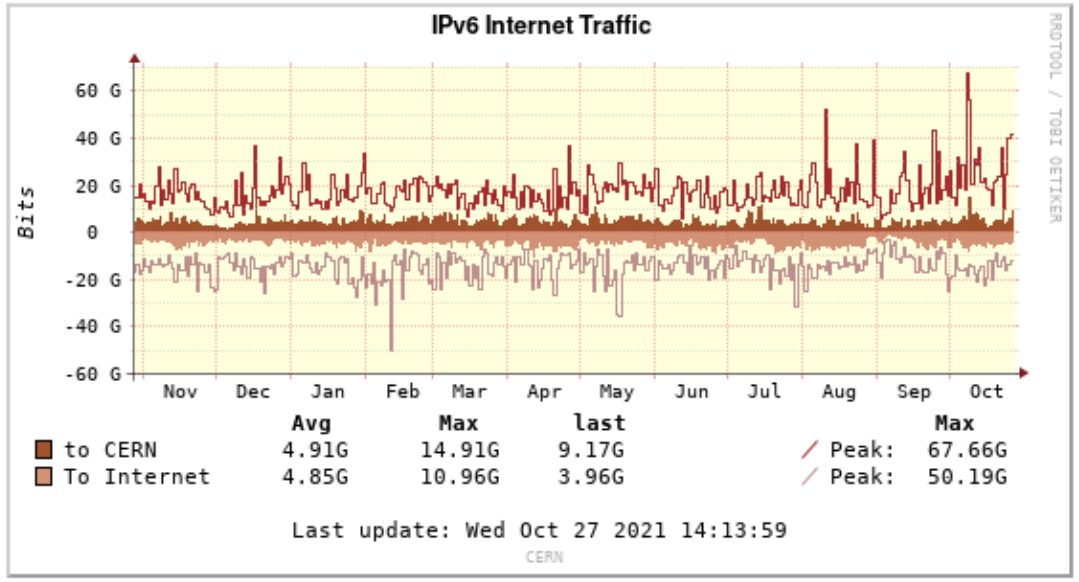
Concepts

- Marking is based on two different approaches
  - **Flow marking** using UDP fireflies (works for both IPv4 and IPv6)
  - **Packet marking** using IPv6 flow label and/or header extensions
- Both carry **flow identifier**, which at present is an encoded representation of experiment/science domain and activity
  - For UDP fireflies flow id can be extended with other fields in the future
  - For packet marking the space is restricted due to number of bits available in the headers
- Experiments and activities need to be registered prior to their usage
  - Registry serves this purpose and ensures RENs and DDMs have consistent view
- Designed to work with proxies, cached proxies and private networks
- Generators, collectors, storage and analytics can evolve independently

8

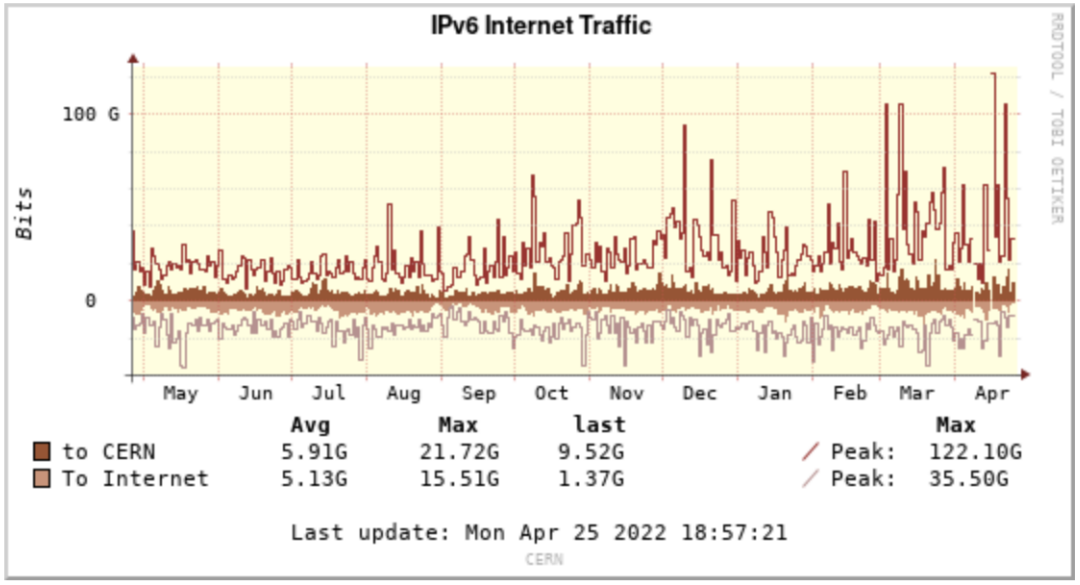
CERN: all traffic to IPv6 Internet (Oct 2021 and April 2022) – steady growth

Yearly



Graphic help ?

Yearly

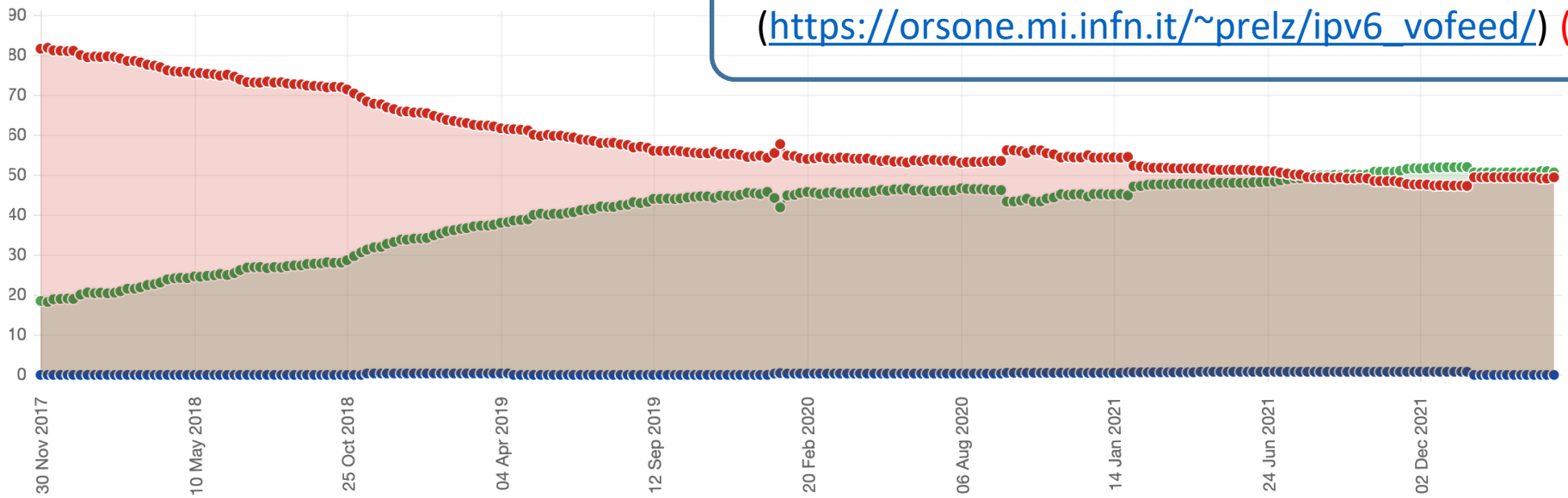


Graphic help ?

Dual-stack services in WLCG (green line)

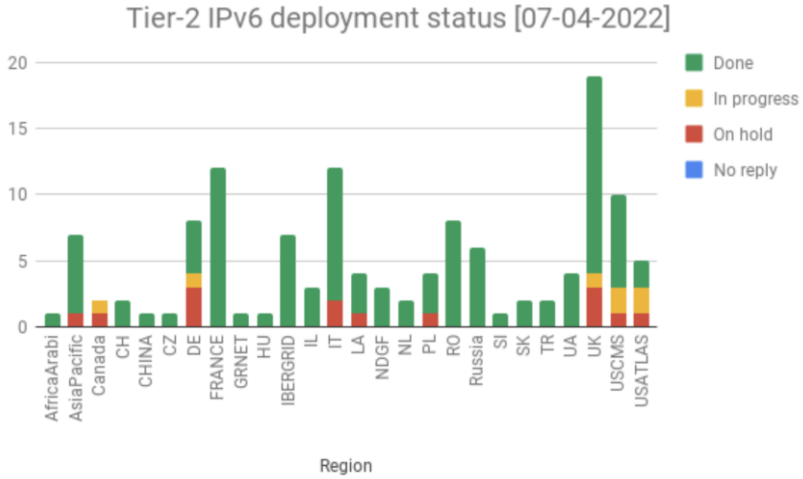
The graphs below record, on a weekly basis (every Thursday at 06:00 CET) the fraction of service endpoints listed in the VO Feeds of the 4 major LHC experiments (Alice, Atlas, CMS, LHC-B) where the DNS returns an IPv4-only (A) resolution (red line), a dual-stack IPv6-IPv4 (A+AAAA) resolution (green line) or an IPv6-only resolution (cyan line). The graph is meant to provide a bird's eye view of the IPv6 transition at WLCG sites.
Comments and complaints → ipv6@hepix.org.

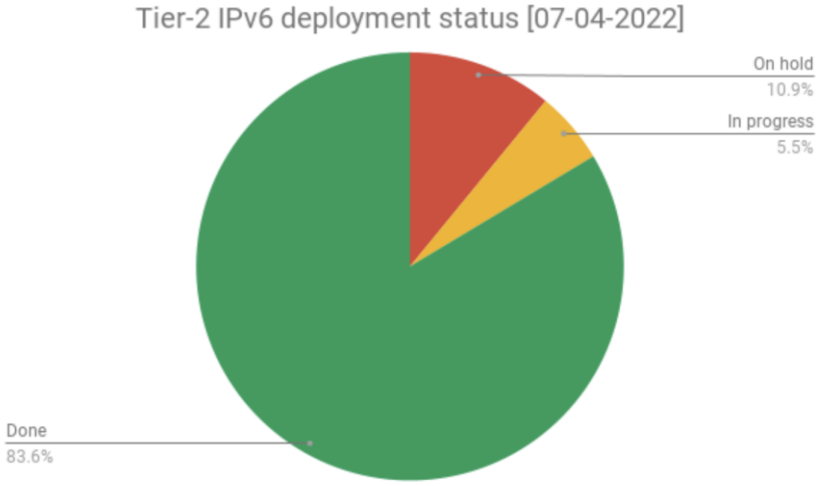
results from the LHC Experiment VO feeds (https://orsone.mi.infn.it/~prelz/ipv6_vofeed/) (~51%)



IPv4/IPv6 deployment at Tier-2 sites

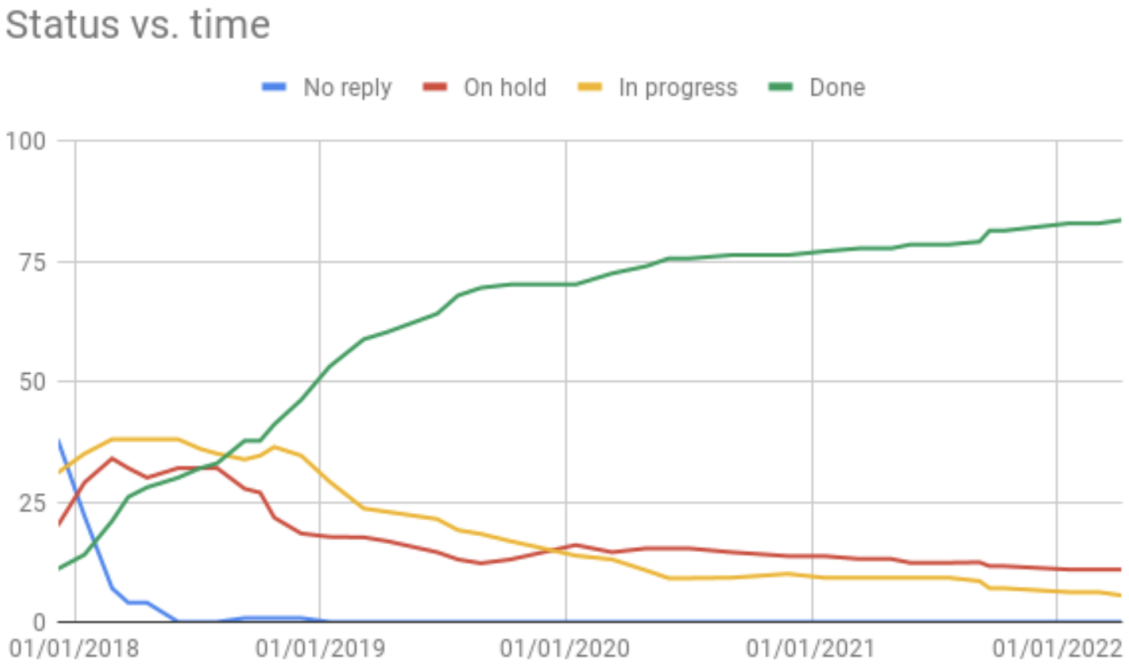
- The deployment campaign was launched in November 2017
- Steady progress (status)
  - ~84% of Tier-2s have dual stack storage
  - 82% of storage

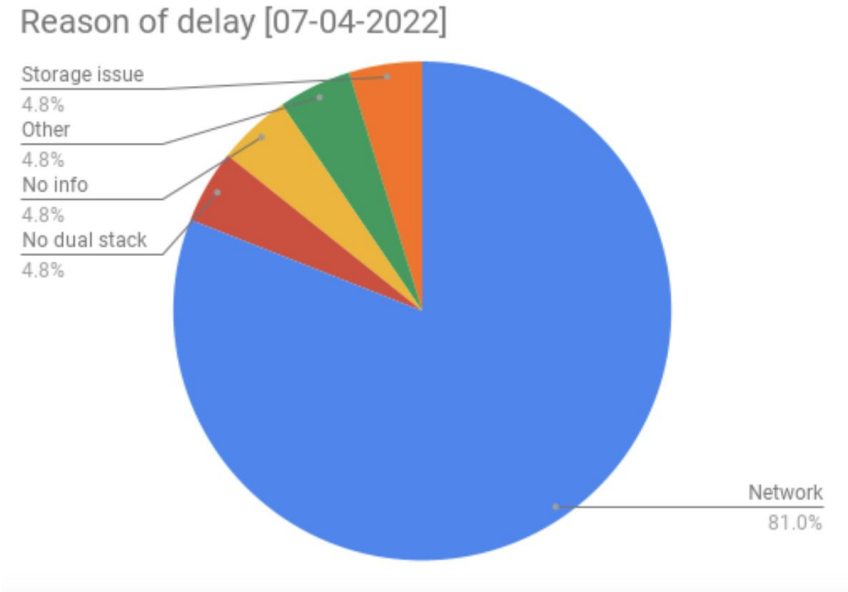




Experiment	Fraction of T2 storage accessible via IPv6
ALICE	89%
ATLAS	72%
CMS	94%
LHCb	79%
Overall	82%

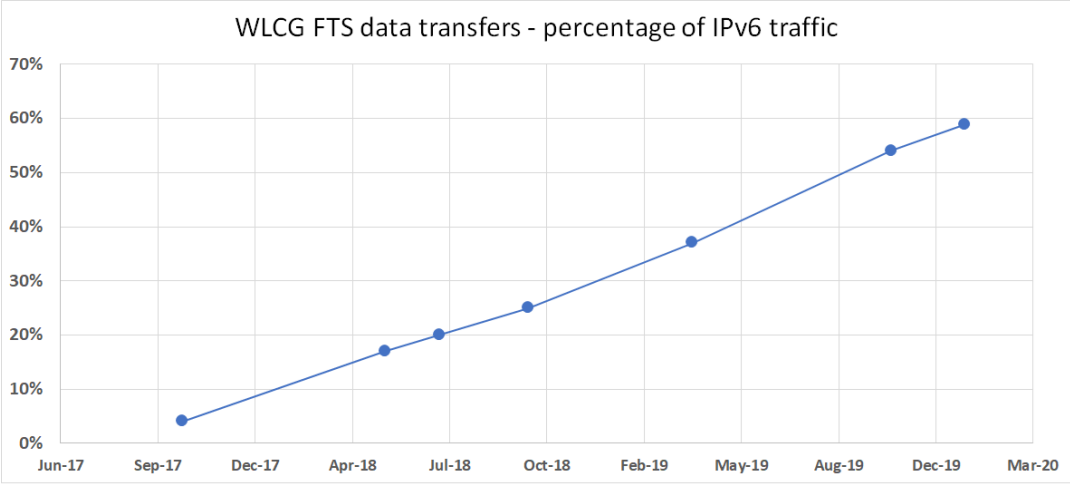
Tier-2 evolution of dual-stack



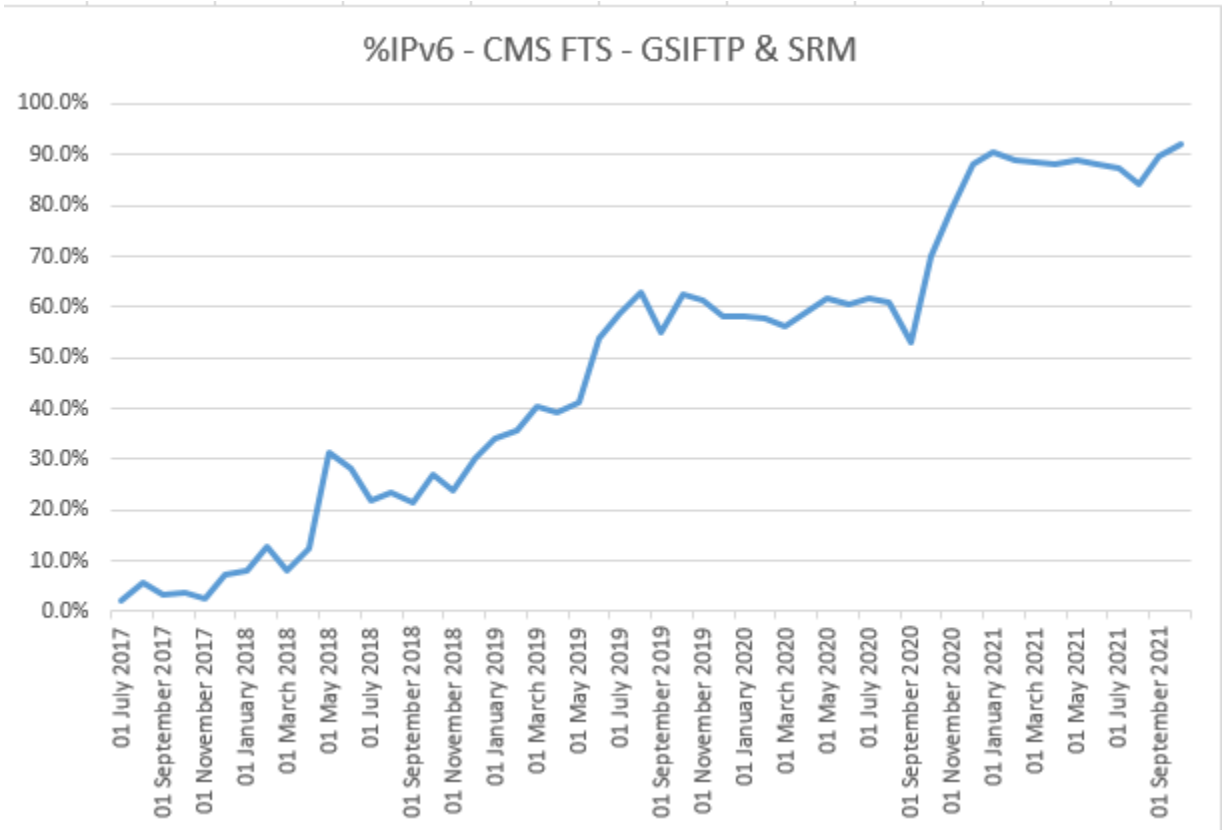


% of (srm & gsiftp) FTS traffic over IPv6 (OLD)

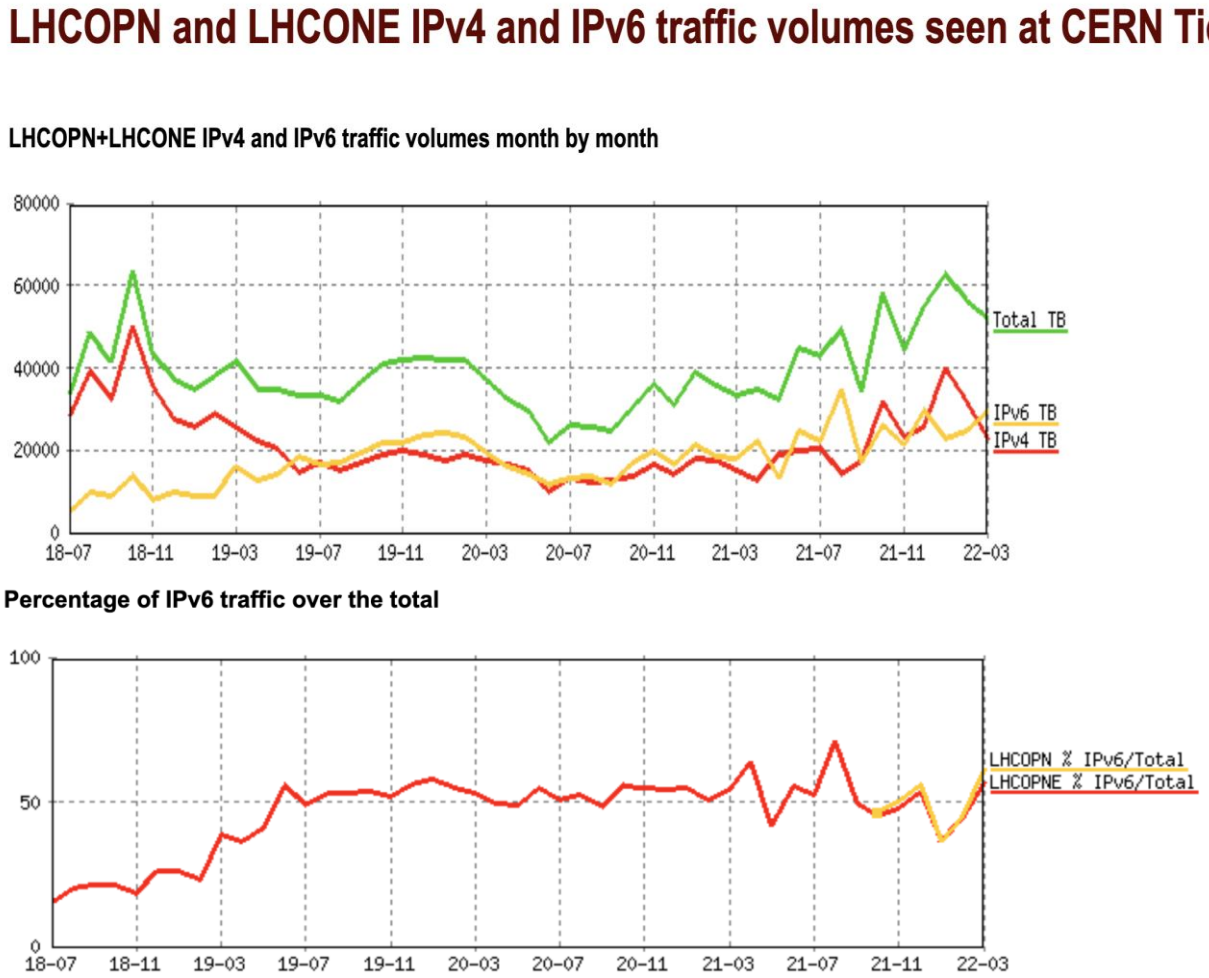
- Some FTS protocols, e.g. DAVS still not able to monitor IPv6 traffic
- these are excluded from this plot
- And explains end dates!



All protocols & all experiments



IPv6 traffic on LHCOPN/LHCONE at CERN

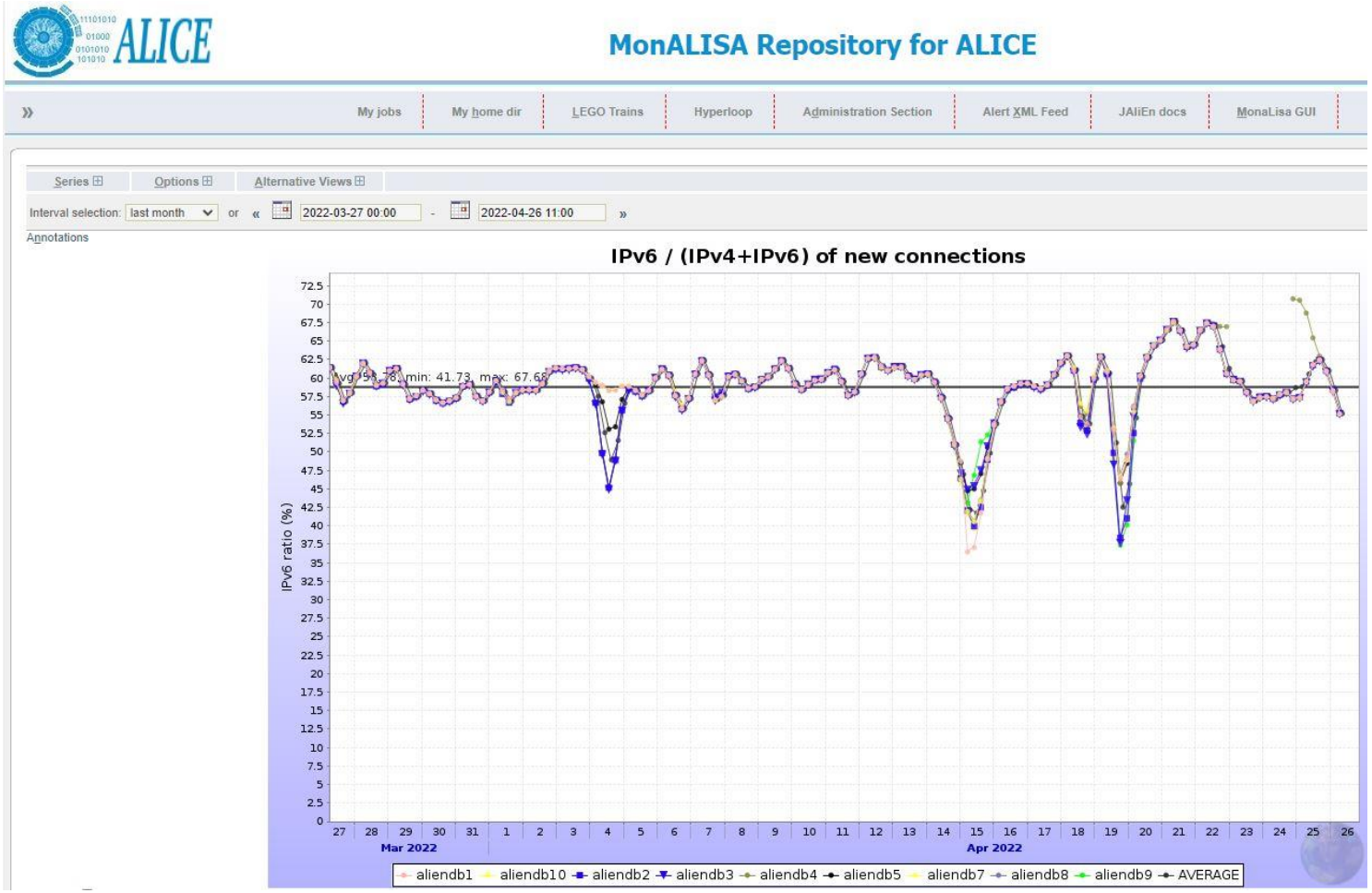


[LINK](#) to these plots

IPv6 traffic on LHCOPN/ONE as seen at CERN

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

Monitoring of ALICE connections to CERN services (average ~60%)

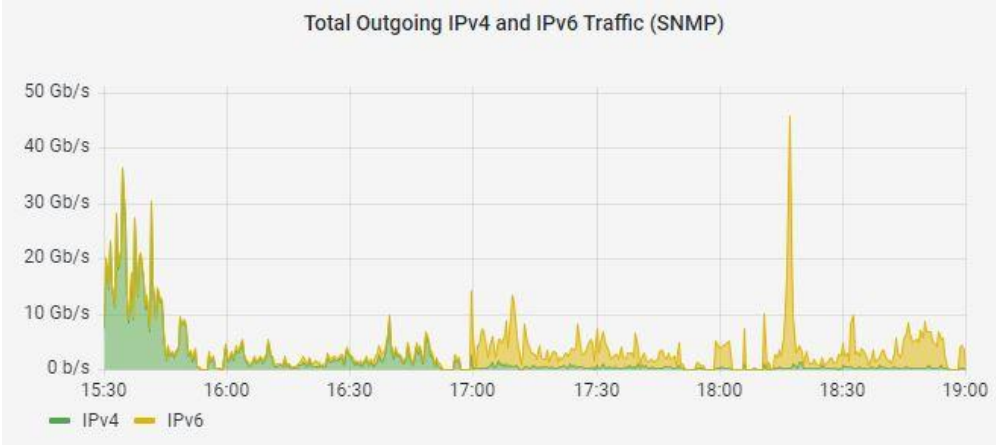


Why: IPv4 transfers on LHCOPN?

- Tier1s are dual-stack, but IPv4 often still used for transfers
  - Site/experiment issues - some storage end-points not yet dual-stack
    - Old software stacks (legacy deployments) are still deployed
  - both ends are dual-stack but configuration or request prefers IPv4
  - transfers are to/from WN's and the WNs are still IPv4-only
    - Difficult to monitor WN IPv6 traffic centrally – can sites monitor?
- IPv6 WG has been analyzing Tier-1 top-talkers over IPv4
  - We should aim to understand behaviour of each Tier-1
- Plan for a CERN technical student to assist with this analysis
- encourage all sites to deploy all WNs as dual-stack
- encourage all sites and all experiments to "prefer" IPv6 transfers

IPv4 choice for dCache/WebDAV transfers

- Found in outgoing transfers from US-ATLAS (BNL, AGLT2, MWT2) and at KIT
  - *Traffic outbound over WebDAV – chooses IPv4 even when dual-stack end-points involved*
- https://docs.oracle.com/javase/7/docs/api/java/net/doc-files/net-properties.html
- **java.net.preferIPv6Addresses** (default: false)
  When dealing with a host which has both IPv4 and IPv6 addresses, and if IPv6 is available on the operating system, the default behavior is to prefer using IPv4 addresses over IPv6 ones. This is to ensure backward compatibility, for example applications that depend on the representation of an IPv4 address (e.g. 192.168.1.1). This property can be set to **true** to change that preference and use IPv6 addresses over IPv4 ones where possible.
- Fix by "This property can be set to **true** to change …"
- **see this plot at AGLT2 (LHC data challenge 14 Feb 2022)**
  - **Preference changed at ~17:00**



dd a custom DnsResolver class that prefers IPv6 addresses.

Fix to dCache (for choice of IPv6)

- dCache V7.2.11
- Https://dcache.org/old/downloads/1.9/release-notes-7.2.shtml#release7.2.11

POOL

- HTTP-TPC transfers now prefer IPv6 address, if both endpoints support it.
- Out-of-the-box, Apache HTTP client will try connections in the order returned by the system's DNS resolver. This order is not guaranteed, but seems to favour IPv4 in real-world environments.
- Added a custom DnsResolver class that prefers IPv6 addresses.

IPv6 WG priority for 2022-23

- Find and fix reasons for still preferring IPv4

WLCG - from dual-stack to IPv6-only (CHEP2019) https://doi.org/10.1051/epjconf/202024507045

- Planning for an IPv6-only WLCG
- To simplify operations
  - dual-stack infrastructure is the most complex!
- Large infrastructures (e.g. Facebook, EE/BT) use IPv6-only internally
- The plan - the goal we are working towards
  - IPv6-only for the majority of WLCG services and clients
  - With ongoing support for IPv4-only clients where needed
- Timetable to be defined

Messages to WLCG sites

WLCG MB statements (July 2021)

- Deployment of dual stack storage remains the priority
  - this is a prerequisite to fully supporting IPv6-only WNs
- All sites and regions should plan accordingly and as soon as possible
- The final goal is IPv6-only (timetable to be agreed later)

Encouragement: IPv6 WG to all sites and experiments

- deploy all WN, VM, containers as dual-stack
- Configure to enable and "prefer" IPv6 transfers

Summary

- WLCG is ready to support use of IPv6-only CPU resources
- Tier-1s all have production storage accessible over IPv6
- Tier-2s ~84% sites done (>82% of Tier-2 storage is dual-stack)
- The WG is keen to see Data Transfer monitors fully instrumented to track IPv6
- We will continue to analyse and fix reasons for IPv4 data transfers
- All sites/experiments:
  - please deploy dual-stack on all clients and services
  - and configure IPv6 as preferred protocol
- WG is planning for the move to IPv6-only services
- Next WG meetings: 12 May 2022, 28/29 June 2022 (In person at CERN)
- ***message to new research communities - build on IPv6 from start!***

Questions, Discussion?