



Science and
Technology
Facilities Council

Ensuring Use of IPv6 after it is deployed

David Kelsey (UKRI/STFC-RAL)
HEPiX/LHCONE, UVIC,
Victoria, BC, Canada - 18 October 2023



On behalf of all members of the HEPiX IPv6 working group - (many thanks all!)



M Babik (CERN), M Bly (RAL), N Buraglio (ESnet), T Chown (Jisc),
D Christidis (CERN/ATLAS), J Chudoba (FZU Prague), P Demar (FNAL), J Flix (PIC),
C Grigoras (CERN/ALICE), B Hoeft (KIT), H Ito (BNL), D P Kelsey (RAL),
E Martelli (CERN), S McKee (U Michigan), C Misa Moreira (CERN),
R Nandakumar (RAL/LHCb), K Ohrenberg (DESY), F Prelz (INFN), D Rand
(Imperial), A Sciabà (CERN/CMS), T Skirvin (FNAL)

(underlined names – here this week)

- Many more in the past, and members join/leave from time to time
- *many thanks also to WLCG operations, WLCG sites, LHC experiments, networking teams, monitoring groups, storage developers...*

Overview

Outline of talk

- IPv6 on WLCG (history) – the HEPiX IPv6 working group
- Dual-stack IPv6/IPv4 storage deployment
- Monitoring
- IPv6-only WLCG
- Ensuring use of IPv6
- Summary





Science and
Technology
Facilities Council

IPv6 on WLCG – history



Science and
Technology
Facilities Council



Why did we start a working group?

2010-11

- Survey of HEPiX Community (Sep 2010) – “IPv6 readiness”
 - National NRENs are ready; Universities and Labs are not ready
 - Some lack of IPv4 address space, including CERN (WLCG wish to avoid use of NAT)
- IANA projecting imminent IPv4 address exhaustion
- Sep 2010 – memo from US Federal CIO to all depts including Department of Energy (HEP national labs) - Deploy dual-stack!
- Offers of opportunistic CPU resources could arrive and be IPv6-only
- Our middleware, software, technology and tools are not yet IPv6 capable
 - This will take lots of time to fix - so started a working group in April 2011!

HEPiX IPv6 Working Group

2011-16 Phase 1

- full analysis of work to be done
 - Applications, middleware, system and network tools, operational security
- Ran a distributed test-bed
- Timetable and plan
 - **Initial plan for support of IPv6-only clients was 2014**
- Test important data transfer protocols, technology and data storage/file systems
 - DPM, dCache, xRootD, OpenAFS, FTS, CASTOR, ...
- Found *many* problems needing work
 - Worked closely with developer community
- **Concluded IPv6 support will be much later than 2014!**

IPv6 Deployment on WLCG (Phase 2)

2017 onwards (as approved by WLCG Management Board)

- All Tier1 storage services in IPv6/IPv4 dual-stack mode from April 2018
- Tier-2 storage services
 - Aim for large number of dual-stack Tier-2s **by end 2018**



Science and
Technology
Facilities Council

Dual-stack IPv6/IPv4 deployment (on WLCG Tier2s)

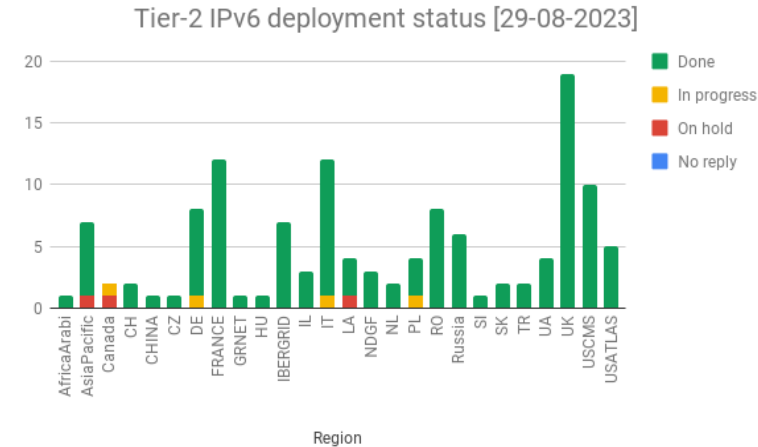
**Tier1 storage has been IPv6-capable for
several years**



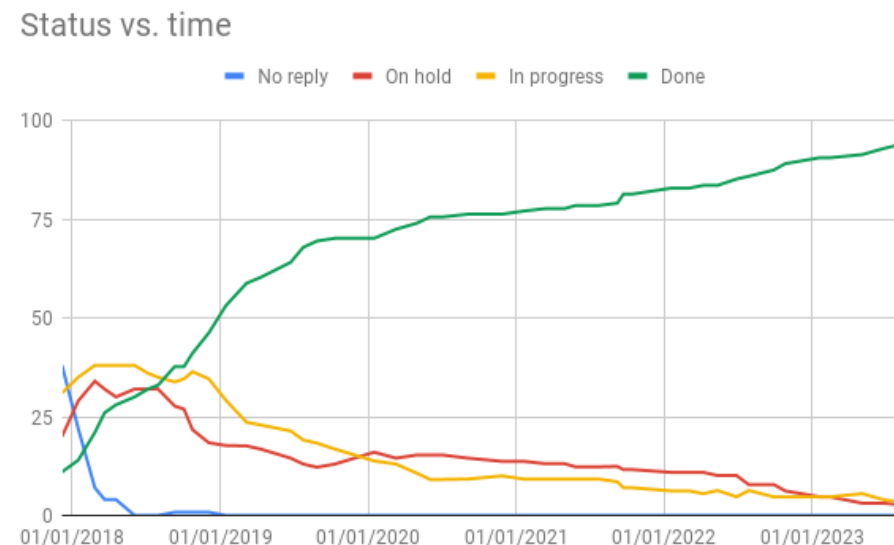
Science and
Technology
Facilities Council

Good news! - IPv6/IPv4 at Tier-1/2 sites

- Tier-1 complete
- Tier-2 deployment from Nov17
- ([status](#)) shows >94% T2 sites
 - **97%** of Tier-2 storage dual stack



Experiment	Fraction of T2 storage accessible via IPv6
ALICE	91%
ATLAS	95%
CMS	100%
LHCb	100%
Overall	97%





Science and
Technology
Facilities Council

Monitoring



Science and
Technology
Facilities Council



Importance of monitoring

- We must monitor
 - deployment of IPv6-capable services
 - fraction of data transfers taking place over IPv6
- Monitoring implementations used for IPv6
 - perfSONAR
 - ETF - experiment test framework
 - FTS (File Transfer Service)
 - Network utilisation and traffic plots
 - e.g. IPv6 versus IPv4 on LHCOPN/LHCONE
- But in recent years some existing **monitoring stopped working**
 - FTS over WebDAV not tracking IPv6 (GSIFTP and SRM was instrumented)
 - work has been done to fix this problem

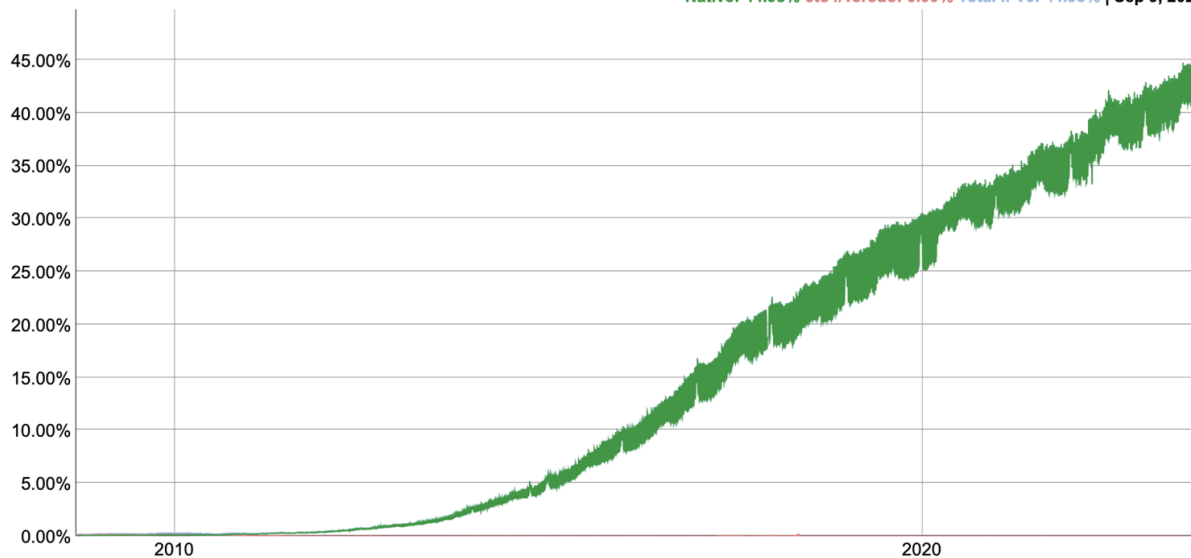
IPv6 traffic continues to grow

Google

IPv6 Adoption

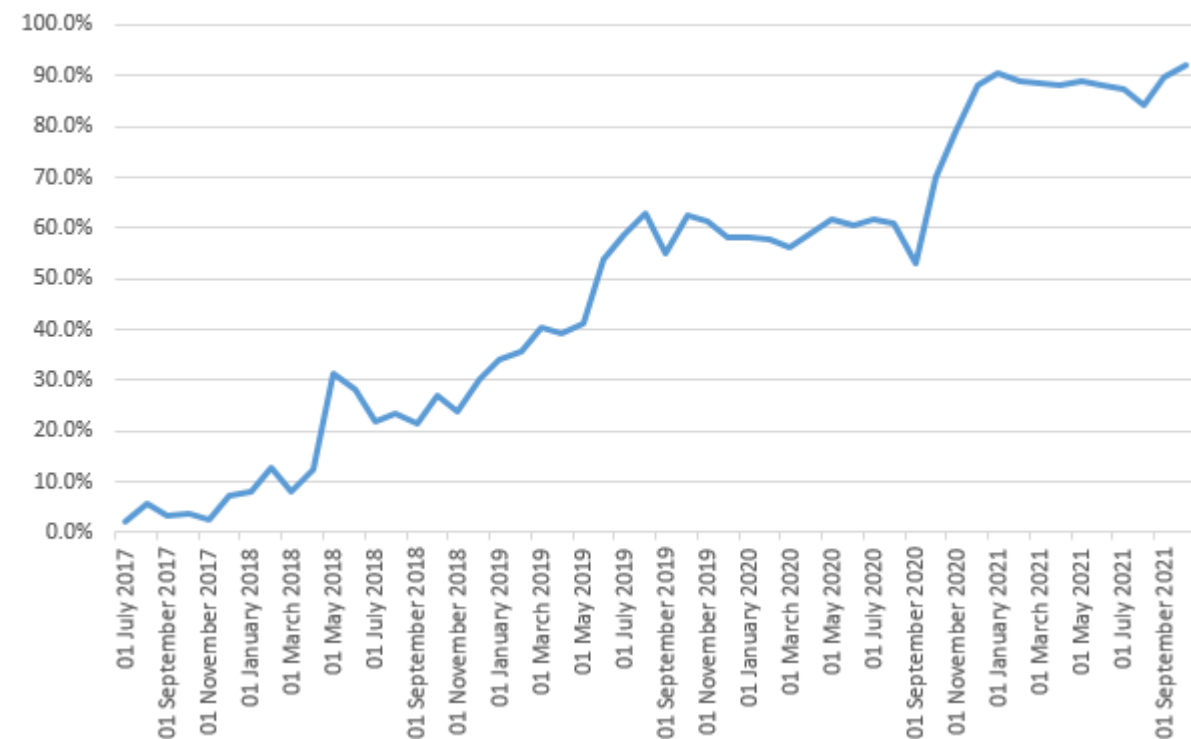
We are continuously measuring the availability of IPv6 connectivity among Google users. The graph shows the percentage of users that access Google over IPv6.

Native: 44.98% 6to4/Teredo: 0.00% Total IPv6: 44.98% | Sep 9, 2023



WLCG Data Transfers

%IPv6 - CMS FTS - GSIFTP & SRM



Stops when CMS moved away from GSIFTP & SRM

Phase 3 - IPv6-only

Drivers for use of IPv6

- Sites running out of routable IPv4 addresses (avoid NAT)
 - Use IPv6 addresses for external public networking
- To be ready to support use of IPv6-only CPU clients
- There are **other drivers** for IPv6:
 - scitags.org – packet marking (in header of IPv6 packets)
 - Research Networking Technical Working Group ([RNTWG](#))
 - USA Federal Government – [directive](#) on “IPv6-only” (Nov 2020)



Use of the IPv6 Flow Label for WLCG Packet Marking

Dale W. Carder - LBNL / ESnet (presenter)

Tim Chown - Jisc

Shawn McKee - University of Michigan

Marian Babik - CERN

draft-cc-v6ops-wlcg-flow-label-marking

IETF 117, San Francisco, 25 July 2023

Rationale

- Complex workflows used by multiple data-intensive science communities
 - ~1.4M x86 cores across ~170 sites w/ ~1.6 EB of storage
 - Individual network flows usually small, but can aggregate to many 10's Gbit/s
- Traffic on purpose-built networks (LHCOPN, LHCONE) as well as R&E Networks
 - **Predominantly IPv6**, working towards **IPv6 exclusively**
- Mark packets to identify traffic owner/purpose.
 - Coarse definitions of community/activity provides insight *in aggregate*
- Track data transfers with *existing* network flow monitoring (IPFIX & sFlow)
 - Quantify global behavior and analyse tradeoffs at scale
 - ex: dataset & storage placement, job scheduling
- Potential future use for traffic engineering

3

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
 - Dual-stack is less secure
- Large infrastructures (e.g. Facebook, Microsoft,...) use IPv6-only internally
- The goal we are working towards
 - IPv6-only for the majority of WLCG services and clients
 - Do we support IPv4-only clients? - still to be decided
 - Plan that this will not be needed
- Timetable still to be defined and agreed with Management Board

How to ensure use of IPv6?

“Obstacles” to IPv6

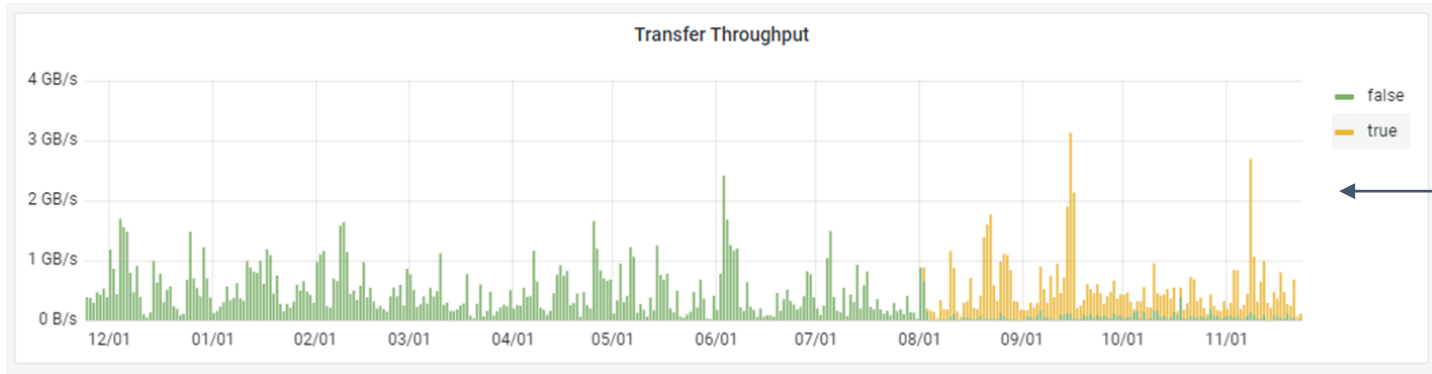
There are many reasons stopping the full use of IPv6/IPv4

- Dual stack is an essential step on the journey to IPv6-only

The Obstacles that we have been addressing:

1. **WLCG Sites not yet deployed IPv6 networking** ~done
2. **Sites have IPv6 but Tier-2 has no dual-stack storage** ~done
3. **IPv6 monitoring not available or broken** see next slide
4. **Service is dual-stack but IPv4 being used** see next slide

Some obstacles fixed (#3 and #4)



Some FTS **monitoring** now able to distinguish IPv6 from IPv4

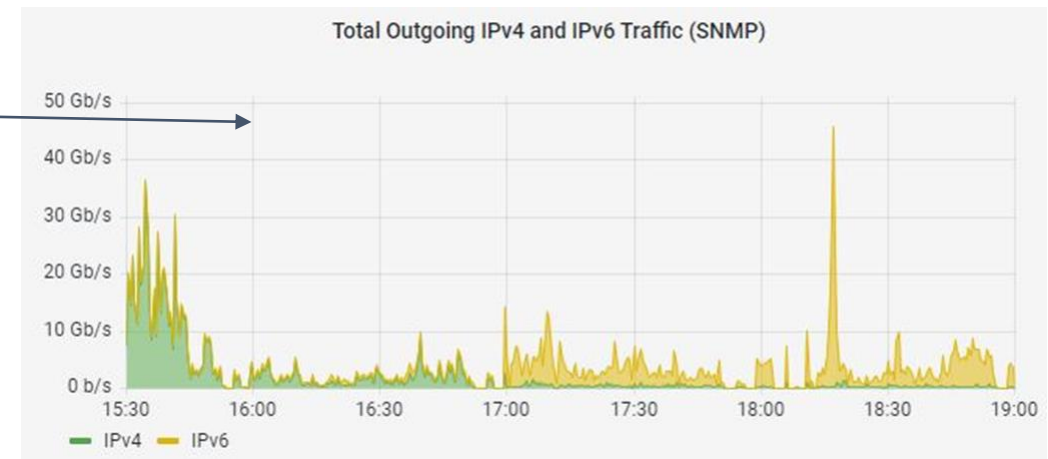
ATLAS & CMS HTTP transfers into CERN (last year)

– IPv6 showing from August 2022 onwards

IPv6 is yellow

Data transfers into USA/ATLAS Great Lakes Tier 2 (AGTL2) Found to **use IPv4** even when both ends dual-stack (dCache/WebDAV) **java.net.preferIPv6Addresses (default: false) - Now set to “true”** Fixed at 17:00 on 14 Feb 2022 (confirmed in the plot!) This fix is essential for all dCache instances - fixed in v7.2.11

Many other uses of IPv4 have been investigated
htcondor, xRootD, FTS and Top-100 talkers LHCOPN using IPv4
Many problems identified and fixed.



WorkerNode migration to IPv6 at KIT

(Bruno Hoefft)

- <https://indico.jlab.org/event/459/contributions/11661/>
- migrate the CPU (WorkerNode) farm towards IPv6
- monitoring of ALL WN network traffic
 - Packetbeat on all nodes storing to OpenSearch and analysed with Kibana
- Initially a small subset of WN, then the whole farm
 - 0.5 TB of data in 6 days
- Apr22 - 28% IPv6; Dec22 - 67% IPv6
- Ongoing detailed work (many applications) to keep improving
- Shows how effective monitoring and fixing can be

Obstacles to IPv6 - still to be addressed

5. **Non-storage services not yet dual-stack**
 - a. ~60% of all WLCG services are dual-stack today
 - b. New GGUS ticket campaign (Compute) will start soon
6. **WLCG client CPU (worker nodes, VMs, containers) many IPv4-only**
7. **Services/clients outside of WLCG Tier-1/Tier-2 not yet considered**
 - a. Tier-3, Public/Commercial Clouds, Analysis facilities, Experiment portals...
8. **Use of new or evolving technologies not yet tested or tracked**
 - a. New CPU architectures (GPU, non-x86, ...), container orchestration, ...
9. **“People” can be the obstacle**
 - a. they do not consider use of IPv6 or refuse to deploy!
10. **Analysis of old data using old software**
 - a. e.g. ALICE analysing Run 2 data with IPv4-only version of xRootD
 - b. possible that all experiments have such a requirement

We will try to fix where possible - but much is outside of our control!

A collection of IPv6 traffic plots (WLCG)

Imperial London - LHCONe - 100 Gbps on IPv6

<https://shapingthefutureofjanet.jiscinvolve.org/wp/uncategorized/100gbps-of-cern-data-over-ipv6-on-the-janet-network/>



Figure 1 — Imperial monitoring shows the two-hour period where the 100G link was filled and where 100% of the LHCONe traffic was IPv6.

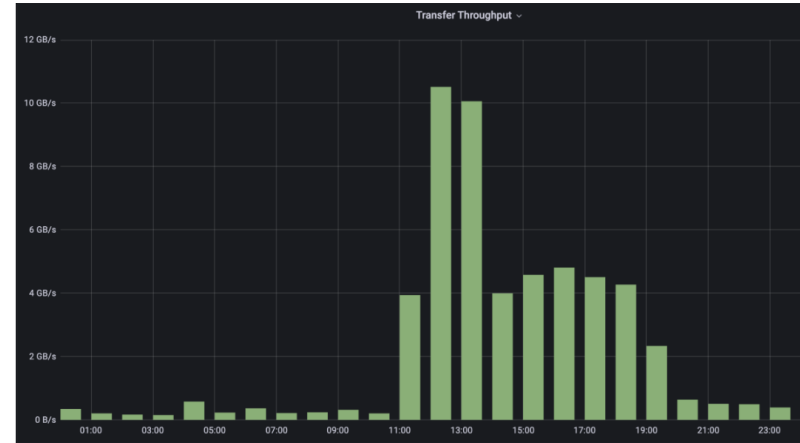


Figure 2 — The traffic levels seen in the network view correspond to those seen by the WLCG File Transfer Service (FTS) visualization tools.



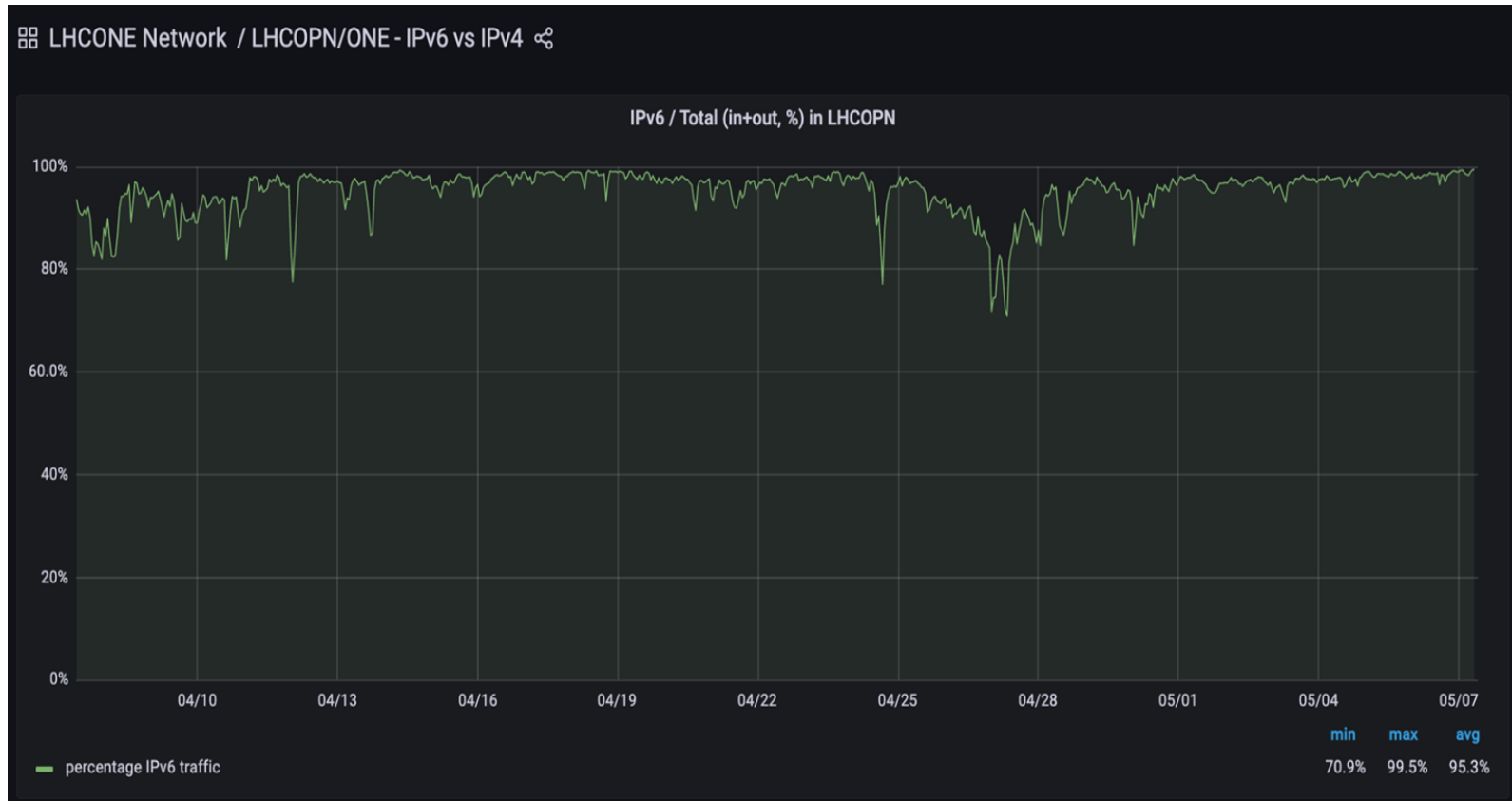
Figure 3 — It was also interesting to see this traffic reflected in the monitoring platform for the GÉANT pan-European research and education backbone network.





IPv6 on WLCG after removing several “obstacles” during the last year

LHCOPN network (at CERN) ~95% IPv6 during 30 days April/May 23



USA/ESnet LHCONE network

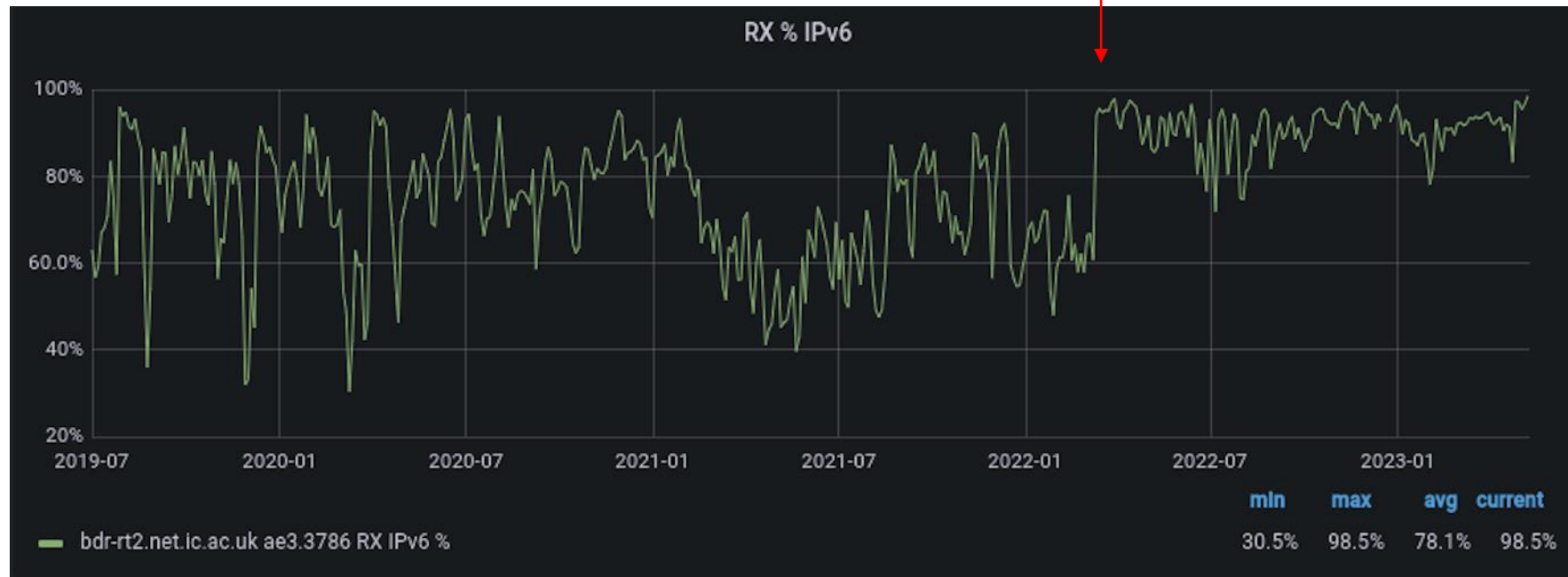
(Dale Carder - talk at Internet2 meeting Sep 2023)

“Can we turn off IPv4 on LHCONE. IPv6 is often over 90%”

%IPv6 on LHCONe (Imperial College London)

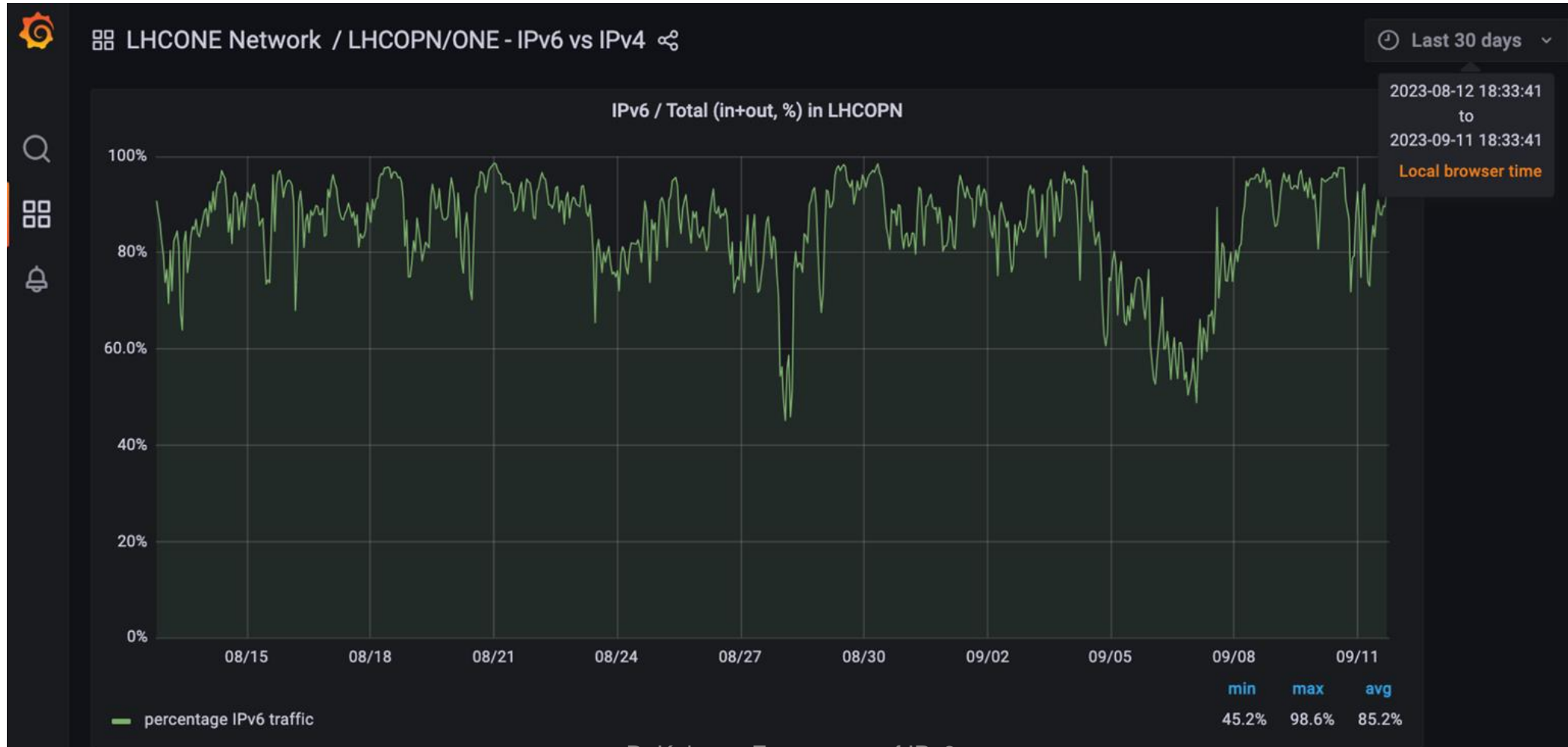


dCache storage preference set to IPv6

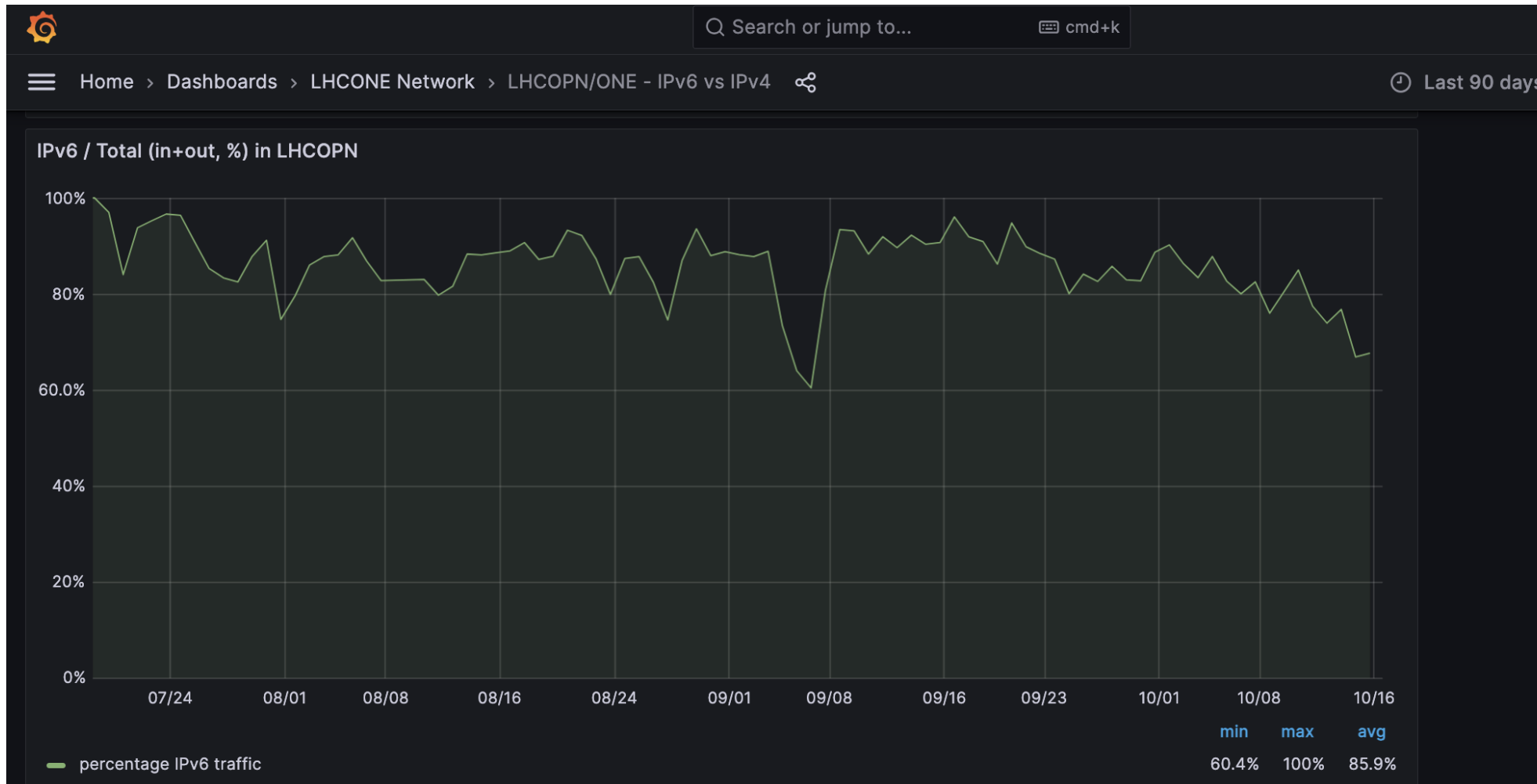


Since Feb 2022
~90% IPv6

LHCOPN traffic - %IPv6 (large drops) August/September 2023



LHCOPN %IPv6 at CERN – Last 90 days (IPv4 only CPU will be a reason – deploy dual-stack!)



Summary

- WLCG now supports IPv6-only clients
- Tier-1s: complete; Tier-2s: 97% storage is IPv6 capable
- Most data transfers use IPv6
- We have concentrated on ensuring use of IPv6
 - LHCOPN/LHCONE can be 90-95% IPv6 - **but not always!**
- We continue to address more obstacles to IPv6 in WLCG
 - All WLCG CPU services and CPU clients to dual-stack is priority now
 - Approved by WLCG MB on 17 Oct 2023
- **End point is still IPv6-only services (IPv4 is “legacy” networking)**
- ***Message to new research communities - build on IPv6 from start***

More information

Some papers from the HEPiX IPv6 working group

a) *“IPv6 Security”*

- M Babik et al 2017 J. Phys.: Conf. Ser. 898 102008
- <http://dx.doi.org/10.1088/1742-6596/898/10/102008>

b) *“IPv6 in production: its deployment and usage in WLCG”*

- M Babik et al, EPJ Web of Conferences 214, 08010 (2019)
- <http://dx.doi.org/10.1051/epjconf/201921408010>

c) *“IPv6-only networking on WLCG”*

- M Babik et al EPJ Web of Conferences 245, 07045 (2020)
- <http://dx.doi.org/10.1051/epjconf/202024507045>

d) CHEP2023 submission - *“Overcoming Obstacles to IPv6 on WLCG”*

- M Babik et al, EPJ Web of Conferences (under review)



Science and
Technology
Facilities Council

Questions?