

# **IPv6-only on WLCG**

## ***update from the IPv6 working group***

Andrea Sciabà (CERN), David Kelsey (STFC, UKRI)  
on behalf of the HEPiX IPv6 Working Group

HEPiX Spring 2021 online Workshop  
16 March 2021

# On behalf of all members of the HEPiX IPv6 working group

## *Recently active in the HEPiX IPv6 Working Group*

- M Babik (CERN), M Bly (RAL), T Chown (Jisc), D Christidis (U Texas/ATLAS), J Chudoba (Prague), C Condurache (EGI.eu), C Grigoras (CERN/ALICE), B Hoeft (KIT), 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), D Stockland (Imperial)
- 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...*

# Outline

- WLCG transition to dual-stack IPv4/IPv6
- WLCG data transfers – fraction over IPv6
- Monitoring (ETF)
- IPv6-only networking (the final phase!)
- Summary

*For WG meetings see (since Feb 2020 – no longer face to face!)*

<https://indico.cern.ch/category/3538/>

*Meetings of IPv6 working group since last HEPiX Workshop (Oct 2020):*

*One-hour meetings: 22 Oct, 26 Nov, 25 Feb*

*Two half-days: 19/20 Jan*

# News since October 2020

- WLCG MB mandates **complete deployment of dual-stack storage** and encourages to be ready for IPv6-only batch nodes (23 Feb 2021 meeting)
- Data transfers – **move from SRM, GSIFTP to XRootD and https/webdav**
  - Should all work over IPv6 but FTS IP monitoring limited to GSIFTP & SRM
  - FTS will extend IP stack monitoring to all protocols in the coming months
- Some data transfers between sites which are both IPv6 capable still go over IPv4 - problems are still to be understood
  - We **encourage deployment of dual-stack worker nodes** everywhere
- New **IPv6-only testbed at CERN**
- Sampling network traffic (INFN) - to **look for IPv6-only problems**
- Dual-stack on Kubernetes requires version 1.20

# IPv6 and Tier-1 storage

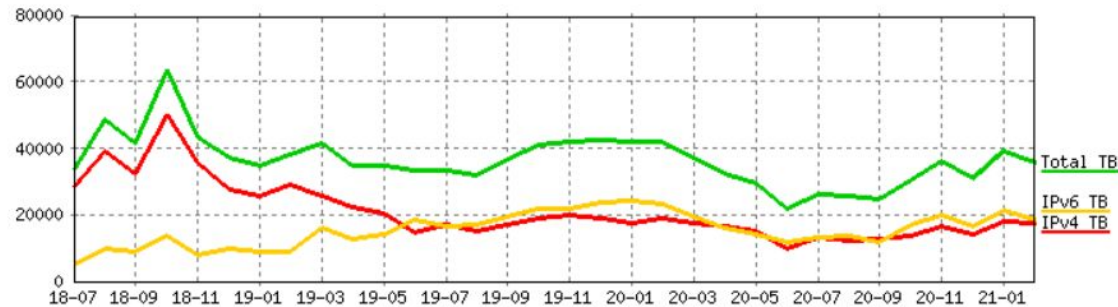
- FTS servers at BNL, CERN, and RAL work in dual stack, while Fermilab is still IPv4-only
  - CMS/Fermilab now use the CERN FTS server, so IPv6 transfers are being used
- FTS transfers happen via IPv6 at most Tier-1s
- IPv6 transfers do not happen at RRC-KI

Experiment	Fraction of T1 storage accessible via IPv6
ALICE	78%
ATLAS	96%
CMS	100%
LHCb	94%
Overall	96%

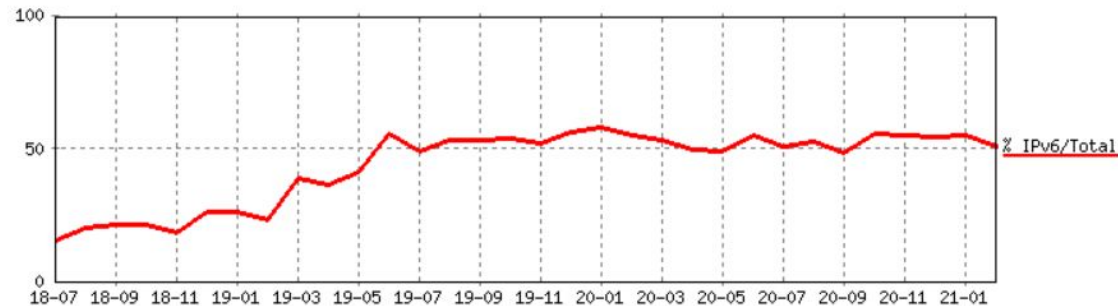
# IPv6 traffic on LHCOPN/LHCONE at CERN

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

IPv4 and IPv6 traffic volumes month by month



Percentage of IPv6 traffic over the total



IPv6 traffic on LHCOPN & LHCONE as seen at CERN

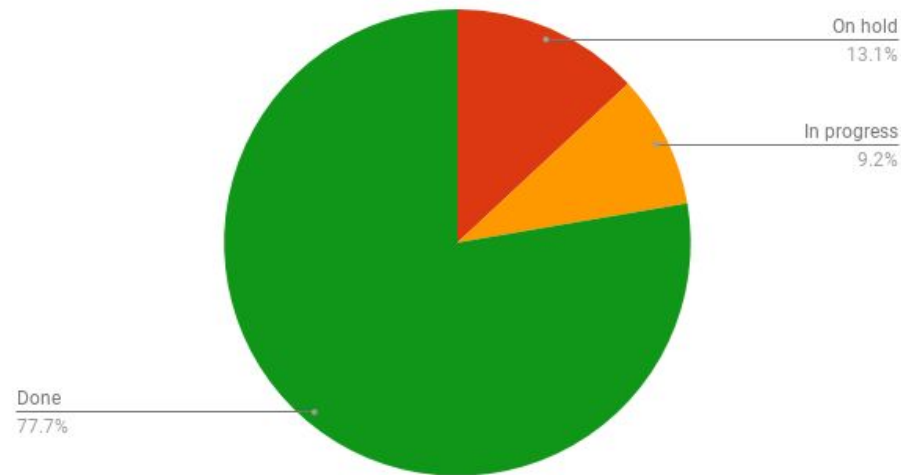
- ~50% of all traffic is IPv6
- From June 2019 onwards

[LINK](#) to these plots

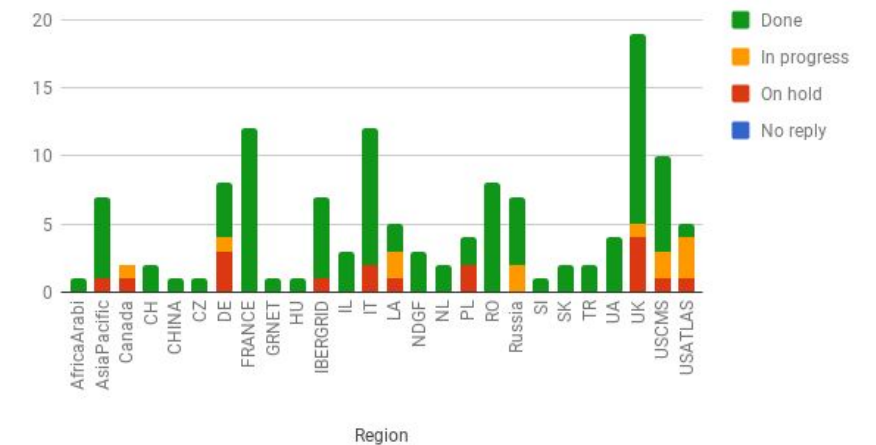
# IPv6 deployment at Tier-2 sites

- The deployment campaign was launched in November 2017
- Steady progress ([status](#))
  - **About 78%** of Tier-2 sites have storage

Tier-2 IPv6 deployment status [12-03-2021]



Tier-2 IPv6 deployment status [12-03-2021]

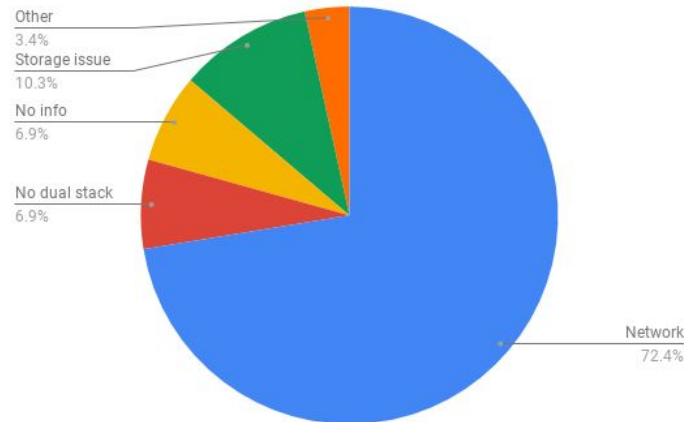


Experiment	Fraction of T2 storage accessible via IPv6
ALICE	86%
ATLAS	62%
CMS	93%
LHCb	75%
Overall	77%

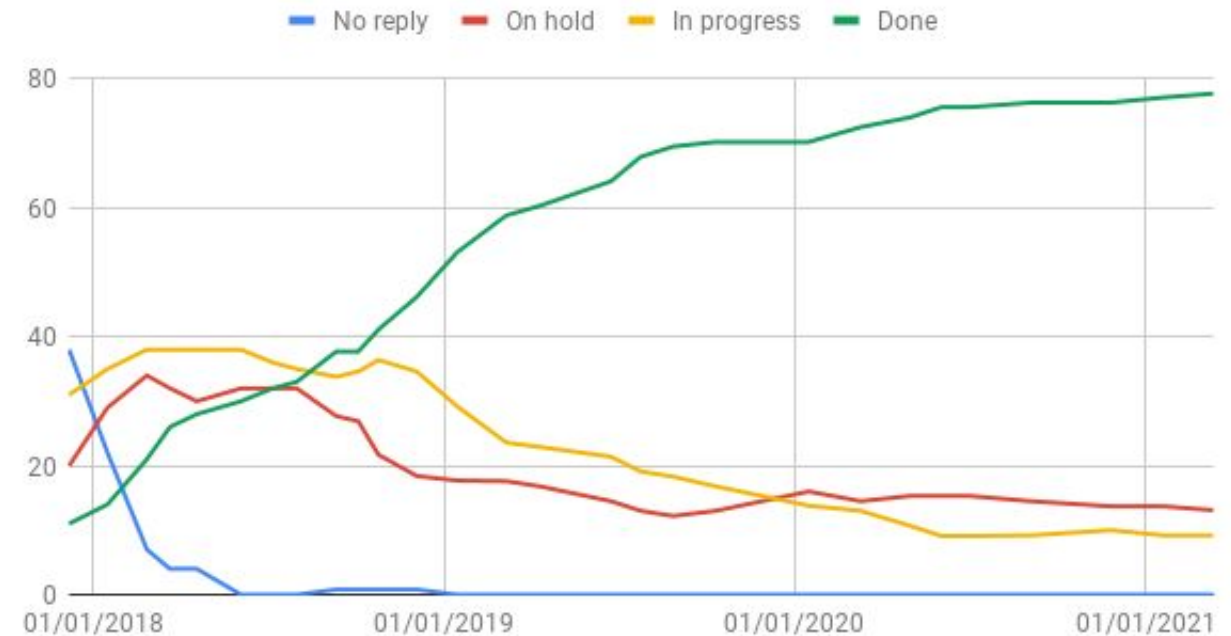
# Tier-2 status (cont'd)

- Progress is slower and slower
  - The cause is that remaining sites are the ones having the most difficulties
    - Typical example: IPv6 not a priority for the campus

Reason of delay [12-03-2021]



Status vs. time





# Data Transfers - FTS transfer monitoring

## Feb-Mar 2021



Approximately **52%** of data transferred via FTS in Feb-Mar 2020 went over IPv6

**But** this includes FTS over DAVS where IPv6 vs IPv4 monitoring is **not available** - all traffic assumed to be IPv4

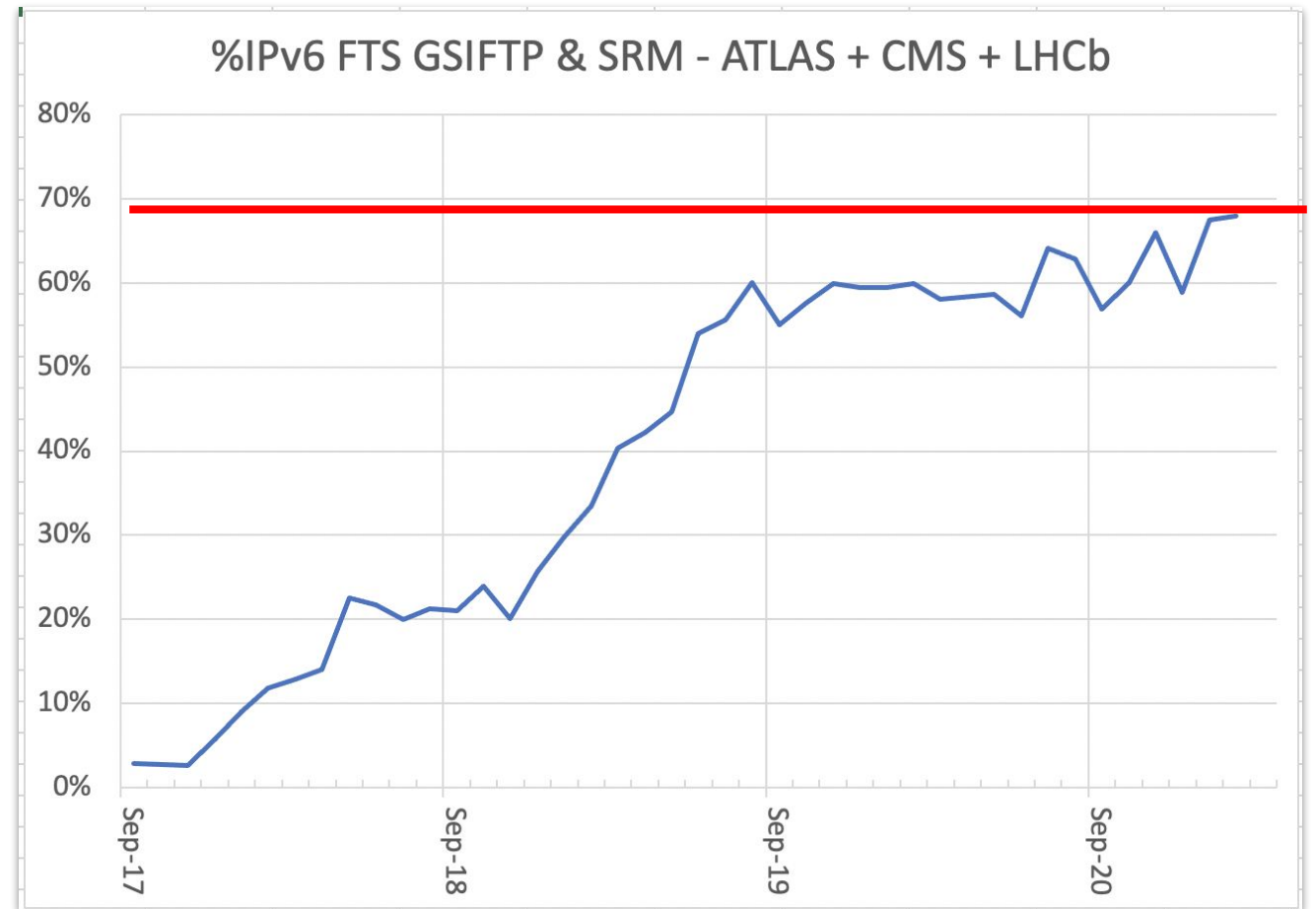


<https://monit-grafana.cern.ch/>

# Fraction of SRM/GridFTP FTS traffic over IPv6



- Some FTS protocols, e.g. DAVS, not yet instrumented to monitor IPv6
- these are excluded from this plot



# Experiment Test Framework (ETF) & IPv6



- ATLAS, CMS **testing from IPv6-only nodes** in pre-production
  - Both **storage and compute** if IPv6 endpoint is available
  - Same tests/metrics used as in production (which is still IPv4-only)
  - Will add a test to **track IPv6 availability on WNs**
  - LHCb also interested
  - ALICE providing its own IPv6 tests directly from Monalisa
- Central instance that can provide a site overview over both IPv4 and IPv6 tests is in pre-production [[link](#)]
  - Few issues still remain to be fixed

## ETF (2)

- Successfully tested IPv6-only submission to **CERN IPv6-only testbed**
  - current ETF job submission works fine over IPv6
- HTTP/webdav tests with IPv6 support available for quite some time
- CMS working on a **new storage probe**
  - to introduce testing of new protocols/features
  - expected to be useful for LHCb and ATLAS
- Plan to move IPv6 testing in production and introduce **IPv6 reporting capability in the WLCG availability/reliability reports**

# IPv6-only networking on WLCG

- Our main use case was and still is: **be ready for use of IPv6-only CPU!**
- But there is a **new reason** for data transfers over IPv6
- Research Networking Technical Working Group (**RNTWG**)
  - The ability to understand WAN network flows is too limited
  - Need new methods to mark and monitor network use
  - **Packet Marking sub-group** started work
  - Requires use of IPv6
  - See Marian's talk on RNTWG

# IPv6-only networking (2)

## Other motivations:

- Running a dual-stack IPv4/IPv6 infrastructure is **complex**
- Large companies (e.g. Facebook, EE/BT) use IPv6-only internally
  - Then use tools like NAT64/DNS64/464XLAT for legacy world
- US Federal Government now requires 80% of its services to be IPv6-only by 2025
- Our proposed plan is to **simplify and move to IPv6-only** in the majority of WLCG services and clients
  - Ongoing support for IPv4-only clients where needed
    - RFC 7755 - SIIT for IPv6-only Data Centres and RFC 7756 extension

# Preparing for an IPv6-only WLCG

- Goal is to test IPv6-only clusters/worker nodes
  - IPv6-only testing at CERN (slides from Ben Jones)
    - JIRA ticket: <https://its.cern.ch/jira/browse/BBC-2820>
    - CMS tests on that cluster
    - Configured not to use IPv4 (Puppet issues)
    - Next step – remove IPv4 completely
- Network traffic analysis (F Prelz)
- FTS changes to monitoring all protocols

# CERN IPv6-only testbed - Ben Jones

## Goals of Pilot

- Can we find current issues with IPv6?
  - We can wait till everything is fixed, or we can just see what's broken
- Gain experience running IPv6 in production
  - Even CERN has some IPv4 pressure
  - Can next DC be IPv6 only
- End-to-end IPv6: submission to execution
  - IPv6 only workers primary goal, but investigating CEs too
- Not necessary to wait for perfect IPv6 infra to make progress

## Limitations / Hacks

- Available networks mean both the workers and CEs are dual stack
  - iptables rules to block & log IPv4
  - ensure we are configured to prefer IPv6
- Various system services require IPv4
  - where we can't configure to use IPv6 we are allowing specific IPv4 traffic
  - one goal is to discover these limitations, but the point is to be able to run jobs



25/02/2021

HEPiX IPv6

4



25/02/2021

HEPiX IPv6

5



# CERN IPv6-only testbed (2)

## Integration with services

Service	Protocol	Issue
Argus	IPv4 Only	Needs configuration
MONIT	IPv4 Only	MONIT alarm infrastructure currently IPv4 (working on dual stack)
OpenStack metadata	IPv4 Only	Support added for IPv6 in Nova (Victoria). Tools like facter will need to be updated
SSSD	IPv4/IPv6	Needs the configuraiton flag <code>lookup_family_order=ipv6_first</code> to pickup IPv6
Chrony	IPv4/IPv6	Works, but not all time servers configured dual stack
BDII	IPv4	BDII queries HTCondor-CE port 9619 via IPv4
fetch-crl	IPv4/IPv6	Needs "inet6glue" config param and dependency
CVMFS	IPv4/IPv6	Needs <code>CVMFS_IPFAMILY_PREFER=6</code>

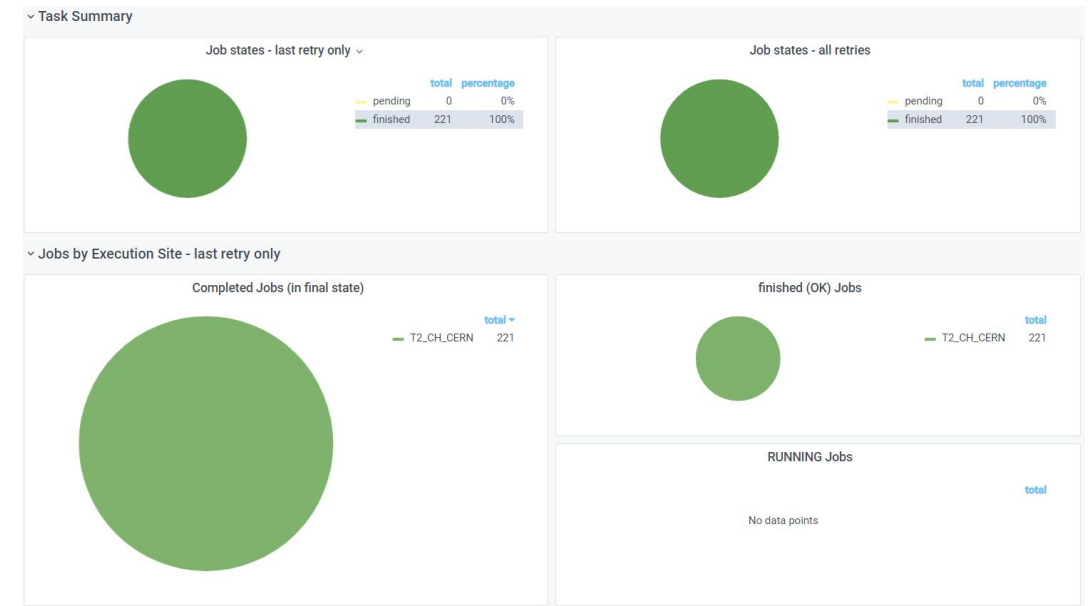
## HTCondor Configuration

- We DENY IPV4 traffic to 9619 (CE) and 9618
- `PREFER_IPV4=False`
  - Needed on CE & on workers
- `ENABLE_IPV4=False` on CE
  - Even with no IPV4 in DNS, the CE MyAddress will have IPV4 address without this setting
  - `_condor_PREFER_IPV4=False` env var for clients can workaround, otherwise global settings on schedds etc

# CMS tests on CERN IPv6-only modes



- Initial hiccups due to network configuration of the involved nodes and a few iterations on the HTCondor configuration
- A HammerCloud test was set up to continuously run analysis-like jobs
  - Success rate is consistently 100%
- The “final” test will be to remove IPv4 altogether from the WNs (not just block it)



# Network traffic analysis - places where protocol translation would fail



- In an **IPV6-only WLCG**, IPv4-only clients will need to use some form of **protocol translation** to connect to IPv6-only services
- Protocol translation is unable to translate IP literal addresses that are handled in the upper stack layers or stored in some database
  - Applications that store, handle or signal literal IPv4 addresses will be broken by protocol translation
- We are running an analysis with INFN and KIT sampled WN traffic to **identify protocols and services** used beyond those we already know
  - Another approach is to use IPv6-only WNs and see what breaks and why
- Then work with experiments and developers to “fix” whatever is broken

# Work to achieve IPv6-only networking on WLCG



- **Fix dual-stack endpoints** that prefer to use IPv4 rather than IPv6
- More **testing of IPv6-only clusters**
- Encourage more sites to **use dual-stack CPU nodes**
  - Many sites have successfully moved to dual-stack worker nodes
- WLCG may need to agree a date for **“end of full support” of IPv4-only clients**
  - e.g. start of LHC Run4?
  - Transition tools such as RFC 7755/6 is one option for connecting IPv4 clients once core is IPv6-only

# Summary



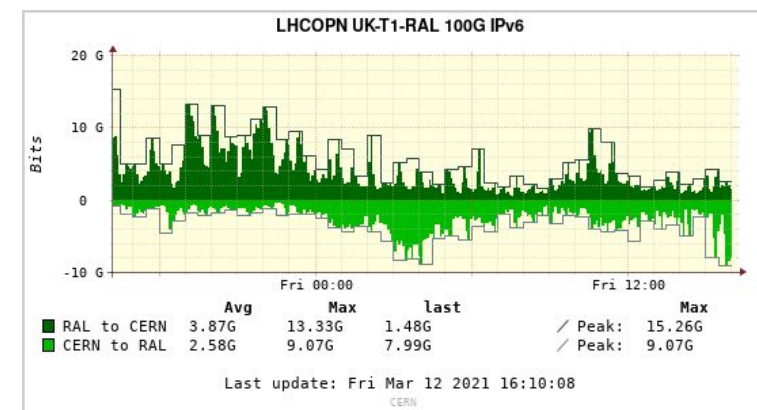
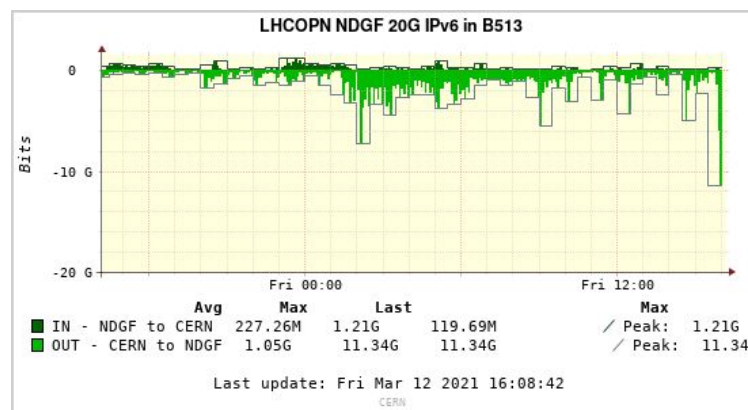
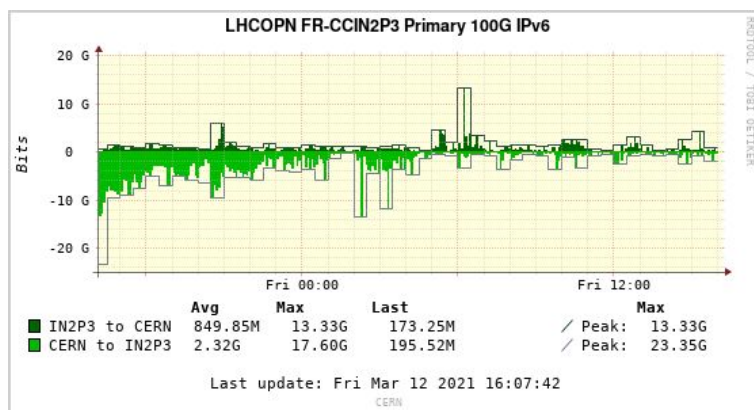
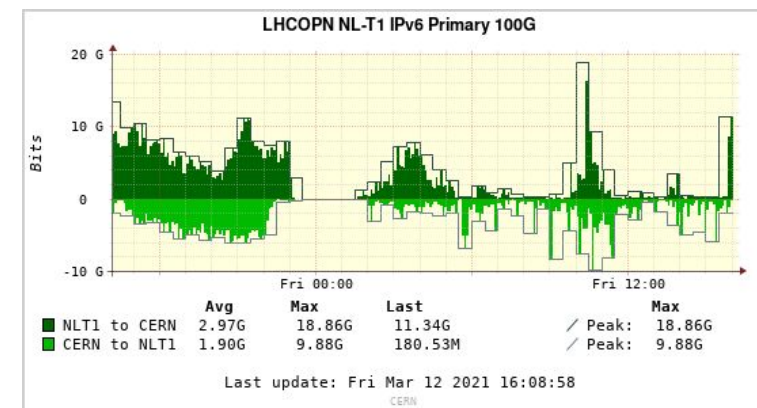
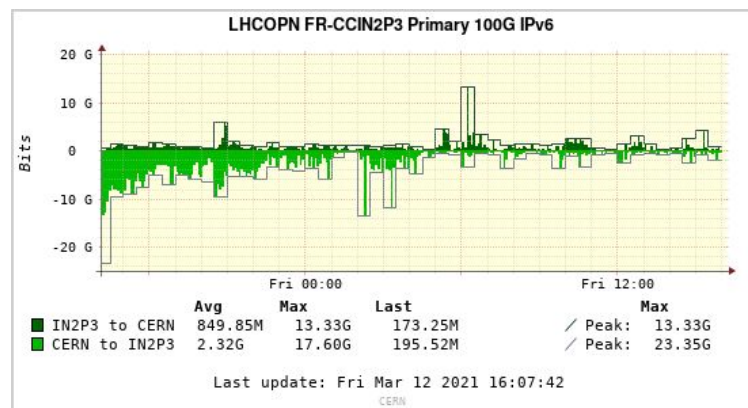
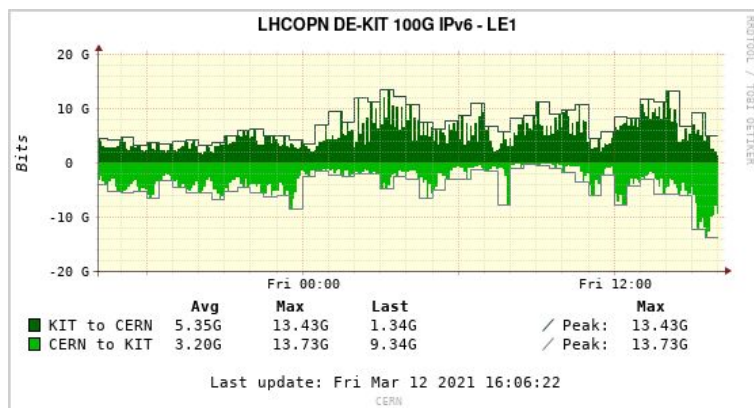
- WLCG is ready to support use of IPv6-only CPU resources
  - Tier-1: 96% of storage is available via IPv6
  - Tier-2: 77% sites & storage are now dual-stack
- ~70% of FTS transfers today over IPv6
  - monitoring is not complete for all protocols
- ~50% LHCOPN+LHCONE traffic observed at CERN is IPv6
  - growth currently “on hold”
- IPv6 WG is now working on the move to **IPv6-only** services
  - Much simpler to manage
  - IPv6 data transfers to help netflow packet marking
- ***Many thanks to all who contribute to this work***

# Questions?

# Backup slides



# LHCOPN IPv6 Traffic





# ATLAS and CMS FTS IPv6 traffic

