



The HEPiX IPv6 working group - an introduction

David Kelsey (STFC UKRI)
HEPiX IPv6 WG meeting, virtual, 14-15 Oct 2021



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), 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), D P Kelsey (RAL), E Martelli (CERN), S McKee (U Michigan), R Nandakumar (RAL/LHCb), K Ohrenberg (DESY), F Prezl (INFN), D Rand (Imperial), A Sciabà (CERN/CMS)
- 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

- History (Phases 1 and 2)
- Current status (see later agenda items for more details)
- The future - Phase 3: IPv6-only networking
- Summary

Worldwide LHC Computing Grid (WLCG)

- The WLCG is a global collaboration
- more than 170 computing centres in 42 countries
- Its mission is to **store, distribute and analyse** the data generated by the LHC experiments
- Sites hierarchically arranged with three tiers:
 - Tier-0 at CERN
 - 13 Tier-1s (mainly national laboratories)
 - >150 Tier-2s (generally university physics laboratories)

WLCG sites



★ Tier-0 📍 Tier-1 📍 Tier-2

Why should WLCG use IPv6?

Why IPv6?

- Survey of 18 major HEP sites (Sep 2010) – IPv6 readiness
 - National NRENs ready, Universities and Labs not ready
 - Some reported lack of IPv4 address space, including CERN
- HEPiX meeting – Cornell, Ithaca NY – Nov 2010
 - Projected IANA IPv4 address exhaustion
 - Sep 2010 – memo from US Federal CIO to all Exec depts (incl DOE)
- Offers of opportunistic CPU resources which could be IPv6-only
 - Experiments want to be able to use them
- Recognition that much of our middleware, software and technology was **not yet IPv6** capable
- HEPiX decided to create a working group (started April 2011)
 - No specific funding – but motivated, competent volunteers!

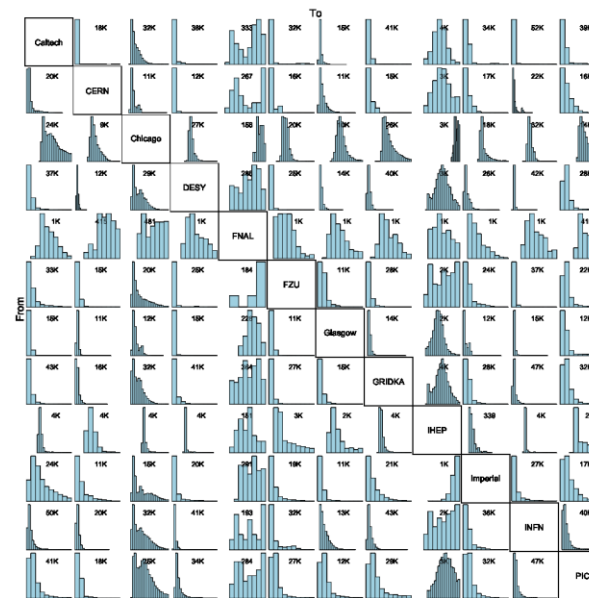
Phase 1: 2011-2016

- Phase 1 – **full analysis of work to be done**
 - Applications, system and network tools, operational security
 - Create and operate a distributed test-bed
 - No interference with WLCG production data analysis!
 - Propose timetable and plan for transition
- CERN announces shortages of routable IPv4 addresses
 - explosion of virtualisation
- Active HEPiX IPv6 test-bed with ~ 12 sites
 - engagement of all 4 LHC experiments
- Testing regular data transfers across the testbed
- Testing dual-stack services (in production) at Imperial College London
- Concluded not able to support IPv6-only clients until **at least 2014**

At CHEP2013 conference

- > 2 PB data transferred over IPv6 in last 6 months
- Success rate > 87%
- Very High!

GridFTP IPv6 data transfer mesh



2013-14 Data Management

- Testing the important data transfer protocols, technology and data storage/file systems
 - For IPv6-readiness
- GridFTP, 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!**

Phase 2 – Dual-stack deployment – 2016-20

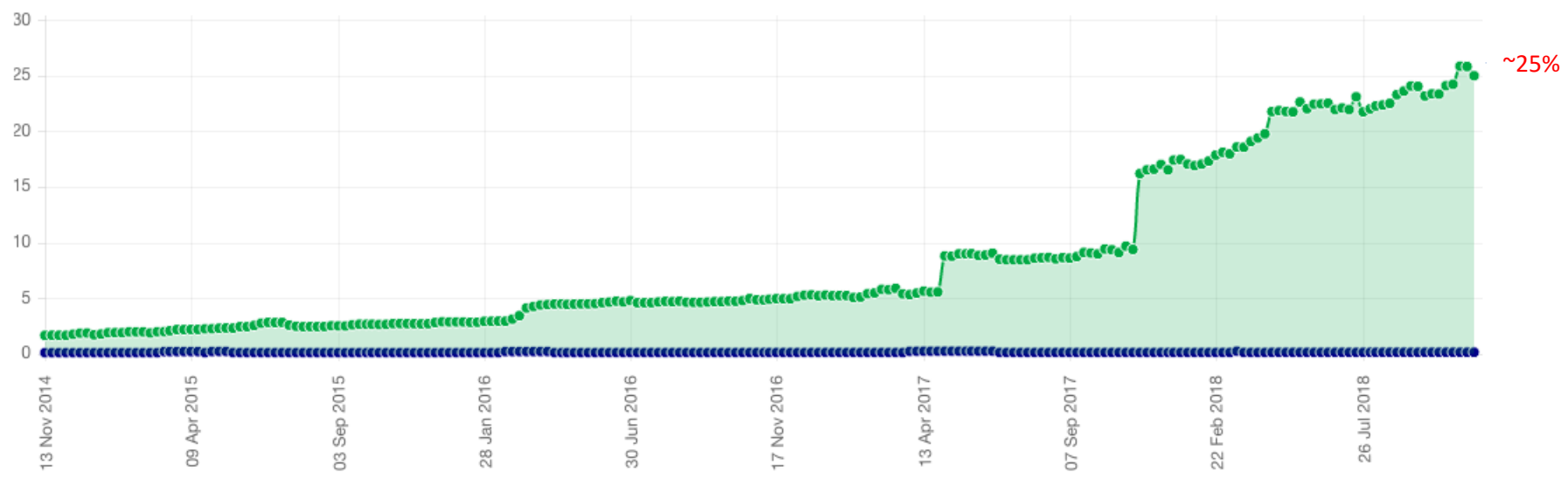
- Most important IPv6 use case
 - Sites, Clouds providing CPU (virtual machines)
 - Opportunistic resources may be IPv6-only
 - **Need dual-stack federated storage services**
 - And dual-stack central WLCG and Experiment services
- LHCOPN (Tier0-Tier1 private network)
 - Encourage IPv6 peering everywhere
- perfSONAR – end to end network monitoring – dual-stack
- Move central services and central monitoring to IPv4/IPv6
- Wrote guidance on **IPv6 security for WLCG sites**
- Deployment timetable approved by WLCG Management Board (Sep 2016)

WLCG – IPv6 deployment

Plan approved by WLCG Management Board

- **April 2017** – support for IPv6-only CPU starts
 - Tier-1s to provide dual-stack storage (in testbed)
- **April 2018**
 - Tier-1 dual-stack storage in production mode
- By end of LHC Run 2 (**end 2018**)
 - A large number of Tier-2s to migrate storage to IPv6
 - All requested to do this

Growth of dual-stack hosts in the WLCG (Jan 2019)



- Percentage of IPv6-only endpoints
- Percentage of dual-stack endpoints

All services, not just storage

Fraction of endpoints listed in the CERN central BDII (lcg-bdii.cern.ch) where the DNS returns a dual-stack IPv6-IPv4 (A+AAAA) resolution (green line) or an IPv6-only resolution (blue line).
http://orsone.mi.infn.it/~prelz/ipv6_bdii/.

Problems & lessons learned

- Many blocking issues outside of our own control
 - Both software and site networking teams
- Developers claim that software is fully IPv6-compliant!
- Software/protocols fixed-size storage for IP addresses
- Software/protocols assume single address (as in IPv4)
- Performance differences between IPv4 & IPv6
 - IPv6 must perform at least as well
- Have to understand cases where fraction of IPv6 is smaller than expected
 - Preference for IPv6 over IPv4 must be established
- Can be lots of development effort and testing is not easy when no other positive change re functionality
- Sys admins, operations staff, security team, developers
 - All need TRAINING and experience

IPv4/IPv6 and Tier-1 storage

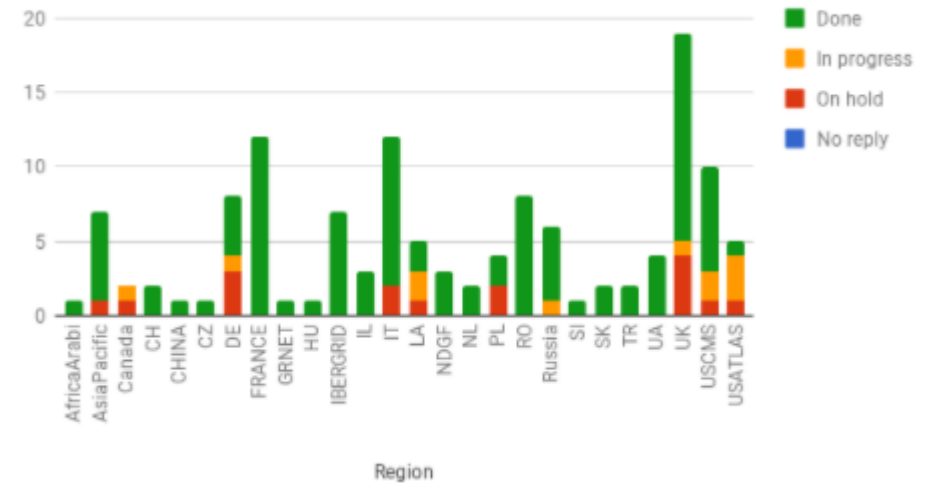
- Sep 2021 (GDB)
- FTS servers at BNL, CERN, Fermilab and RAL all now work in dual stack
- FTS transfers do happen via IPv6 at most Tier-1s (but not all transfers)
- IPv6 transfers do not (yet) happen at RRC-KI but IPv6 is enabled

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

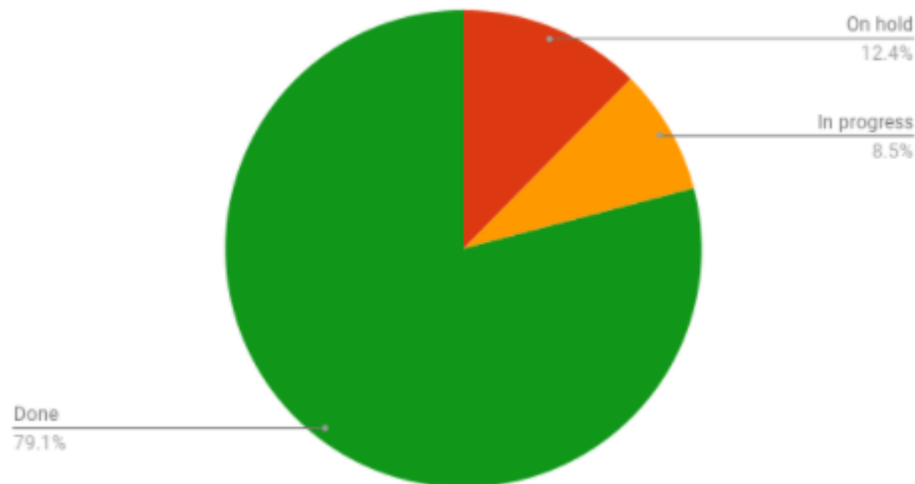
IPv4/IPv6 deployment at Tier-2 sites

- The deployment campaign was launched in November 2017
- Steady progress ([status](#))
 - **~79%** of Tier-2 sites have storage on dual stack

Tier-2 IPv6 deployment status [07-09-2021]



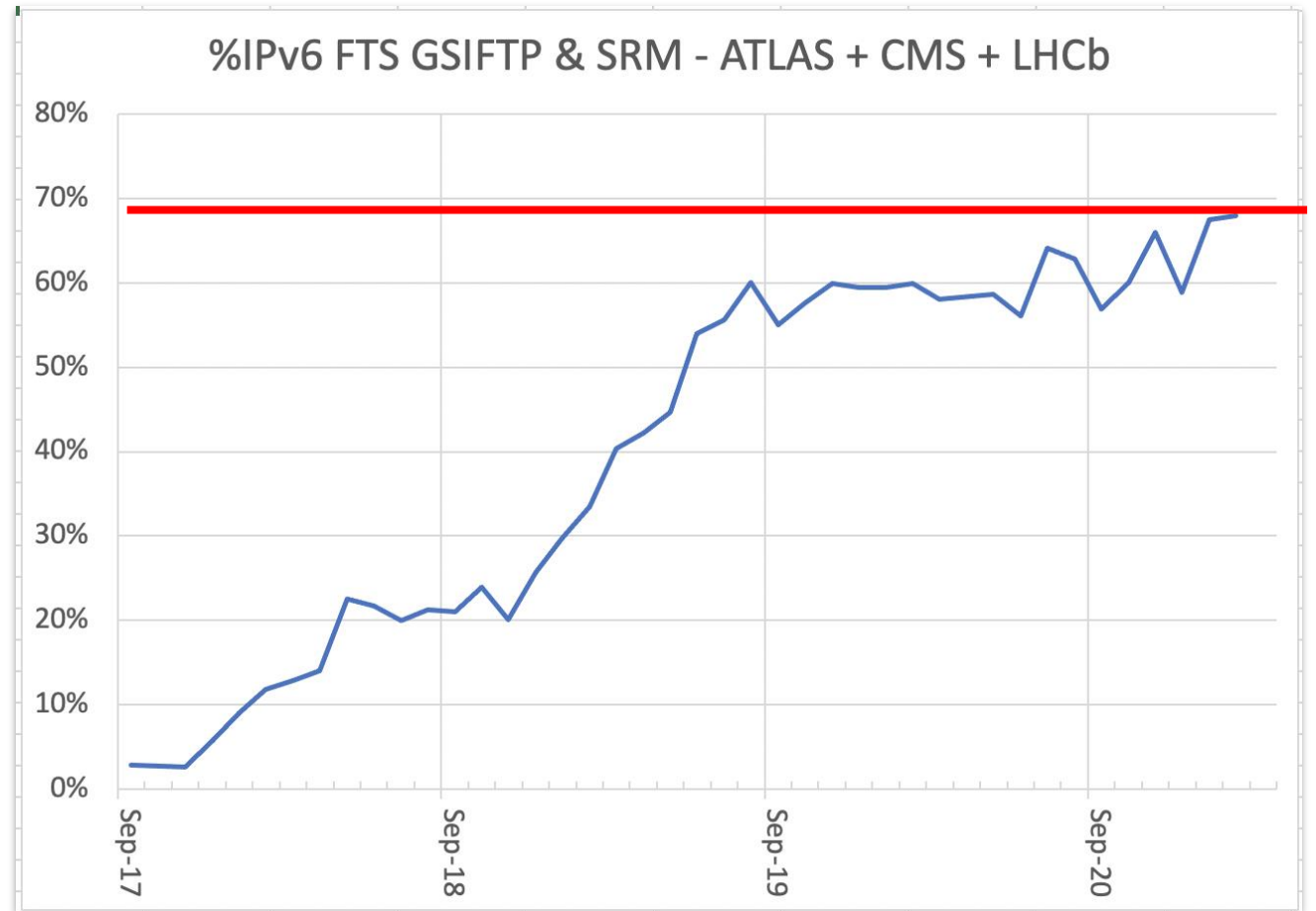
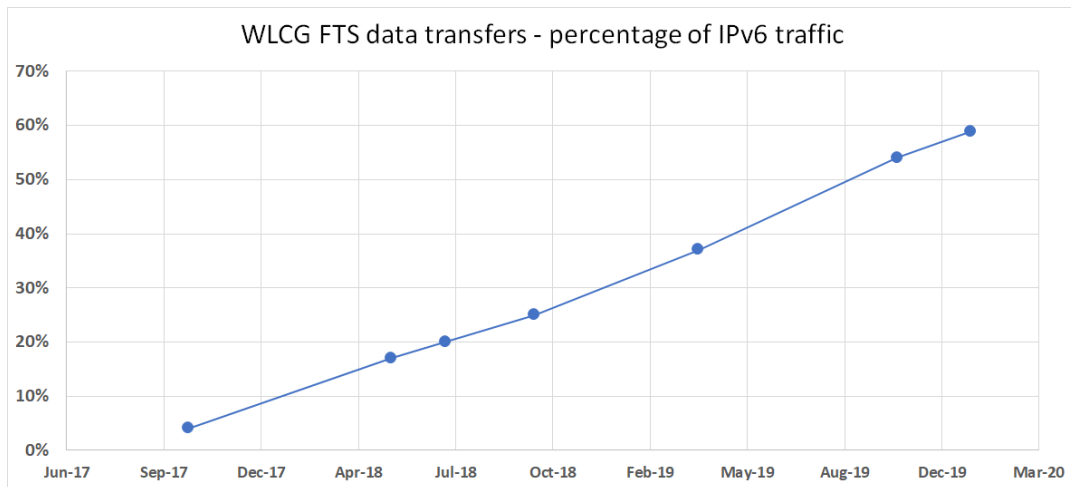
Tier-2 IPv6 deployment status [07-09-2021]



Experiment	Fraction of T2 storage accessible via IPv6
ALICE	87%
ATLAS	62%
CMS	94%
LHCb	75%
Overall	77%

Fraction of SRM/GridFTP FTS traffic over IPv6

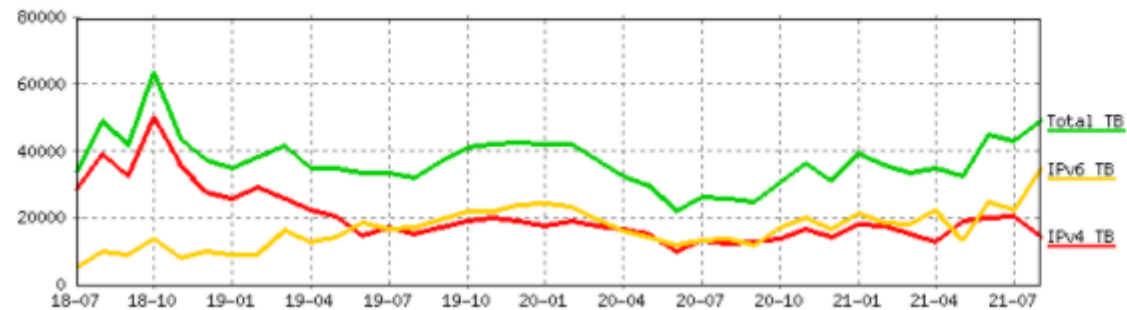
- Some FTS protocols, e.g. DAVS, not yet instrumented to monitor IPv6
- these are excluded from this plot



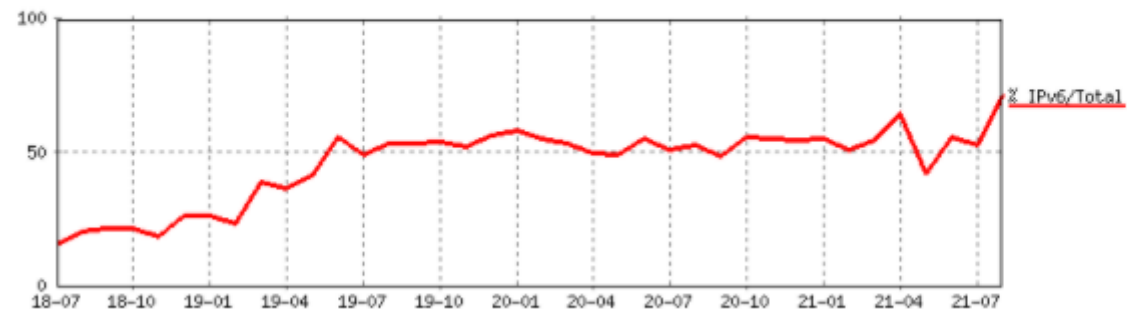
IPv6 traffic on LHCOPN/LHCONE at CERN

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/ONE as seen at CERN

- ~70% of all traffic is IPv6
- > 50% from June 2019 onwards

[LINK](#) to these plots

IPv4 transfers on LHCOPN/ONE?

- Tier1s are dual-stack, but IPv4 can still be 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
 - Investigation of ATLAS/CMS IPv6 traffic has been proposed in UK
 - transfers are to/from WN's and the WNs are still IPv4-only
 - Difficult to monitor WN IPv6 traffic centrally – can sites monitor?
- **encourage all** sites to deploy all WNs and all services as dual-stack
- **encourage all** sites and all experiments to "prefer" IPv6 transfers

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
 - See WLCG GDB May 2021 – [LHCOPN/ONE \(E. Martelli\)](#) slide 15
 - And report today from Marian Babik
 - US Federal Government - 80% of services to be IPv6-only by 2025-26
 - See report today from Phil DeMar

Phase 3: 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
 - e..g. via use of RFC 7755/7756
 - Stateless IP/ICMP Translation for IPv6-only Data Centres
- Timetable to be defined

IPv6-only timetable

- A possible first place to turn-off IPv4?
 - LHCOPN (Tier0 and all Tier1's are all dual-stack)
 - When the amount of IPv4 traffic becomes very small
 - Can we make LHCOPN IPv6-only?
 - IPv4 traffic will flow via general internet links
 - WLCG Management Board (July 2021) – “must not happen too early”
- **WLCG IPv6-only plan** - is during LS3 and before start of Run 4 possible?
 - MB still to agree date for end of “full support” for IPv4 on WLCG Clients
 - Core services at that point can become IPv6-only

WLCG MB Conclusions (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 agreed to be IPv6-only
 - But the timetable to be agreed later!

Summary



- WLCG is ready to support use of IPv6-only CPU resources
 - **Good steady progress towards this goal!**
- Tier-1s should all have production storage accessible over IPv6
 - >96% of Tier-1 storage is available via IPv6
- Tier-2s ~80% sites done
 - >75% of Tier-2 storage is dual-stack
- WG Phase 3 - we are planning for move to IPv6-only services
- ***message to new research communities - build on IPv6 from start***



Questions, Discussion?