



IPv6 & WLCG – an update from the HEPiX IPv6 working group

David Kelsey (STFC UKRI), Andrea Sciabà (CERN)

WLCG GDB @ ISGC2019, Taipei

(Also presented at HEPiX, UCSD, 26 Mar 2019)



On behalf of all colleagues in the HEPiX IPv6 working group

Active in HEPiX IPv6 Working Group – last 6 months

- M Babik (CERN), M Bly (RAL), T Chown (Jisc), J Chudoba (Prague), C Condurache (RAL), T Finnern (DESY), C Grigoras (CERN/ALICE), B Hoeft (KIT), D P Kelsey (RAL), F López Muñoz (PIC), E Martelli (CERN), A Manzi (CERN), R Nandakumar (RAL/LHCb), K Ohrenberg (DESY), F Prelz (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...

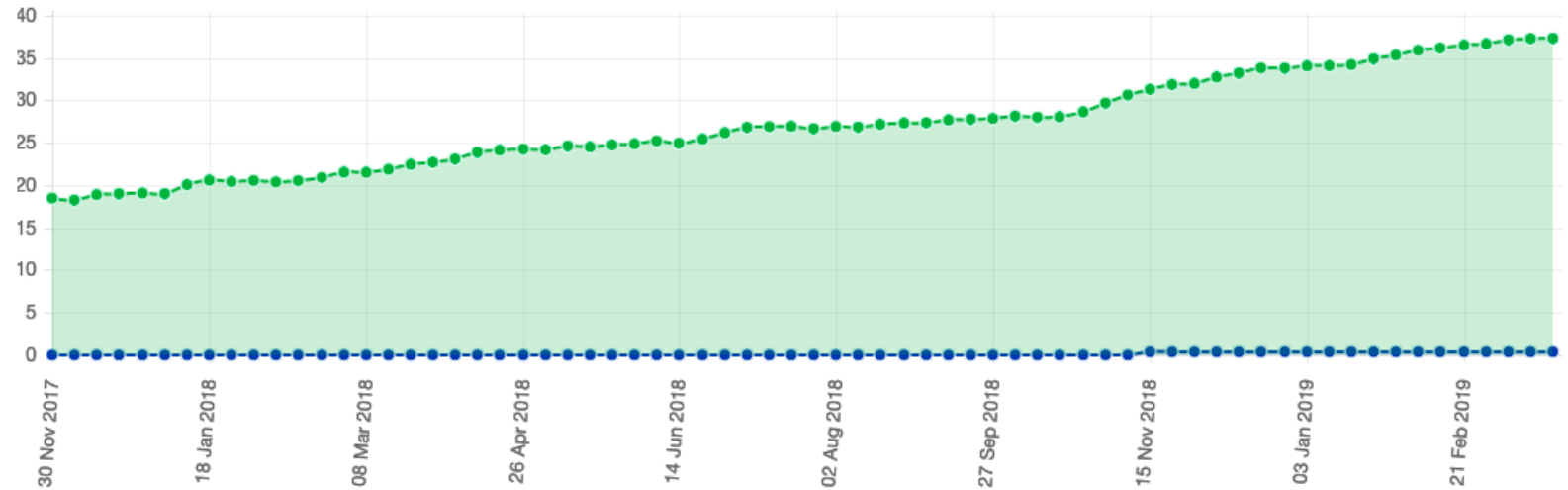
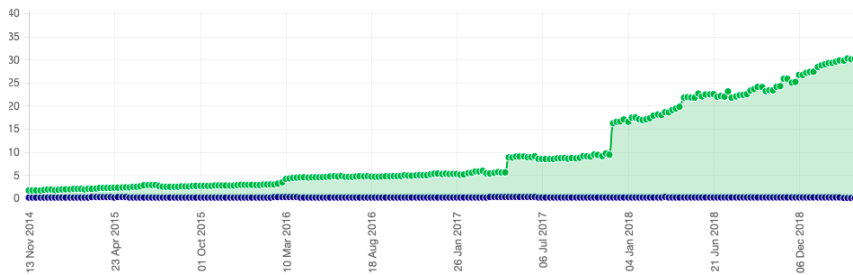
Outline

Since HEPiX in Barcelona (8-12 Oct 2018)

- Finalised our CHEP2018 paper – **“IPv6 in production: its deployment and usage in WLCG”**
 - One F2F meeting at CERN (24/25 Jan 2019)
 - 3 meetings by Vidyo
-
- WLCG dual-stack service endpoint tracking
 - Tier-0/Tier-1/LHCOPN/LHCONE status
 - Tier-2 status
 - FTS transfers
 - Monitoring
 - Work in progress
 - Summary

WLCG services status (dual-stack)

results from the LHC Expts VO feeds
(http://orsono.mi.infn.it/~prelz/ipv6_vofeed/) (~37%)



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://orsono.mi.infn.it/~prelz/ipv6_bdii/).

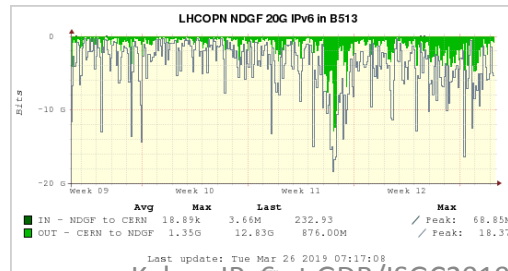
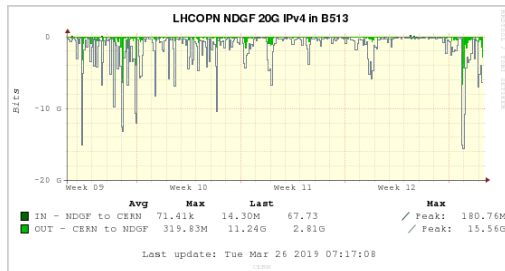
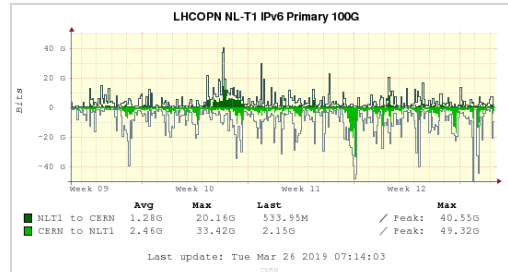
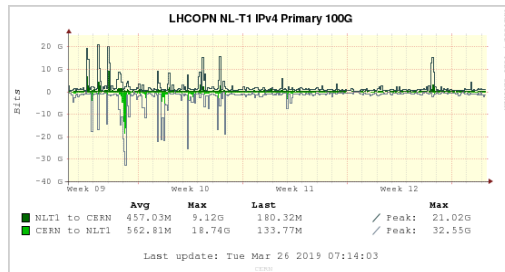
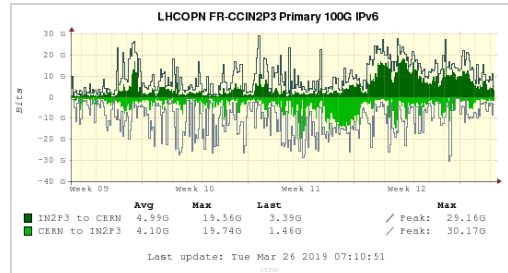
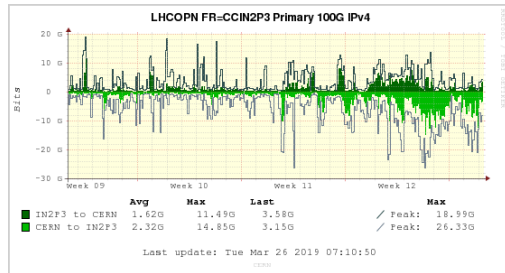
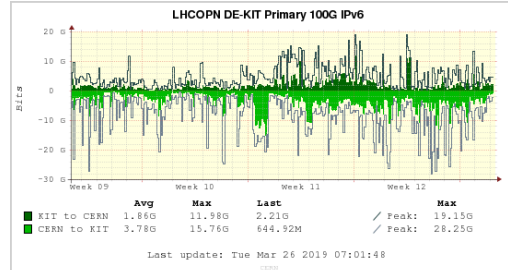
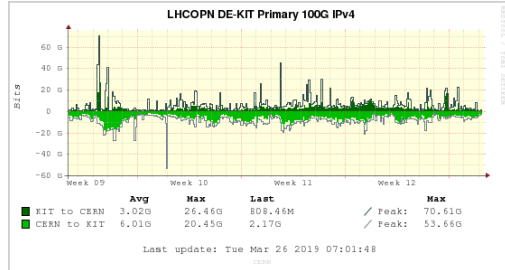
Tier-0/Tier-1/LHCOPN/LHCONE status (Bruno Hoeft)

Network and pS at Tier-1's

- All sites connected to LHCONE except RRC-KI
- TRIUMF, CERN, KISTI and RRC-KI still have some problem with IPv6 on pS

Tier-1	LHCOPN	LHCONE	IPv6 Perfsonar
CA-TRIUMF	OK	OK	??
CH-CERN (Tier-0)	OK	OK	??
DE-KIT	OK	OK	OK
ES-PIC	OK	OK	OK
FR-CCIN2P3	OK	OK	OK
IT-INFN-CNAF	OK	OK	OK
KR-KISTI	OK	OK	??
NGDF	OK	OK	OK
NL-T1 - NIKHEF	OK	OK	NO
NL-T1 - Sara-Matrix	OK	OK	OK
RRC-KI-T1	OK	NO	??
RRCC-JINR-T1	OK	OK	OK
TW-ASGC	OK	OK	OK
UK-T1-RAL	OK	OK	OK
US-T1-BNL	OK	OK	OK
US-T1-FNAL	OK	OK	OK

IPv6 traffic to selected T1's



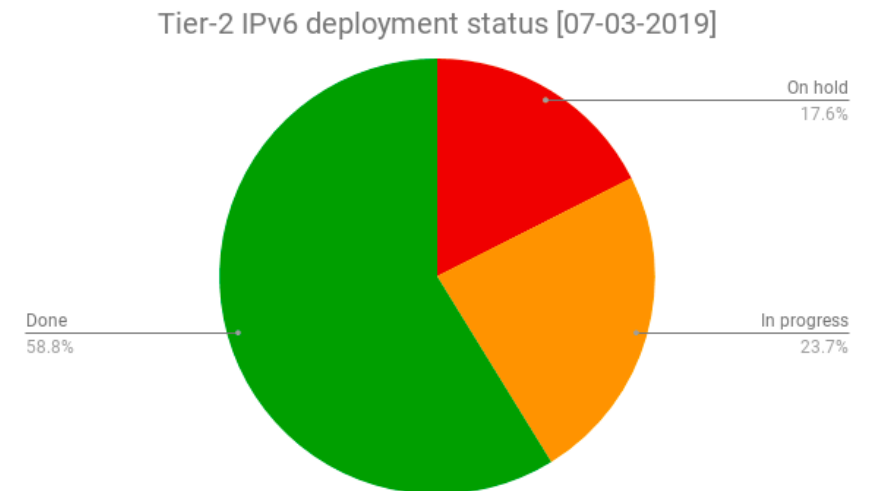
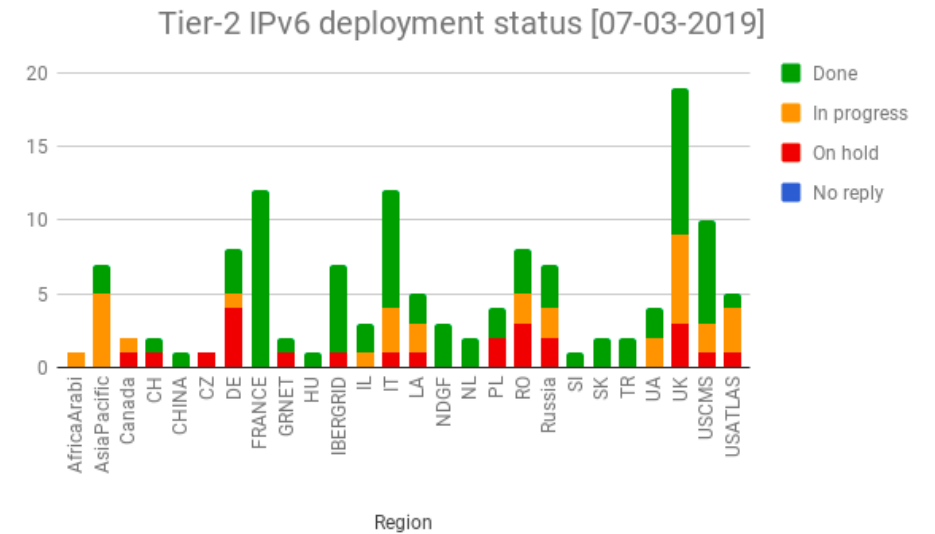
Link	IPv4	IPv6
KIT→CERN	62%	38%
CERN→KIT	61%	39%
IN2P3→CERN	25%	75%
CERN→IN2P3	36%	64%
NL-T1→CERN	26%	74%
CERN→NL-T1	19%	81%
NDGF→CERN	79%	21%
CERN→NDGF	19%	81%

IPv6 on FTS and at Tier-1's

- FTS servers at BNL, CERN and RAL work in dual stack, while at FNAL they use IPv4 only
- GridFTP transfers happen also via IPv6 at
 - IN2P3, JINR, NDGF, RAL, SARA-MATRIX, NIKHEF, CNAF, ASGC, PIC, TRIUMF, KIT
- IPv6 transfers do not happen at
 - FNAL, RRC-KI
- Fraction of disk storage on IPv6:
 - ALICE: 73%
 - ATLAS: 96%
 - CMS: 71%
 - LHCb: 94%
 - All VOs: 86%

Tier-2s: GGUS tickets to all Tier-2 sites

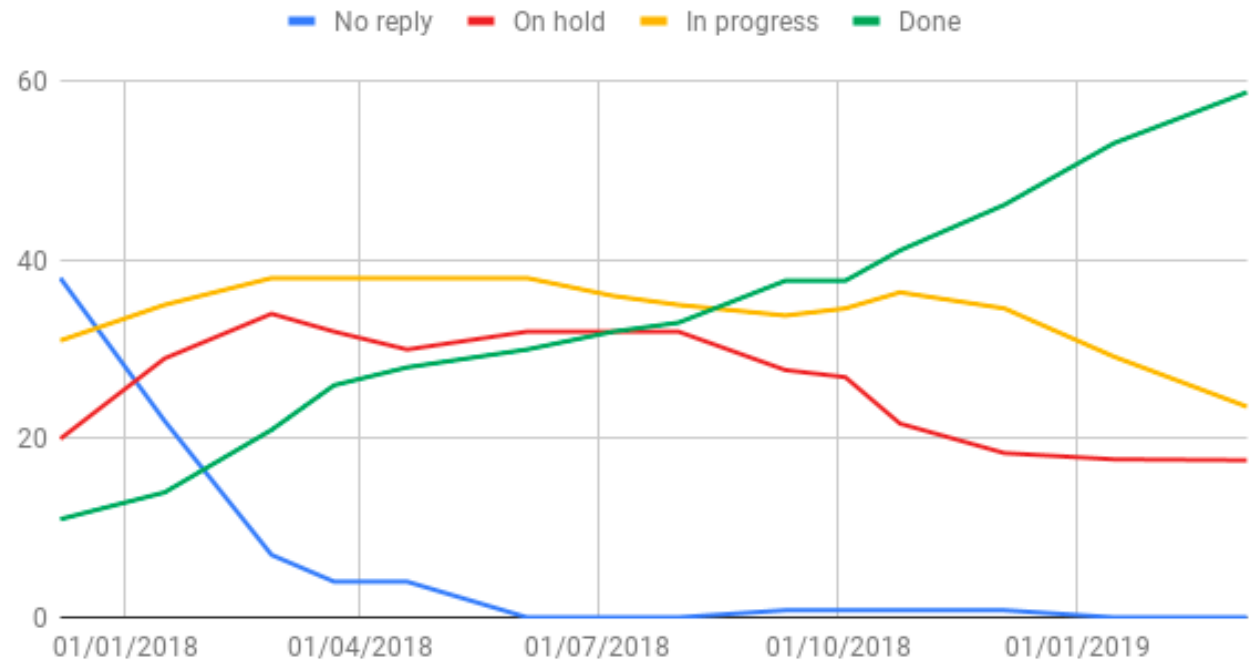
- WLCG set a target for end 2018 for deploying IPv6 on storage systems (and perfSONAR)
- The deployment campaign was launched in November 2017
 - GGUS tickets sent to all non-US sites
 - US sites are tracked via the experiments
 - Sites made aware of the WLCG plans and asked to report plans and give updates
- Steady progress ([status](#))
 - About 59% of T2 sites have storage on dual stack



Tier 2 status (cont'd)

Experiment	Fraction of T2 storage accessible via IPv6
ALICE	64%
ATLAS	46%
CMS	83%
LHCb	63%
Overall	62%

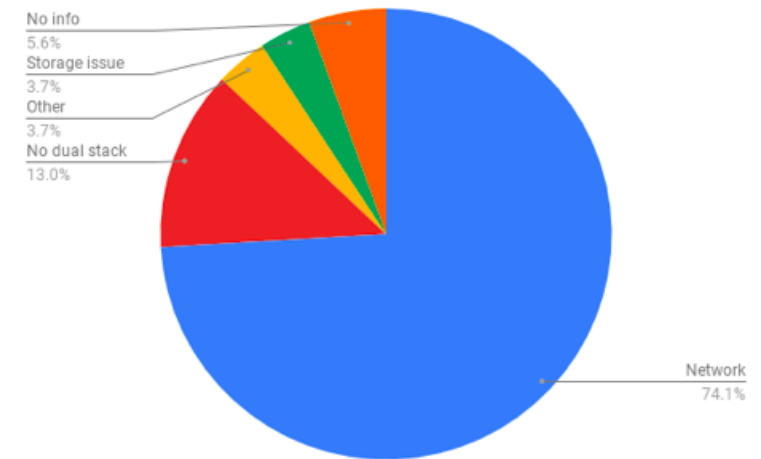
Status vs. time



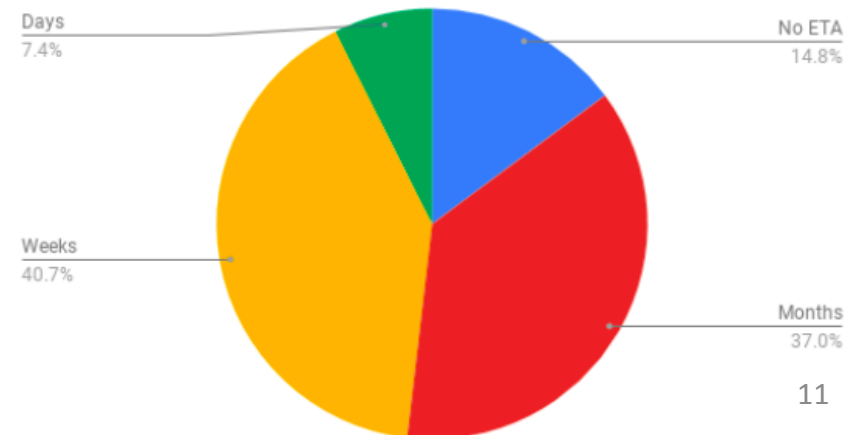
Tier-2 status: reasons for delays

- Checked why sites are late in their IPv6 deployment and the expected ETA
 - **Network**: waiting for interventions on the site infrastructure or network issues to be fixed
 - **No dual stack**: infrastructure network is fine but IPv6 needs enabling on storage or perfSONAR
 - **pS/storage issues**: encountered problems that need fixing
 - **Other**: other types of problems (e.g. the site needs relocating)
 - **No info**: no meaningful information given by the site on their plans

Reason of delay [07-03-2019]



Time scale for deployment [07-03-2019]



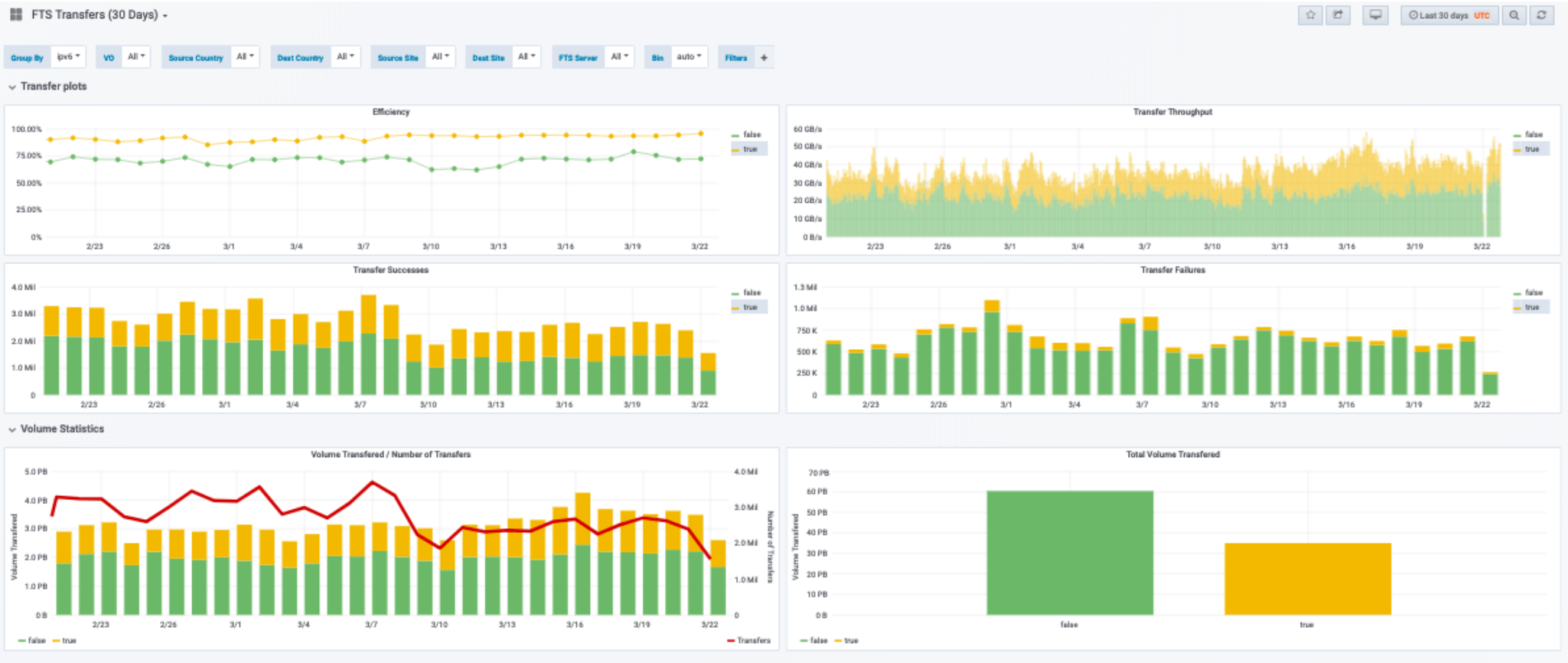
T2 deployment observations

- Tentative extrapolation to the future
 - Within days: 62%
 - Within weeks: 79%
 - Within months: 94%
- Most sites are responsive and provide detailed information
 - For some however regular pinging is a must...
- Several sites must wait for their campus infrastructure to become IPv6-ready
- It is evident that IPv6 is being deployed or will soon be deployed at the vast majority of the remaining sites and that the WLCG request is taken seriously

FTS – transfers over IPv6
(xRootD transfers not yet instrumented
to track IPv6 vs IPv4)

FTS transfer monitoring

Approximately 37% of data transferred via FTS in the last 30 days went over IPv6



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



Monitoring (Marian Babik)

ETF News

- Experiments instances ready
 - CMS and LHCb IPv6 instances ready, currently dual-stack
 - ATLAS has an IPv6-only instance, CMS and LHCb will soon follow
- Other news
 - protoDUNE plans to use ETF to monitor its infrastructure
 - CMS/DODAS project and FNAL HEPCloud also being discussed; CMS/DODAS is working in QA (it's VAC-model; no remote API; jobs are pulled from ETF HTCondor pool)
 - ETF is being migrated to Docker and Kubernetes

perfSONAR news

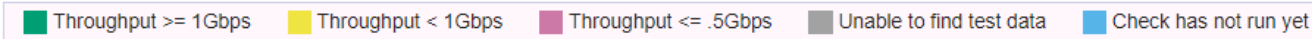
Marian Babik & Duncan Rand

- Campaign to update all perfSONARs to CC7 and 4.1 ongoing
 - Contacting sites directly (will issue tickets if necessary)
 - SL6 unsupported since Q4 2018
- CERN new perfSONAR servers
 - 2 servers with 40 Gbps/100 Gbps
- Good progress with maddash
 - Solved most pS problems at Tier-1s
 - Found some bugs, already fixed by developers
- IRIS-HEP started
 - Has tasks/objectives related to perfSONAR

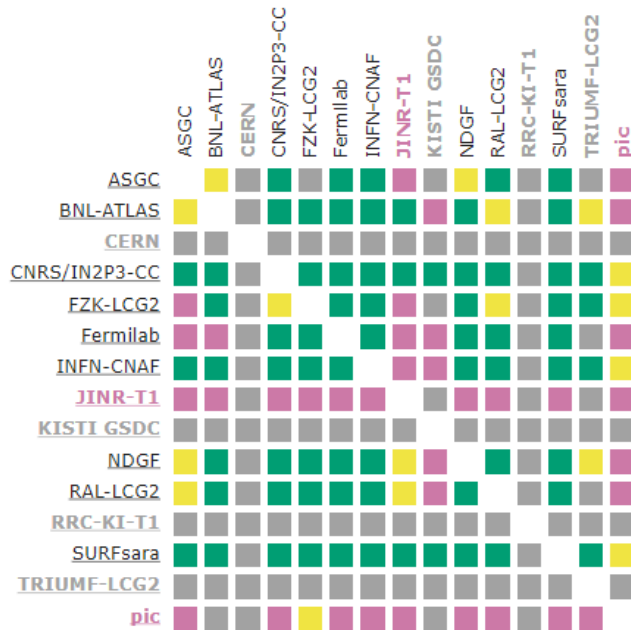
perfSONAR IPv6 Mesh



OPN Mesh Config - OPN IPv6 Bandwidth - Throughput



⚠ Found a total of 8 problems involving 6 hosts in the grid



- Re-configured as was agreed
 - IPv6 pS tests now included in all meshes
 - Throughput and traceroute tests always for both IPv4 and IPv6, latency tests for either depending on what's available on both ends
 - Dual-stack mesh retired
- New monitoring is in place which reports test “efficiency” wrt IPv4 and IPv6
 - Efficiency = number of working destinations / total destinations
 - For each perfSONAR node and each mesh, reports missing results in the perfSONAR’s local measurement archive ([link](#))

Work in progress

- PIC and FTS team – investigating why transfers made over IPv4 when both ends are IPv6?
 - Lots of places where the IPv6 versus IPv4 preference is configured
- ATLAS and CMS are testing IPv6-only worker nodes
 - ATLAS: SiGNET (Slovenia) – production work on IPv6-only WN – works well
 - CMS: Nebraska T2 – See Garhan’s talk

IPv6 deployment on WLCG - Summary



- WLCG to support use of IPv6-only CPU resources
- **Good steady progress towards this goal!**
- Tier-1s should already have production storage accessible over IPv6
 - > 85% of data is available via IPv6
- Tier-2s 59% sites done
 - 62% of Tier-2 storage is dual-stack
- 37% of FTS transfers today over IPv6
 - No monitoring yet for xRootD
- >50% perfSONAR hosts now reporting IPv6-enabled
- There are still some issues to fix and/or understand



Questions?



Backup slides

WLCG deployment plan: timeline

- By April 1st 2017
 - – Sites can provide IPv6-only CPUs if necessary
 - – Tier-1's must provide dual-stack storage access with sufficient performance and reliability
 - At least in a testbed setup
 - – Stratum-1 service at CERN must be dual-stack
 - – A dedicated ETF infrastructure to test IPv6 services must be available
 - – ATLAS and CMS must deploy all services interacting with WNs in dual-stack
 - – All the above, without disrupting normal WLCG operations
- By April 1st 2018
 - – Tier-1's must provide dual-stack storage access in production with increased performance and reliability
 - – Tier-1's must upgrade their Stratum-1 and FTS to dual-stack
 - – The official ETF infrastructure must be migrated to dual-stack
 - – GOCDB, OIM, GGUS, BDII should be dual-stack
- By end of Run2
 - – A large number of sites will have migrated their storage to IPv6
 - – The recommendation to keep IPv4 as a backup will be dropped

T2 tracking page

<https://twiki.cern.ch/twiki/bin/view/LCG/Wlclpv6>

Site	Region	ALICE	ATLAS	CMS	LHCb	Status	perfSONAR	Storage	Ticket	Details
UKI-GridPP-Cloud-IC	UK				Y	Done	NA	NA	GGUS:131599	The site is an extension of UKI-LT2-IC-HEP, has not pS or storage and all services are IPv6-enabled
UKI-LT2-Brunel	UK		Y	Y	Y	Done	NA	Tested	GGUS:131600	Dual stack on all services since years
UKI-LT2-IC-HEP	UK		Y	Y	Y	Done	NA	Tested	GGUS:131601	pS not deployed by choice of the site
UKI-LT2-QMUL	UK		Y	Y	Y	Done	Dual stack	Tested	GGUS:131602	
UKI-LT2-RHUL	UK		Y	Y	Y	In progress	Dual stack	IPv4	GGUS:131603	Still waiting for the central IT service to provide DNS, which is outsourced to JANET
UKI-LT2-UCL-HEP	UK		Y			Done	Dual stack	NA	GGUS:131604	
UKI-NORTHGRID-LANCS-HEP	UK		Y		Y	Done	Dual stack	Tested	GGUS:131605	
UKI-NORTHGRID-LIV-HEP	UK		Y		Y	On hold	Dual stack	IPv4	GGUS:131606	Unable to enable IPv6 on storage for fears of overloading routers and firewalls; need to wait for the upgrades planned for the next financial year
UKI-NORTHGRID-MAN-HEP	UK		Y		Y	In progress	Dual stack	IPv4	GGUS:131607	pS OK; storage: no ETA yet. We got another range of IPv6 addresses, now putting it to production
UKI-NORTHGRID-SHEF-HEP	UK		Y		Y	In progress	IPv4	IPv4	GGUS:131608	IPv6 available at site, will deploy on pS mid November
UKI-SCOTGRID-DURHAM	UK		Y		Y	On hold	Dual stack	IPv4	GGUS:131609	No reverse IPv6 DNS; ETA mid-2019 at the earliest

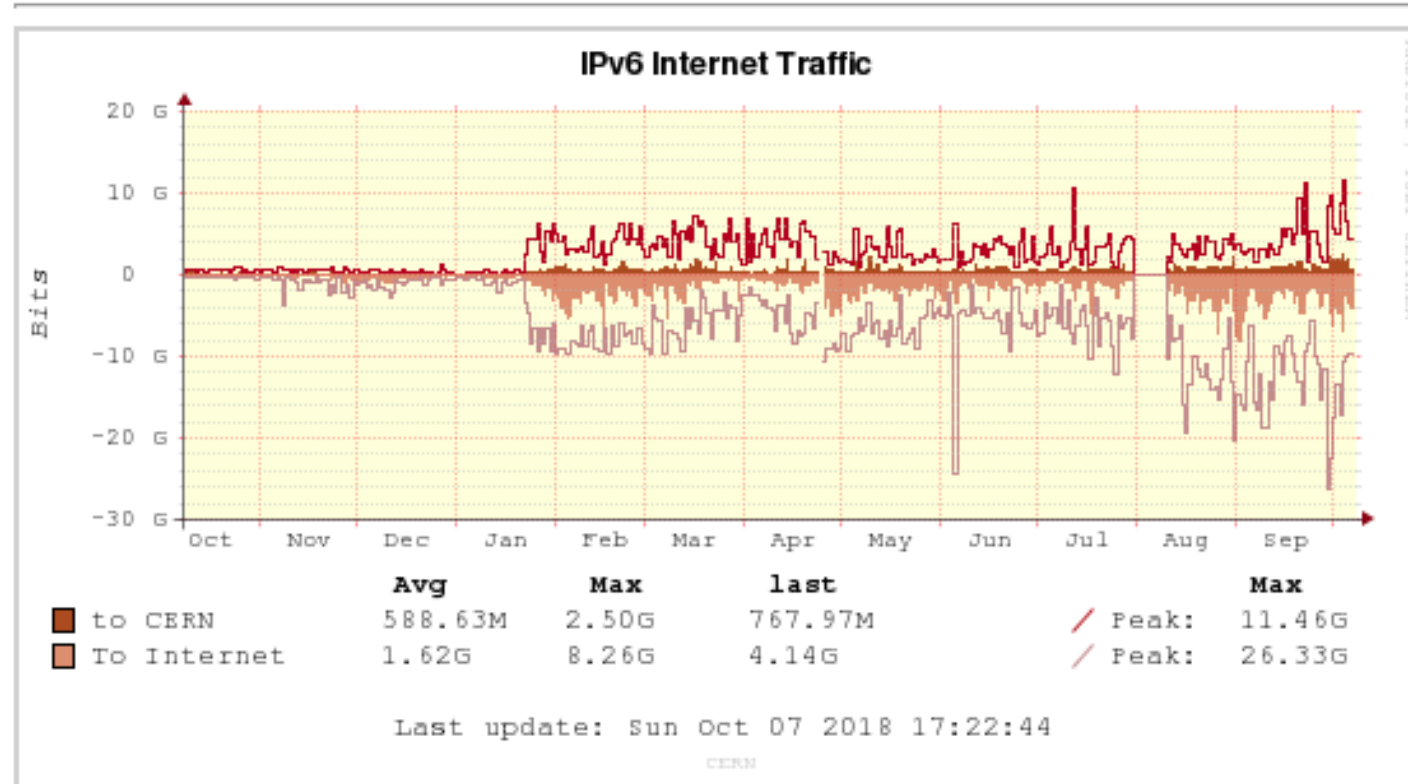
CERN Non-LHCOPN traffic – showing dual-stack EOS turn-on in Jan 2018



IPv6: Internet

ALL Daily Weekly Monthly Yearly

Yearly



Graphic help ?

<https://netstat.cern.ch/monitoring/network-statistics/ext/?q=IPv6&p=EXT&mn=Internet&t=Yearly>

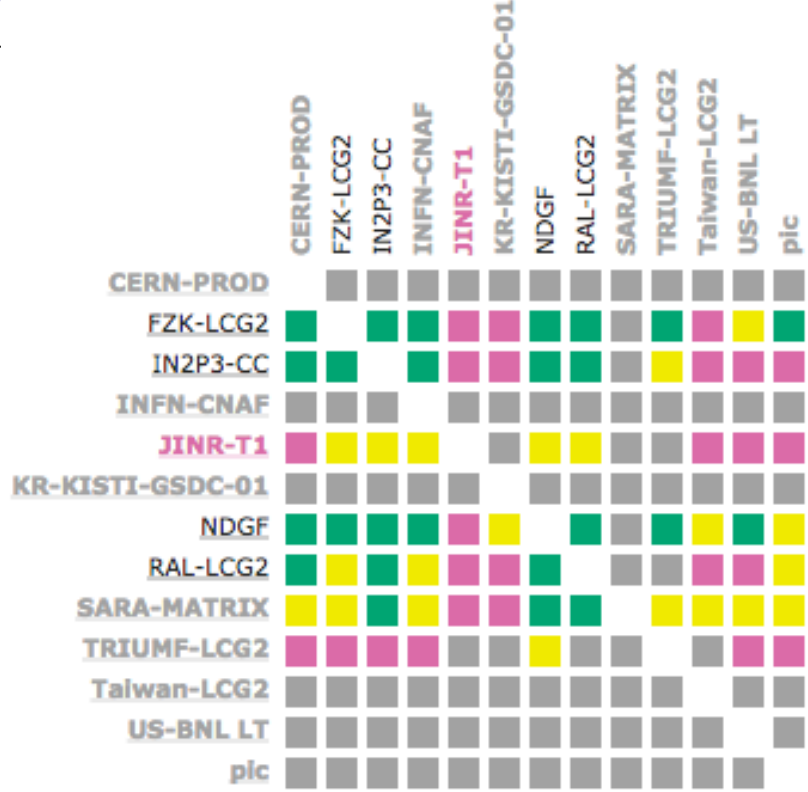
- perfSONAR



OPN Mesh Config - IPv6 TCP BWCTL Test Between OPN Bandwidth Hosts

■ Throughput >= 900Mbps
 ■ Throughput < 900Mbps
 ■ Throughput <= 500Mbps
 ■ Unable to retrieve data
 ■ Check has not yet run

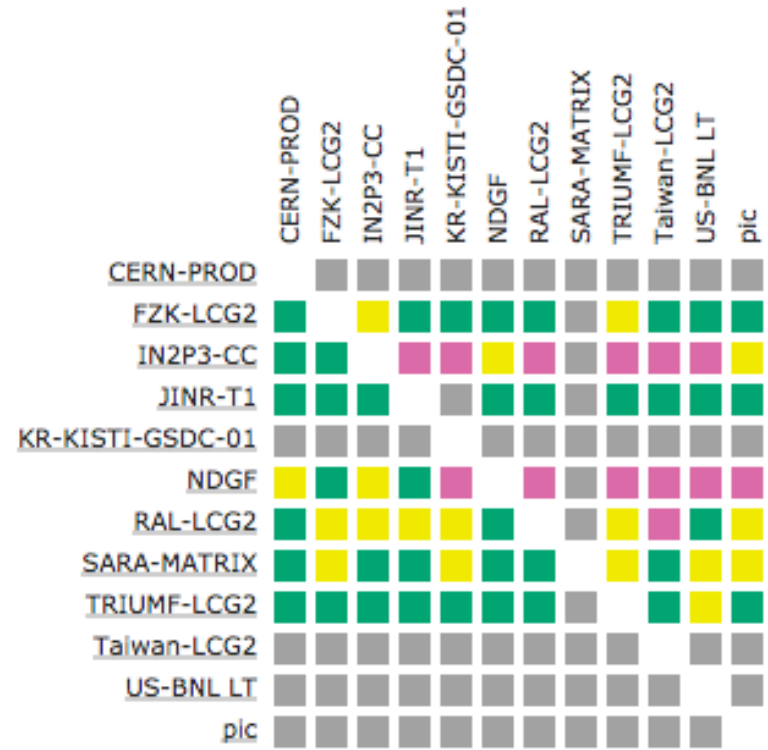
! Found



OPN Mesh Config - IPv6 Traceroute Test Between OPN Bandwidth Hosts

■ Number of Paths is <= 1
 ■ Number of Paths is >= 1
 ■ Number of Paths is >= 2
 ■ Unable to retrieve data
 ■ Check has not yet run

✓ No



<http://psmad.grid.iu.edu/maddash-webui/index.cgi?dashboard=OPN%20Mesh%20Config>