

*Deployment of IPv6-only CPU on  
WLCG – an update from the HEPiX  
IPv6 WG*

Alastair Dewhurst  
Dave Kelsey  
Andrea Sciabà (CERN)  
on behalf of the HEPiX IPv6 WG

HEPiX Spring Workshop 2017  
24-28 April, Budapest

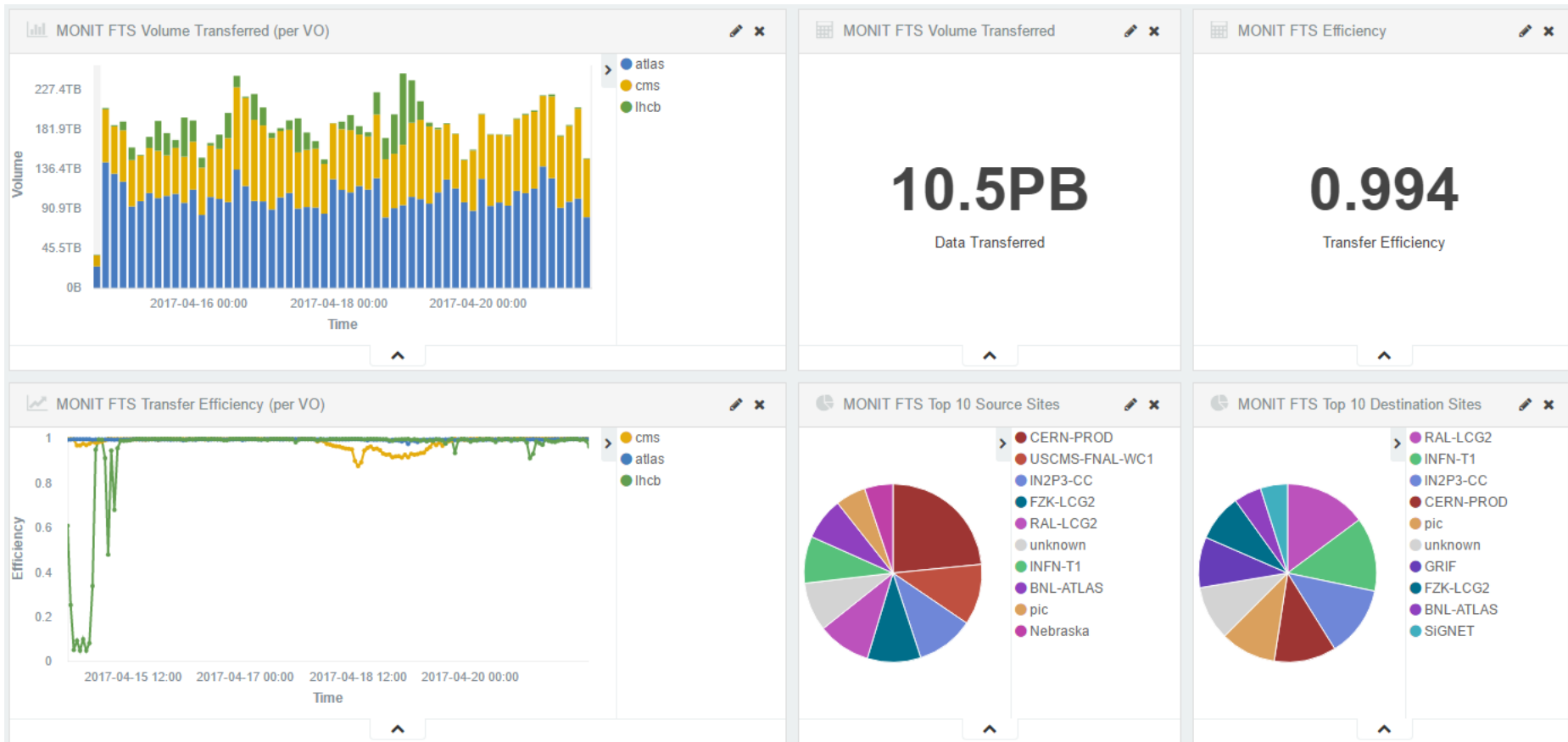
# Outline

- Introduction
- WLCG deployment plan: reminder
- Status update from the experiments
- Status update from the Tier-0/1 sites
- Status update from Tier-2 sites
- Current issues
- Monitoring
- Conclusions

# HEPiX IPv6 WG meetings

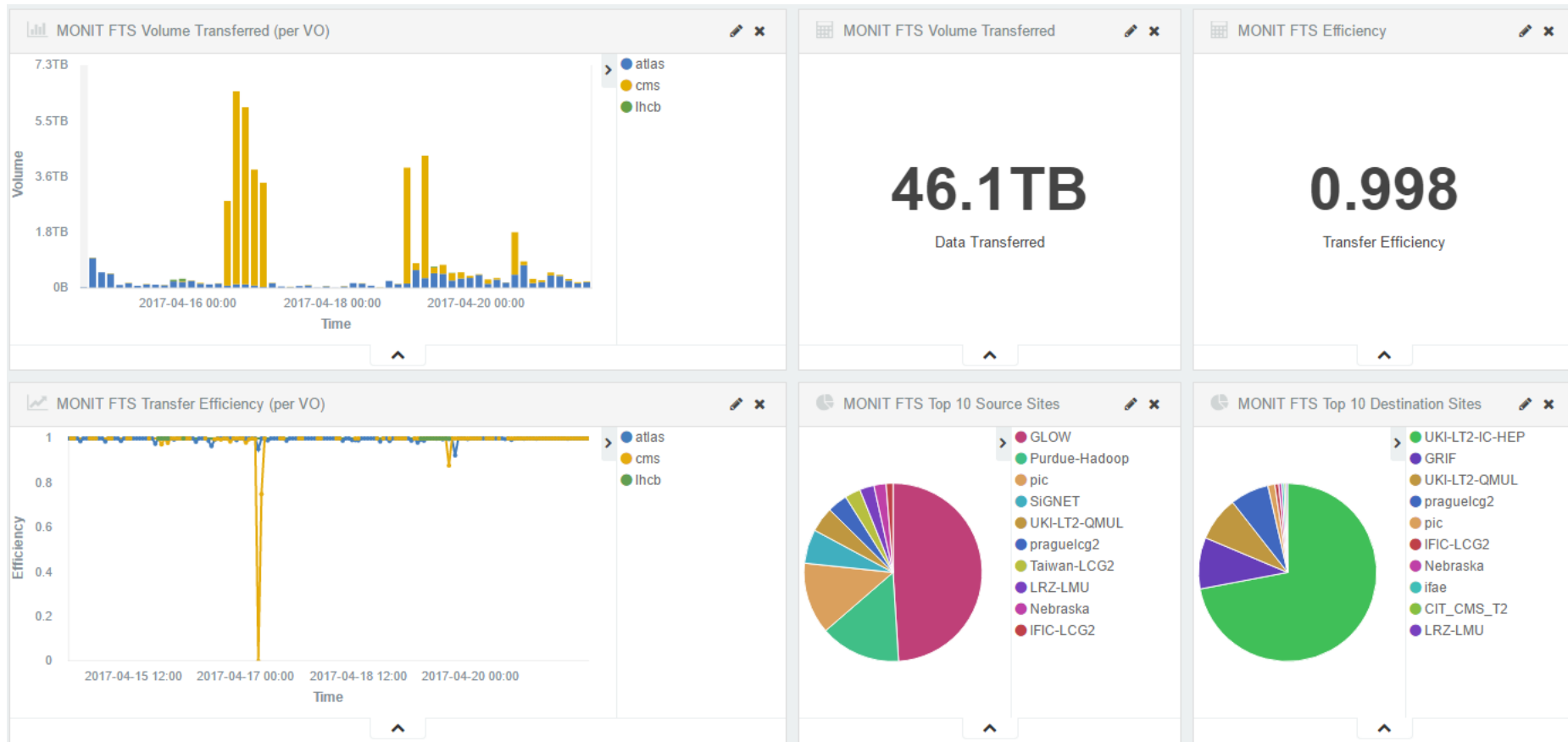
- Meetings held monthly
  - Last F2F at CERN 2/3 Feb 2017
  - Next F2F at CERN 15/16 May 2017
- Participation of all LHC experiments, Tier-0/1 sites and some Tier-2 sites
  - Participation from more sites warmly welcome (e.g. not many US people in the list!)
  - Write to [ipv6@hepixon.org](mailto:ipv6@hepixon.org) to join
- Discuss technical issues, progress reports
  - Best way to get involved and contribute

# FTS transfers (last week) total



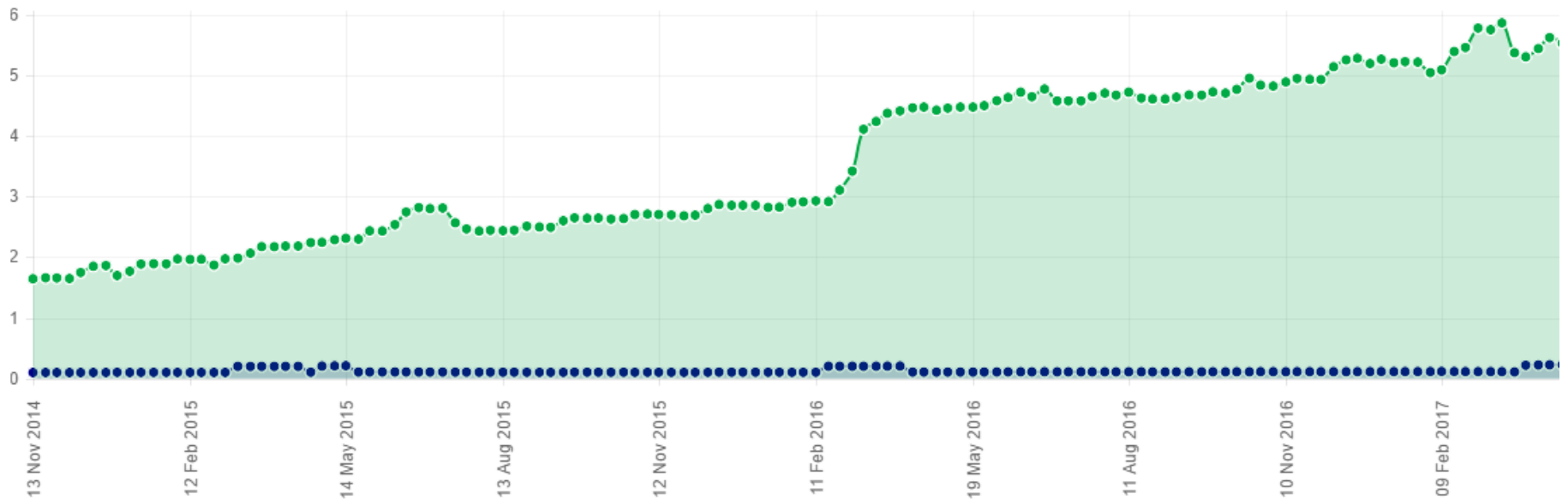
<https://monit.cern.ch/>

# FTS transfers (last week) IPv6



(add "data.ipv6:true" to any FTS monitor page for IPv6 only traffic)

# % BDII services (dual-stack)



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

# WLCG deployment plan: timeline

- By April 1<sup>st</sup> 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 1<sup>st</sup> 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

# Experiments



# ALICE

- ALICE central services have been dual stack for more than a year
- Storage is fully federated: any site can access any data from any other site
  - To support IPv6-only resources, all data must be available on some IPv6-enabled storage
- ALICE can support IPv6-only CPU resources as soon as enough sites have upgraded their storage to dual stack
- F2F meeting with sites in May

# ATLAS

- Computing infrastructure already supports IPv6-only CPUs
- Working on making Frontier servers dual-stack
  - But not critical if the local squid is dual-stack
- Plan to upgrade all externally facing central service boxes to dual stack by April 2018
  - Those not directly accessed by jobs
- ATLAS encourages all sites to upgrade their storage to dual stack
- By the end of Run2, ATLAS will require sites to provide dual stack storage
  - If this is not possible, the data sent to the site and types of workflows may be restricted

# CMS

- CMS expects Tier-2 sites to dual-stack their xrootd-enabled storage from now, and complete by the end of Run2
  - A few volunteer sites identified to pave the way for others
- CMS services are being upgraded to dual-stack
  - Main current activity is to test CPU-only resources on the testbed infrastructure

# LHCb

- Central services fully support dual-stack and IPv6-only CPUs
  - though notice needed for pure-IPv6-only WNs
- Look forward for more dual-stack storage sites

# Tier 0/1s

## Sites IPv6 connectivity

Name	Type	LHCOPN IPv6 peering	LHCONE IPv6 peering	LHCONE IPv6 peers	NREN IPv6 peers	IPv6 LAN	dualstack perSONAR	dualstack storage percentage	Network Statistics	Notes	date last update
CH-CERN	Tier0	Yes	Yes	GEANT, ESnet, CERNlight	GEANT, SWITCH, RENATER, SURFnet, NORDUnet, ASGCnet, KREOnet, Internet2, CANARIE	Yes	Yes	0	<a href="https://netstat.cern.ch/monitoring/network-statistics/ext/?q=IPv6&amp;p=EXT&amp;...">https://netstat.cern.ch/monitoring/network-statistics/ext/?q=IPv6&amp;p=EXT&amp;...</a>		02/16/2017
INFN CNAF	Tier1	Yes	Yes	GARR	GARR	Yes	Yes	0	<a href="https://gins.garr.it/Statistics/viewer.php?stroke_ipv6=on&amp;target%5B%5D=I...">https://gins.garr.it/Statistics/viewer.php?stroke_ipv6=on&amp;target%5B%5D=I...</a>		02/16/2017
RAL	Tier1	No			JANET	No	Yes	0			02/16/2017
FNAL	Tier1	Yes	Yes	ESnet	ESnet	No	Yes	0			02/17/2017
NDGF	Tier1	Yes	Yes	NORDUnet	NORDUnet	Yes	Yes	100	<a href="http://stats.nordu.net/stat-q/load-map/ndgf,,traffic,peak.r">http://stats.nordu.net/stat-q/load-map/ndgf,,traffic,peak.r</a>	percentage of dualstack storage to be checked	
BNL	Tier1	Yes	Yes	ESNet	ESNet	No	Yes	0			02/16/2017
NL-T1 SARA-MATRIX	Tier1	Yes	Yes	GEANT	SURFnet	Yes	Yes	0		Dual stack test SE available. Need to implement IPv6 peering with NIKHEF and KIAE.	03/24/2017
DE-KIT/GridKa	Tier1	Yes	Yes	DFN	DFN	Yes	Yes	2	N/A	dualstack storage → testbed only	02/16/2017
ES-PIC	Tier1	Yes	Yes	RedIRIS	RedIRIS	Yes	Yes	100	N/A		02/16/2017
ASGC	Tier1	Yes	Yes	GEANT, ESnet	GEANT, ESnet	Yes	Yes	0		Dual stack storage will be soon available.	03/21/2017
FR-CCIN2P3	Tier1	Yes	Yes	RENATER	RENATER	Yes	Yes	50	N/A		03/23/2017
TRIUMF	Tier1	Yes	Yes	Canarie	BCNET	Yes	Yes	0		dualstack storage - testbed only	09/01/2016

<http://hepix-ipv6.web.cern.ch/sites-connectivity>

Bruno Hoefft

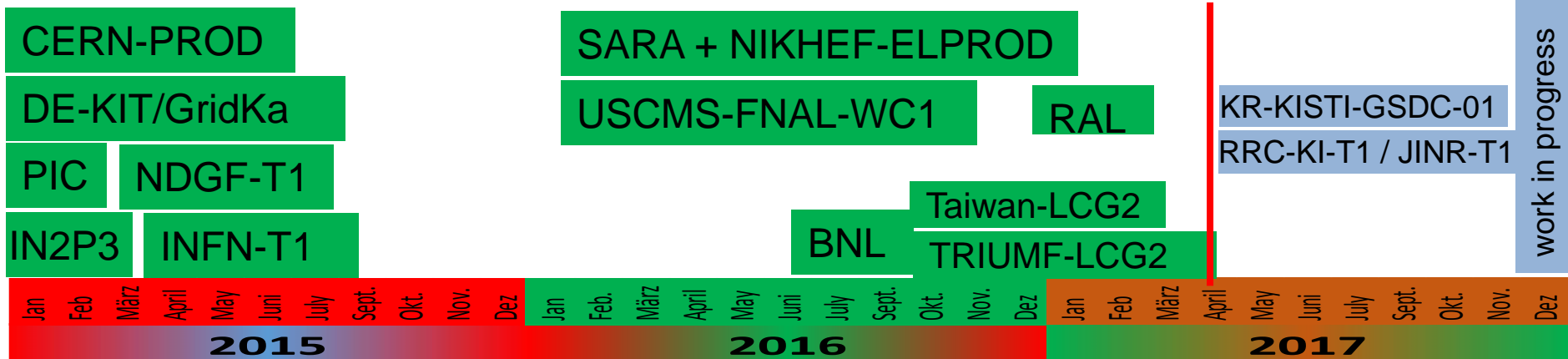
# IPv6 tier-1 site readiness

status at April 2017

- Good IPv6 adoption
  - 11 Tier1s and the Tier0 peering over IPv6
  - dual-stack perfSONAR installed in all of them
- LHCOPN IPv6 still missing from:
  - KR-KISTI (new router hardware needed - by June 2016?)
  - RRC-KI-T1 KIAE (IPv6 deployment started)
  - RRC-KI-T1 JINR (will follow KIAE)

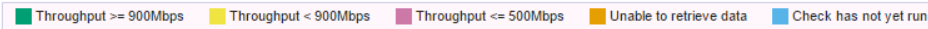
IPv6 peering + dual stack perfsonar

work in progress



# Dual-stack mesh

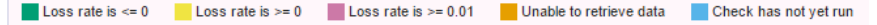
Dual-Stack Mesh Config - IPv6 Bandwidth Test



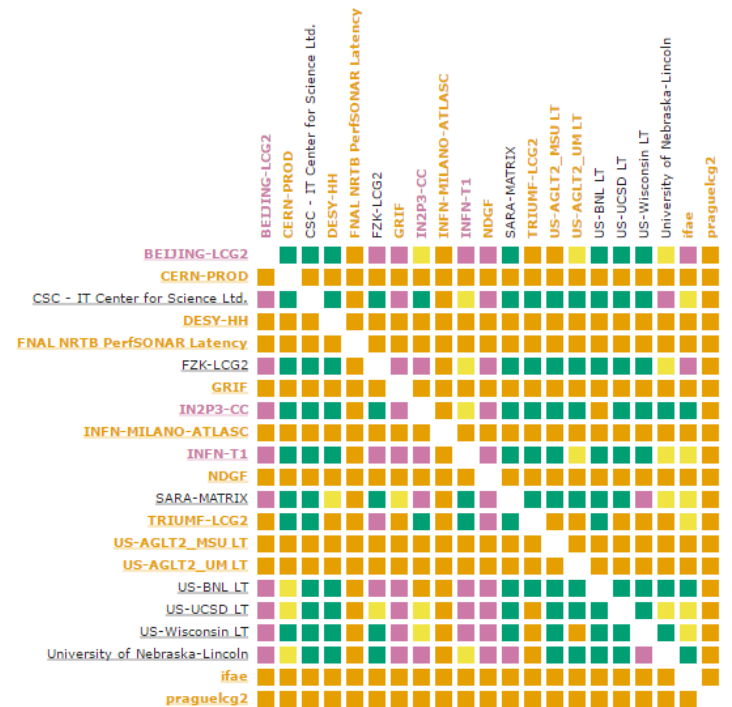
! Found a total of 29 problems involving 16 hosts in the grid



Dual-Stack Mesh Config - IPv6 Latency Test



! Found a total of 14 problems involving 14 hosts in the grid



<http://maddash.aglt2.org/maddash-webui/index.cgi?dashboard=Dual-Stack%20Mesh%20Config>



# CERN

## Input from storage group

- Latest version of EOS (not yet in production) already validated for IPv6 months ago
- Plan to upgrade LHCb production instance to it and put in dual stack in early May
  - Other experiments will follow but only at the end of the 2017 run
  - EOS upgrade also a prerequisite to an EOS catalogue upgrade
- CVMFS fully tested to work on IPv6
- CASTOR will not go to IPv6
  - The long term strategy is to replace it with CTA

# Tier-1 news (1/2)

- PIC
  - All production squid servers now dual-stack
  - Aiming for 100% of WNs before summer
    - Now: nodes 51%, slots 53%, HS06 57%
  - Starting next week migration of batch system to HTCondor [2 CE (HTCondorCE in dual-stack)]
  - All of the dCache storage is in dual stack
- RAL
  - Reported yesterday by James Adams
- SARA
  - Storage dual-stack in production by May 1<sup>st</sup>
  - Setting up peering with NIKHEF, to be finished before summer
- FNAL
  - Tested dual-stack xrootd servers attached to dCache
  - Dual-stack setup for the testbed of the submission infrastructure (ITB)
  - Moving xrootd redirectors to dual stack
  - Have preproduction-level IPv6 capability, now focused at making it production safe

# Tier-1 news (2/2)

- CNAF
  - Dual-stack for CEs almost completed, for storage, to be completed by the end of May
- TRIUMF
  - TRIUMF advertised three IPV6 subnets to the LHCOPN, LHCONE. Perfsonar-PS was configured as a dual stack at the beginning of 2017
  - A three-node testbed with dual stack has been set up
    - Two dCache nodes and one user interface
    - Will test data transfers through IPv6 network with the grid client tools including GFAL client, dCache client and Globus tools
- BNL
  - dCache fully operational in dual-stack
- CCIN2P3
  - dCache and xrootd storage accessible in dual stack
  - Only issue is load balancing on the xrootd redirector, which does not support IPv6
- KIT
  - New storage being setup in dual-stack. Will replace the old one over a period of one year
  - WLCG services will get gradually deployed in dual stack

# Tier 2s

# Sites with IPv6-only CPUs

- Brunel
  - Has dedicated IPv6-only CPU queue on their batch farm
  - ATLAS, CMS and LHCb jobs have all been tested and shown to work there
- QMUL
  - Mix of IPv6-only and IPv4 nodes
    - The IPv6 nodes can talk over IPv4 via a NAT64 if needed
    - Useful to provide logs of what machines are being contacted over IPv4 still

# DPM issues

- Observed transfers from NDGF (dual stack) to Marseilles (IPv4) via FTS3 (dual-stack) to fail ([GGUS:127285](#))
  - Communication from NDGF to FTS3 via IPv6
  - FTS3 returns to NDGF the port but not the IP of the destination disk server!
- Cause is the way GridFTP redirection is implemented in DPM
  - The EPSV command does not have to contain the IP of the server according to RFC 2428 but this is a mistake
    - dCache adds the IP and Globus, GFAL, ARC, FTS, expect it
- Impact is relatively low for WLCG sites
  - Normally redirection happens via SRM and is not affected by this issue
  - Temporary workaround is to disable GridFTP redirection in DPM
  - The best solution would be for Globus to allow “RFC-breaking” but working behaviour

# IGTF CA – CRLs & IPv6

- Ulf Tigerstedt has been monitoring status for an year
- <http://cvmfs-6.ndgf.org/ipv6/overview.php>
- Lowest: 31 working, 8 "has AAAA-record but the network does not work" and 56 IPv4-only CRL servers
- 5<sup>th</sup> April 2017: 44 good/1/53
- So 12 more dual-stacked CAs during one year
- IGTF is currently pushing for all!

# “How to” deploy IPv6 - Tier 2

- Started work collecting knowledge and advice
  - <https://hepixon.web.cern.ch/content/how-deploy-ipv6-wlcg-tier-2-site>
  - The goal is to provide easy to follow recipes and advice on how to make your site IPv6-ready
- Half day “hands on” tutorial in WLCG Workshop confirmed
  - Manchester – 19-22 June 2017
  - Tutorial will be on Thursday morning



# EFT IPv6

- ETF IPv6 instance provides dual-stack testing support for SAM
- Works for all experiments (only ATLAS and CMS configured now)
- Using experiment production topologies
  - it parses a list of CEs/SEs from the experiments feeds and only monitor those that have an IPv6 entry
- To help sites understand status/availability of their IPv6 resources vs IPv4
- This instance can be added to the ETF central, which provides an overview of site services across all experiments
  - so it can be used to compare how site services perform (wrt IPv4 vs IPv6)

State	Service	Icons	Status detail	Age	Checked	Perf-O-Meter
OK	emi cream glExec WN-glExec-atlas/Role=pilot		tds219.pic.es: OK: Success	11 hrs	5 hrs	
OK	IPv6 DNS		DNS OK - 0.005 seconds response time (ce01.pic.es. 1969 IN AAAA 2001:67c:1148:500::11)	2016-11-11 20:14:09	5 hrs	
OK	IPv6 TCP Check		TCP OK - 0.032 second response time on ce01.pic.es port 8443	2017-04-11 21:07:58	5 hrs	
OK	org.atlas.WN-cvmsf-/atlas/Role=lcgadmin		tds219.pic.es: OK	2016-11-11 15:51:02	2017-04-18 16:18:34	
OK	org.atlas.WN-swspace-/atlas/Role=lcgadmin		tds219.pic.es: OK	2016-11-11 15:51:02	2017-04-18 16:18:34	
OK	org.atlas.WN-swspace-/atlas/Role=pilot		tds219.pic.es: OK	2016-11-11 15:55:19	5 hrs	
OK	org.cms.glExec WN-glExec-cms/Role=pilot		tds219.pic.es: OK: Success	2017-03-15 20:01:28	2017-03-28 16:01:28	
OK	org.cms.WN-analysis-/cms/Role=lcgadmin		tds219.pic.es: OK: OK	2017-02-10 19:57:06	5 hrs	
OK	org.cms.WN-basic-/cms/Role=lcgadmin		tds219.pic.es: OK: OK	2017-02-10 19:57:06	5 hrs	
OK	org.cms.WN-cvmsf-/cms/Role=lcgadmin		tds219.pic.es: OK: cvmsf vers 2.1.19 (probe 1.3-pre2)	2017-02-27 13:57:07	5 hrs	
OK	org.cms.WN-env-/cms/Role=lcgadmin		tds219.pic.es: OK: OK	2017-02-10 19:57:06	5 hrs	
OK	org.cms.WN-frontier-/cms/Role=lcgadmin		tds219.pic.es: OK	2017-02-10 19:57:06	5 hrs	
OK	org.cms.WN-mc-/cms/Role=production		tds219.pic.es: OK: OK	2017-03-28 09:05:43	2017-03-28 16:05:44	
OK	org.cms.WN-remote-stageout-/cms/Role=lcgadmin		tds219.pic.es: OK: OK	2017-02-10 19:57:06	5 hrs	
OK	org.cms.WN-squid-/cms/Role=lcgadmin		tds219.pic.es: OK	2017-02-10 19:57:06	5 hrs	
OK	org.cms.WN-swinst-/cms/Role=lcgadmin		tds219.pic.es: OK: OK	2017-02-10 19:57:06	5 hrs	
CRIT	org.cms.WN-xrootd-access-/cms/Role=lcgadmin		tds219.pic.es: CRITICAL: Failed cmsRun: exit code 21504	2017-02-10 19:57:06	5 hrs	
WARN	org.cms.WN-xrootd-fallback-/cms/Role=lcgadmin		tds219.pic.es: WARNING: Failed cmsRun: exit code 23552	2017-02-10 19:57:06	5 hrs	

[https://etf-ipv6.cern.ch/etf/check\\_mk/index.py](https://etf-ipv6.cern.ch/etf/check_mk/index.py)

# Summary

- Much improved engagement by Tier 1s
- Still limited dual-stack storage by April 1<sup>st</sup> at Tier-1 but coming soon everywhere
- A good number of Tier 2s run dual-stack
  - But still a small minority, overall
- **WLCG Tier 2s must start planning NOW**
- Very few issues left (notably with DPM, not critical)
- Automatic endpoint monitoring in place
- Still a lot of work to do
- How best to track/urge/encourage/support the Tier 2's
  - Produce documentation!

# Links

- HEPiX IPv6 web

<http://hepix-ipv6.web.cern.ch>

- Working group meetings

<http://indico.cern.ch/categoryDisplay.py?categId=3538>

- WLCG Operations IPv6 Task Force

<http://hepix-ipv6.web.cern.ch/content/wlcg-ipv6-task-force-0>