

Plans to support IPv6-only CPU on WLCG

David Kelsey (STFC-RAL)
Chair of HEPiX IPv6 working group

EGI OMB meeting

15 Dec 2016

Members of the HEPiX IPv6 group 2016

M. Babik (CERN), J. Chudoba (FZU), A. Dewhurst (RAL), S. Fayer (Imperial), T. Finnern (DESY), T. Froy (QMUL), C. Grigoras (CERN), K. Hafeez (RAL), B. Hoefl (KIT), T. Idiculla (RAL), D. Kelsey (RAL, Chair), F. Lopez Munoz (PIC), E. Martelli (CERN), R. Nandakumar (RAL), K. Ohrenberg (DESY), F. Prelz (INFN), D. Rand (Imperial), A. Sciaba (CERN), U. Tigerstedt (CSC), D. Traynor (QMUL)

And recently joined: H. Ito (BNL), A. Zaytsev (BNL), X. Espinal (CERN), J. Belleman (CERN), A. Falabella (CNAF), M. Bly (RAL), C. Condurache (RAL)

Plus many others in earlier times

Many thanks to them all!

Overview

- Working Group timeline
- IPv6 global status
- Dual-stack storage and IPv6 data transfers
- Support for IPv6-only CPU

HEPiX IPv6 Working Group

Timeline

- 2011: Started work

CHEP2012: “From IPv4 to eternity” – a talk presented the working group plans

- 2012-2014: Testbed activities and working with Storage developers

CHEP2013: Paper “WLCG and IPv6 - the HEPiX IPv6 working group” (testbed)

- 2015: Requested LHCOPN/ONE IPv6 peering & dual-stack perfSONAR

CHEP2015: Paper “The production deployment of IPv6 on WLCG” (dual-stack)

- 2016: Now encouraging more dual-stack storage

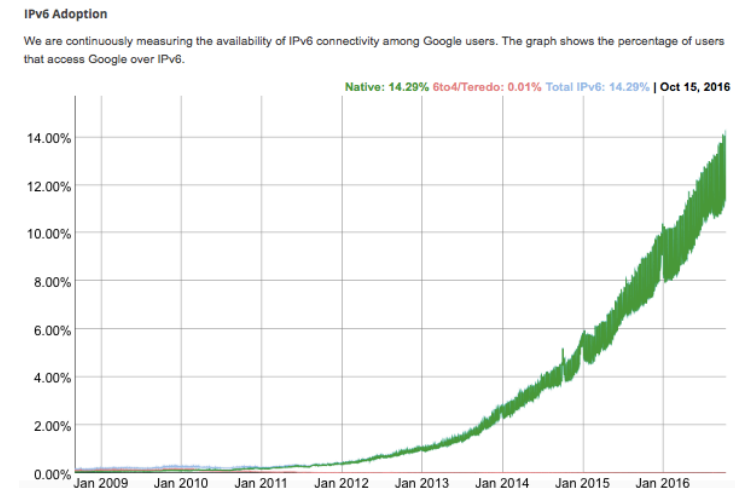
CHEP2016: Plans for IPv6-only CPU

- 2017: First support for IPv6-only CPU

Global IPv6 status

*Based on slide from
Alastair Dewhurst*

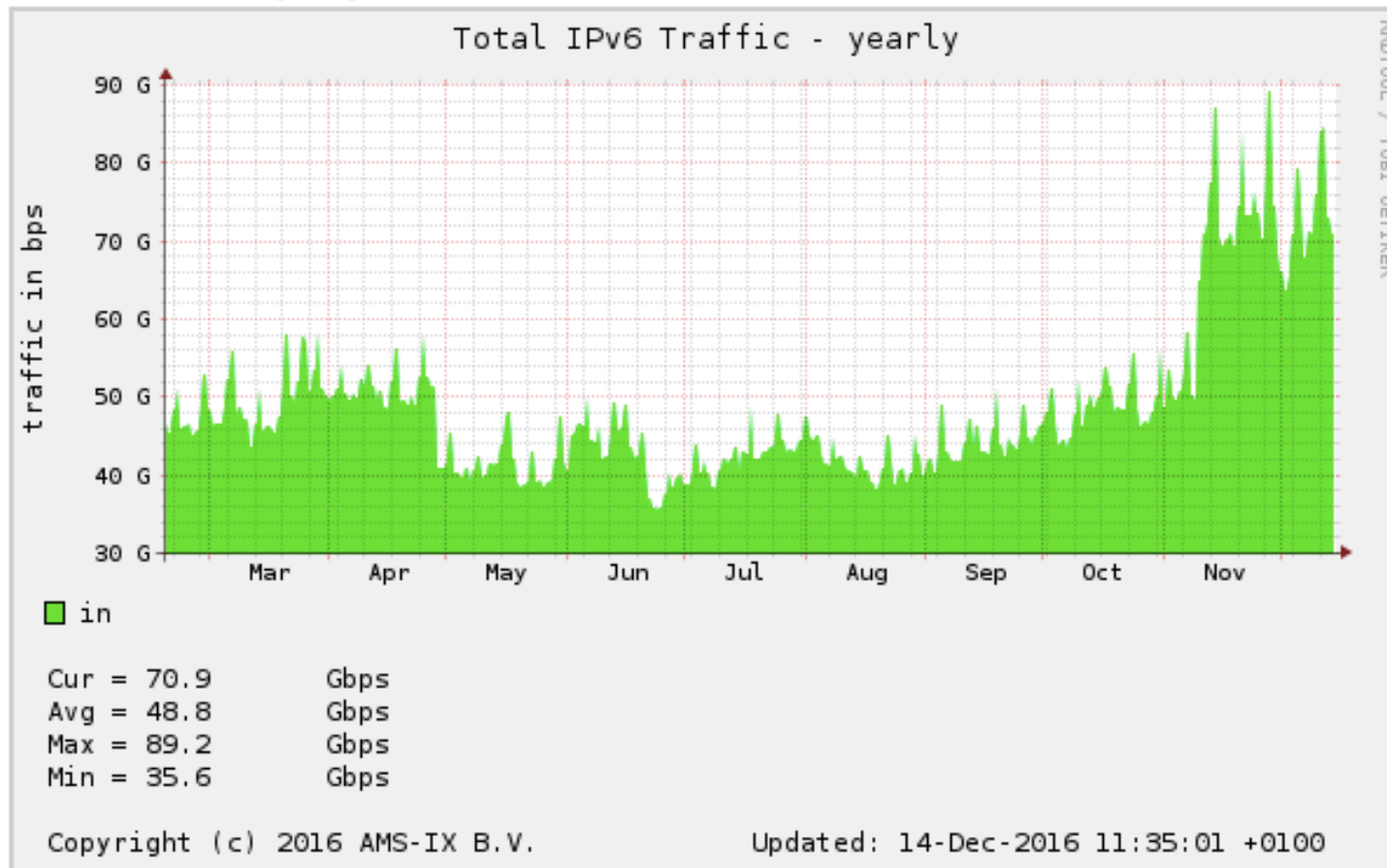
- Over 14% of access to Google services now via IPv6
 - <https://www.google.com/intl/en/ipv6/statistics.html>
- Some commercial hosting companies offer cheaper IPv6 only services
- June 2016, Apple now mandates all Apps submitted to the App Store must support IPv6-only networking
- September 2016, Microsoft Azure VMs now available as IPv6
- For HEP:
 - 11 sites currently provide dual stack storage.
 - 2 sites running IPv6 WN (With NAT64/DNS64).



IPv6 traffic at AMS-IX 2016

<https://ams-ix.net/technical/statistics/sflow-stats/ipv6-traffic>

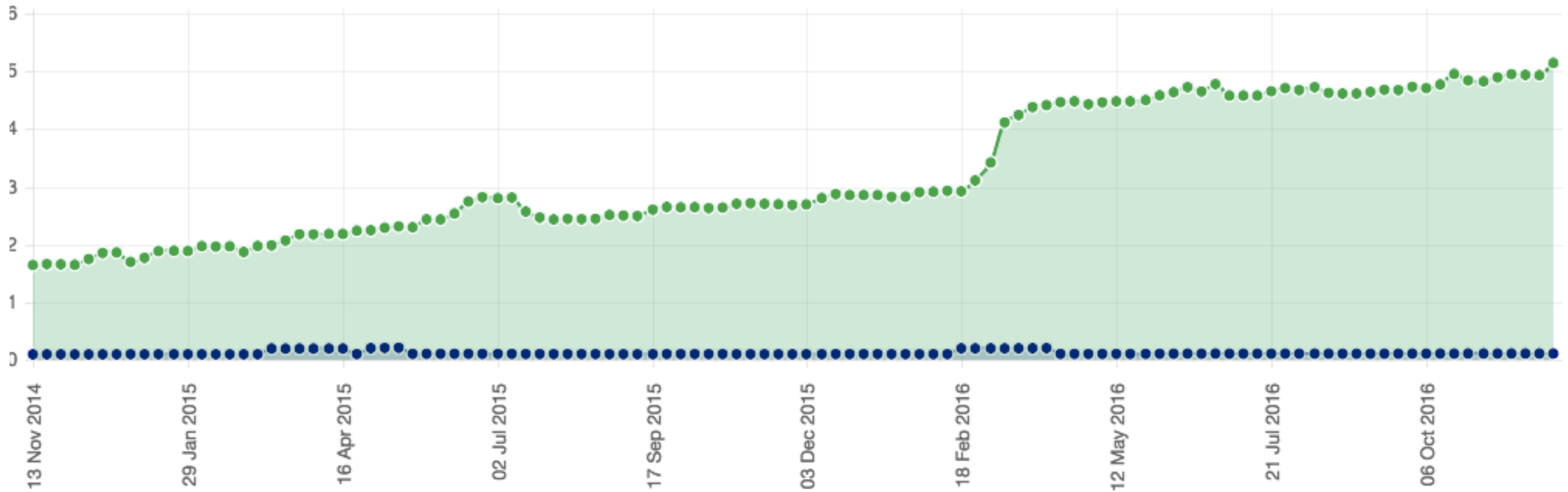
Total IPv6 Traffic - yearly



WLCG Dual-Stack services

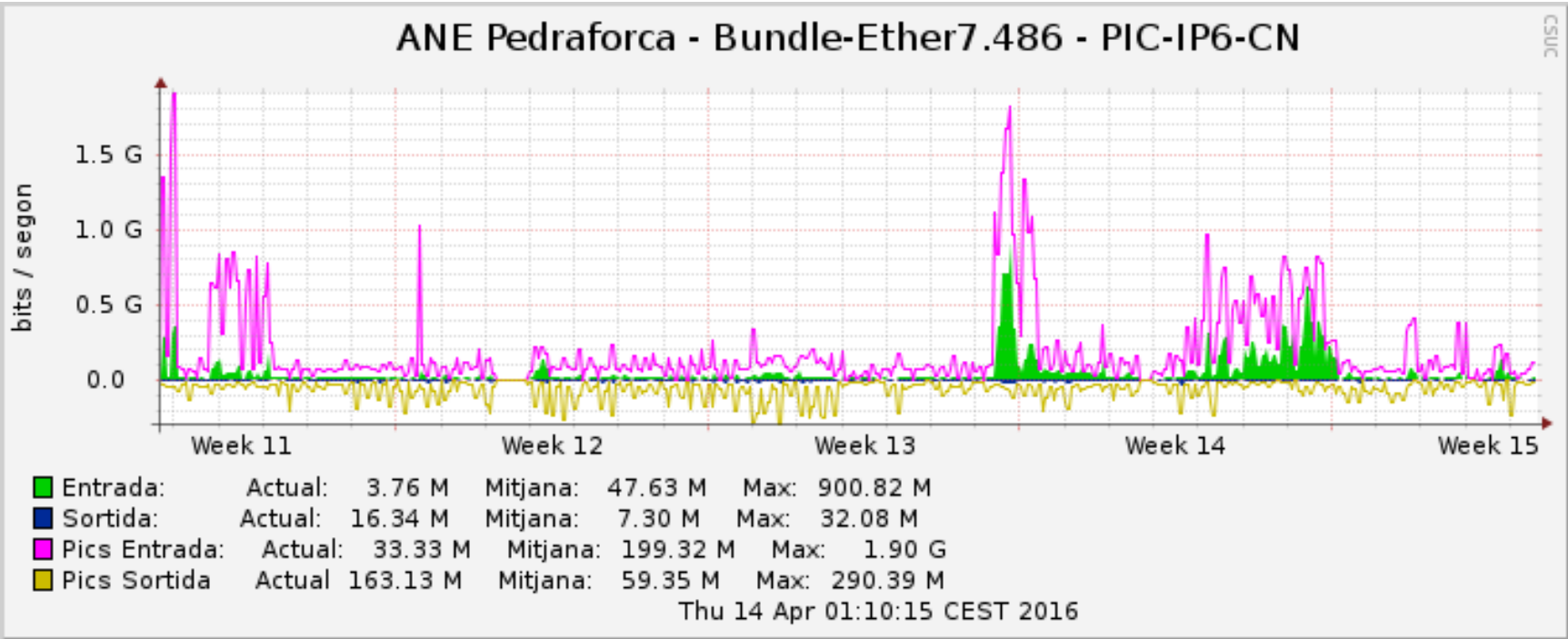
Services in BDII

Fraction of IPv6 reached ~5%

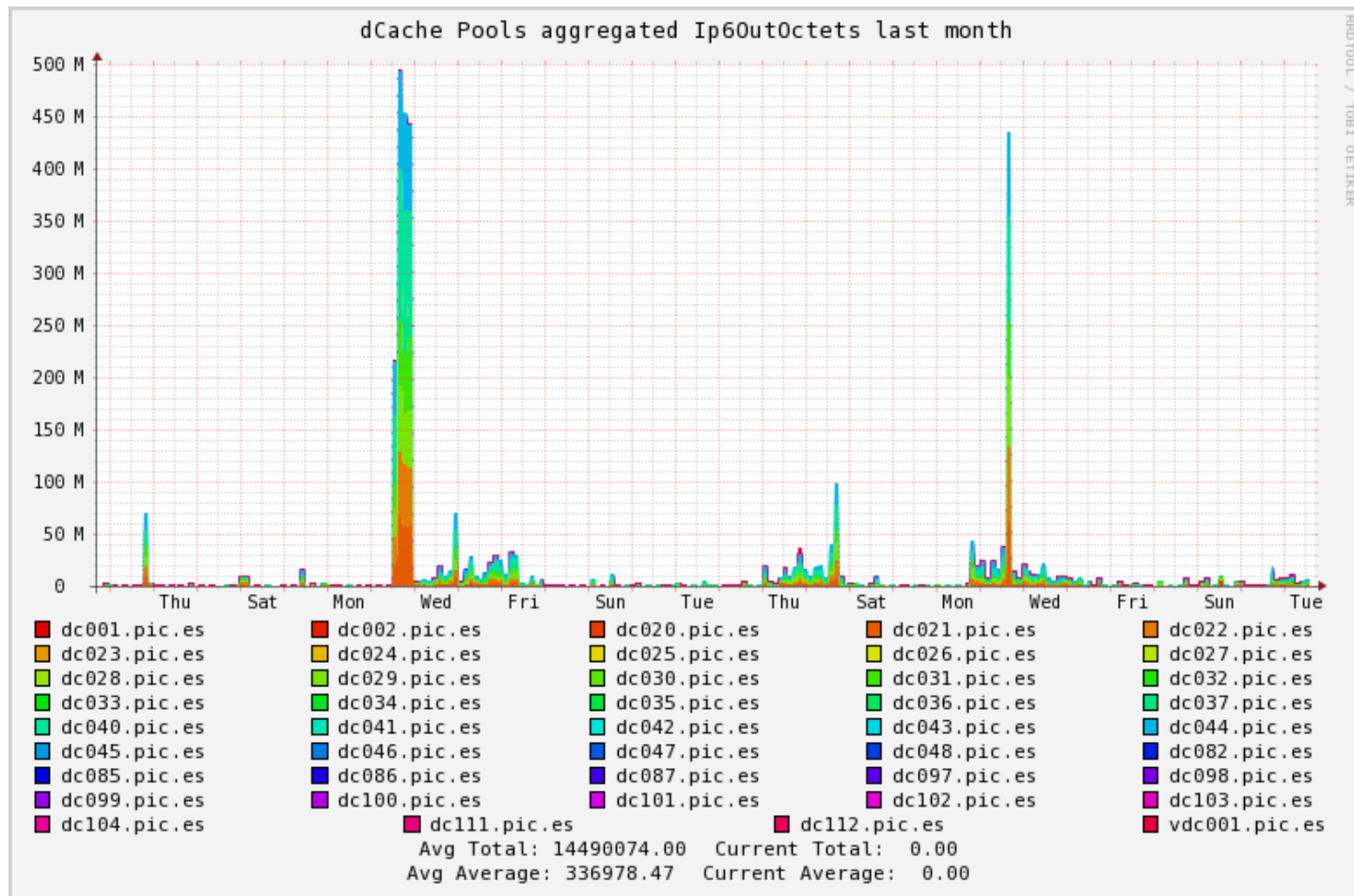


PIC IPv6 traffic 2016

dCache IPv6 traffic from PIC to Tier 2 over general internet



PIC dCache IPv6 traffic (2)

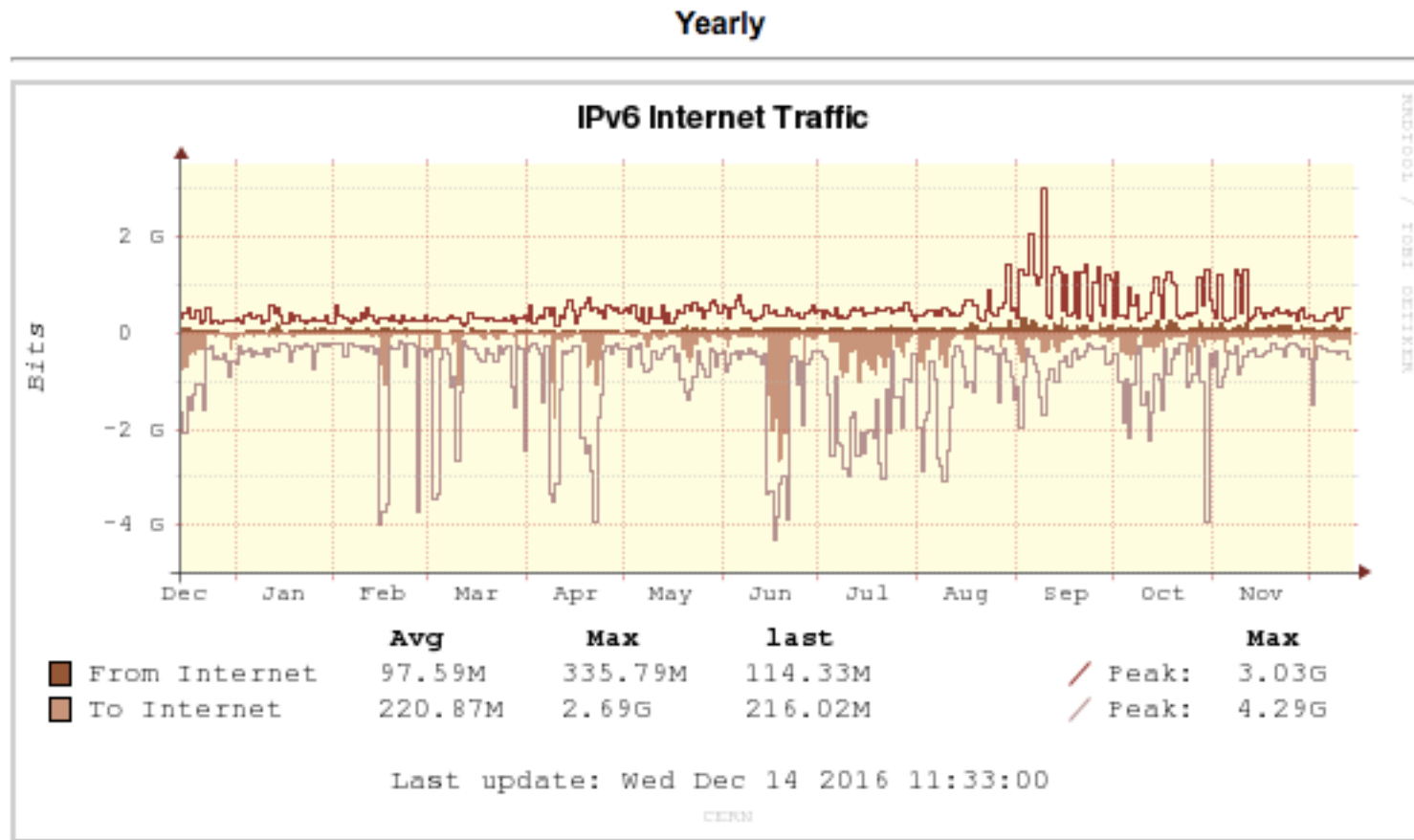


dCache IPv6 traffic from PIC to other sites in Aug 2016 (MBps)

Peaks of ~500 MBytes/s
 (4 Gbps average)
 - to Imperial, London

Tuesday 30th August with a duration of about 12 hours

IPv6 internet traffic CERN



[Graphic help ?](#)

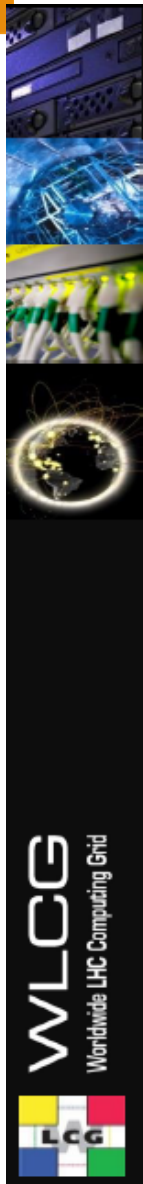
Dual-stack Storage & IPv6 data transfers

- Most storage solutions and protocols now work in dual-stack mode
 - dCache, DPM, StoRM, FTS
 - XrootD 4, GridFTP, http
- Several sites have been running dual-stack for some time
- Production data transfers over IPv6 are happening!
- We are working on per experiment monitoring over time
 - What fraction of data transfers are IPv6?
- It works!

Support for IPv6-only worker nodes/VMs

Introduction

- IPv6 requirements of the experiments have been discussed at the last GDB.
- A document has been submitted to the MB for **approval**.
 - A summary is presented here.
- The main aims:
 - To provide a viable migration path for sites needing to switch to IPv6.
 - To allow sites to make long term planning decisions regarding their network setup.
 - To allow VOs to make use of IPv6-only CPU resources should they become available in future.

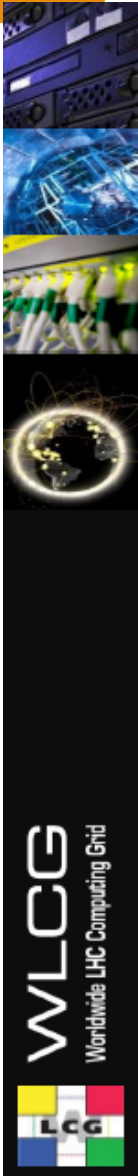


WLCG
Worldwide LHC Computing Grid



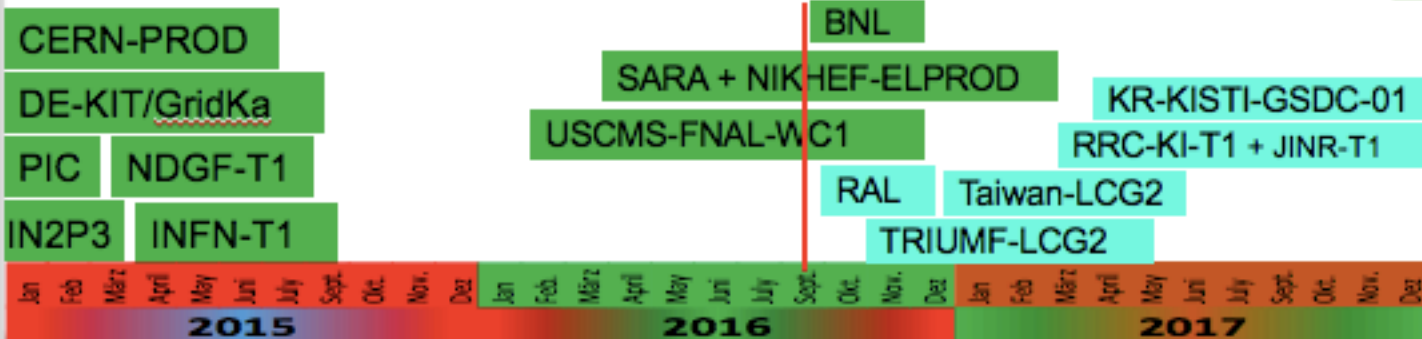
Proposed timeline

- 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
 - GOADB, 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



IPv6 tier-1 site readiness ticket status at Monday 19. Sept.

Ticket-ID	Type	VO	Site	Priority	Resp. Unit	Status	Last Update	Subject	
121896		none	BNL-ATLAS	very urgent	OSG(Prod)	solved	2016-09-14	Tier-1 LHCOPN IPv6 Peering, incl. dualst...	✓
121895		none	USCMS-FNAL-WC1	very urgent	OSG(Prod)	closed	2016-05-21	Tier-1 LHCOPN IPv6 Peering, incl. dualst...	✓
121894		none	RAL-LCG2	top priority	NGI_UK ▶ assigned	closed	2016-06-30	Tier-1 LHCOPN IPv6 Peering, incl. dualst...	ok
121893		none	Taiwan-LCG2	top priority	ROC_Asia/Pacific	closed	2016-06-30	Tier-1 LHCOPN IPv6 Peering, incl. dualst...	ok
121892		none	SARA-MATRIX	top priority	NGI_NL	closed	2016-06-30	Tier-1 LHCOPN IPv6 Peering, incl. dualst...	✓
121891		none	RRC-KI-T1	top priority	ROC_Russia ▶ assigned	closed	2016-05-30	Tier-1 LHCOPN IPv6 Peering, incl. dualst...	ok
121890		none	NIKHEF-ELPROD	top priority	NGI_NL	closed	2016-05-15	Tier-1 LHCOPN IPv6 Peering, incl. dualst...	✓
121889		none	NDGF-T1	top priority	NGI_NDGF	closed	2016-06-16	Tier-1 LHCOPN IPv6 Peering, incl. dualst...	✓
121888		none	KR-KISTI-GSDC-01	top priority	ROC_Asia/Pacific	closed	2016-06-30	Tier-1 LHCOPN IPv6 Peering, incl. dualst...	ok
121887		none	JINR-T1	top priority	ROC_Russia	closed	2016-06-30	Tier-1 LHCOPN IPv6 Peering, incl. dualst...	ok
121886		none	INFN-T1	top priority	NGI_IT	closed	2016-05-15	Tier-1 LHCOPN IPv6 Peering, incl. dualst...	✓
121885		none	IN2P3-CC	top priority	NGI_FRANCE	closed	2016-05-15	Tier-1 LHCOPN IPv6 Peering, incl. dualst...	✓
121884		none	pic	top priority	NGI_IBERGRID	closed	2016-06-15	Tier-1 LHCOPN IPv6 Peering, incl. dualst...	✓
121883		none	FZK-LCG2	top priority	NGI_DE	closed	2016-06-15	Tier-1 LHCOPN IPv6 Peering, incl. dualst...	✓
121882		none	CERN-PROD	top priority	ROC_CERN	closed	2016-06-23	Tier-1 LHCOPN IPv6 Peering, incl. dualst...	✓
121881		none	TRIUMF-LCG2	top priority	ROC_Canada	closed	2016-05-21	Tier-1 LHCOPN IPv6 Peering, incl. dualst...	ok



IPv6 peering + dual stack personal

work in progress

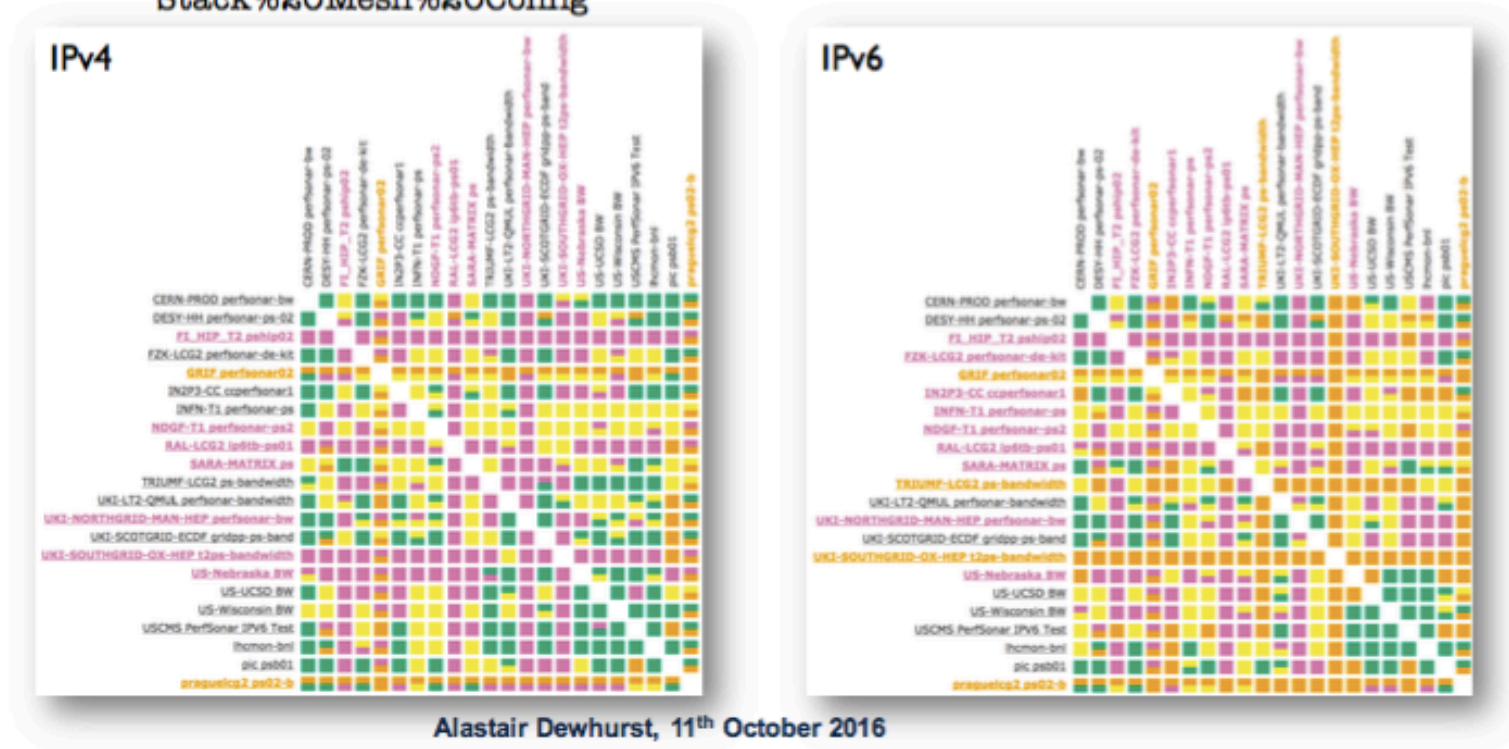
no response

LHCOPN/IPv6 updates

- TRIUMF now peering over IPv6
- ASGC/Taiwan peering over IPv6 for perfSONAR and testing
- Dual-stack perfSONAR
 - new sites: BNL, RAL, TRIUMF, ASGC
- LHCOPN peering still to come
 - RAL
 - KISTI
 - Russia

PerfSonar

- Meshes show dual stack PerfSonar endpoints
- <http://maddash.aglt2.org/maddash-webui/index.cgi?dashboard=Dual-Stack%20Mesh%20Config>



Alastair Dewhurst, 11th October 2016





Deployment of IPv6 only CPU resources at WLCG sites

Alastair Dewhurst, on behalf of the HEPiX
IPv6 Working group

Slides shown at CHEP2016, San Francisco, Oct 2016

IPv6 only CPU

5

The HEPiX IPv6 group made a proposal[1] to the WLCG Management Board which has now been accepted to allow sites to migrate their CPU resources to IPv6 only by April 2017.

- All VOs encourage sites to upgrade their storage to dual stack.
- All VOs are working towards making their central services required by Grid jobs dual stack by April 2017.
- Shared central services (e.g. CVMFS) should be accessible via IPv6 by April 2017.
- Tier 1s should provide functional dual stack access to storage (and other services they may run):
 - 1GB/s and 90% availability by April 2017.
 - 10Gb/s and 95% availability by April 2018.

[1] <https://indico.cern.ch/event/467575/contributions/1145552/attachments/1311592/1962831/WLCGIPv6Deployment.pdf>

Alastair Dewhurst, 11th October 2016



Central Services

7

- The FTS service is used to transfer large volumes of data between sites.
 - Even though third-party copy is used, the FTS service needs to be dual stack to allow transfers to go via IPv6.
 - Both sites need to have dual stack storage
 - FTS3 pilot service at CERN supports IPv6, ongoing discussion on how best to upgrade the rest.
- CVMFS is used by all the LHC VOs to distribute software to WN.
 - Assuming site squid is dual stack then stratum 1 does not need to be.



Alastair Dewhurst, 11th October 2016



Conclusion

12

- There is now a WLCG Management Board endorsed plan to allow sites to run IPV6 only CPU.
- The LHC VOs all have plans to upgrade their infrastructure to meet this goal.
- We expect to review the progress towards IPv6 adoption in 2018 and expand the plan to allow sites to not need IPv4 at all.



Alastair Dewhurst, 11th October 2016



Some final words

- Any site wishing to deploy IPv6-only CPU (WN/VM)
 - Must discuss with the working group
 - Must discuss with the Expts/VOs they support
- If you notice IPv6 problems
 - Please investigate (often down to config. errors)
 - Ask ipv6@hepik.org for help/advice
 - Don't just turn off IPv6 (or make IPv4 default)
- WLCG Tier 2s need to plan NOW for IPv6 deployment
 - Work together with your site network team
 - Dual-stack perfSONAR
 - Then dual-stack storage
- Request to EGI for all central services and webs/wikis to be dual-stack

Links

- HEPiX IPv6 web

<http://hepixonweb.cern.ch>

- Working group meetings

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

- WLCG Operations IPv6 Task Force

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

- Papers published in proceedings of CHEP2013 and CHEP2015 (and CHEP2016 to come)

Questions?

ALICE

8

- ALICE uses a fully federated storage model based on XrootD.
 - XrootD 4 will need to be provided by the sites
- 100% of data must be available via IPV6 before any WN can be IPv6 only.
 - As data is duplicated not all sites need to upgrade (but the vast majority will).
- ALICE central services have been dual stack for over a year.
- Around 5% of ALICE sites have upgraded their storage to dual stack.
 - Aim for all sites to upgrade by end of Run 2.



Alastair Dewhurst, 11th October 2016



ATLAS

9

- ATLAS are focusing on making services required by IPv6 only WN dual stack first. These all use https:
- The PanDA servers.
 - Pilots contact the Panda servers to pull in jobs as well as send regular updates.
- The DDM headnodes:
 - These are contacted by the jobs to register files that have been uploaded.
- Some services such as pilot factories have been already made dual stack.



Alastair Dewhurst, 11th October 2016



CMS

10

- The job submission middleware, GlideinWMS has been validated as IPv6 compliant.
- The data management system, PhEDEx, uses the Oracle client for communication between local site agents and the central service.
- CMS is upgrading all its central services to dual stack.
 - Focus on the services needed directly by WN - aim to complete by April 2017.
 - All services by end of run 2.



Alastair Dewhurst, 11th October 2016



LHCb

11

- LHCb uses the DIRAC framework to submit jobs to the Grid.
- DIRAC officially supports IPv6 and other users are already running with dual stack instances.
- LHCb has 40 VO boxes at CERN.
 - Rationalising these machines down to ~10.
 - All new machines will be dual stack, currently 4 deployed and tested.



Alastair Dewhurst, 11th October 2016

