

WLCG and IPv6: The HEPiX IPv6 Working Group

David Kelsey (STFC-RAL)

CHEP2013, Amsterdam

17 Oct 2013

Co-authors

Thanks to all my colleagues

K. Chadwick (FNAL), S. Campana (CERN), G. Chen (IHEP), J. Chudoba (FZU), P. Clarke (Edinburgh), M. Elias (FZU), A. Elwell (CERN), S. Fayer (Imperial), Q. Fazhi (IHEP), T. Finnern (DESY), L. Goossens (CERN), C. Grigoras (CERN), B. Hoefft (KIT), T. Kouba (FZU), F. Lopez Munoz (PIC), E. Martelli (CERN), M. Mitchell (Glasgow), A. Nairz (CERN), K. Ohrenberg (DESY), A. Pfeiffer (CERN), F. Prelz (INFN), D. Rand (Imperial), M. Reale (GARR), S. Rozsa (Caltech), A. Sciaba (CERN), R. Voicu (Caltech), C. J. Walker (QMUL), T. Wildish (Princeton)

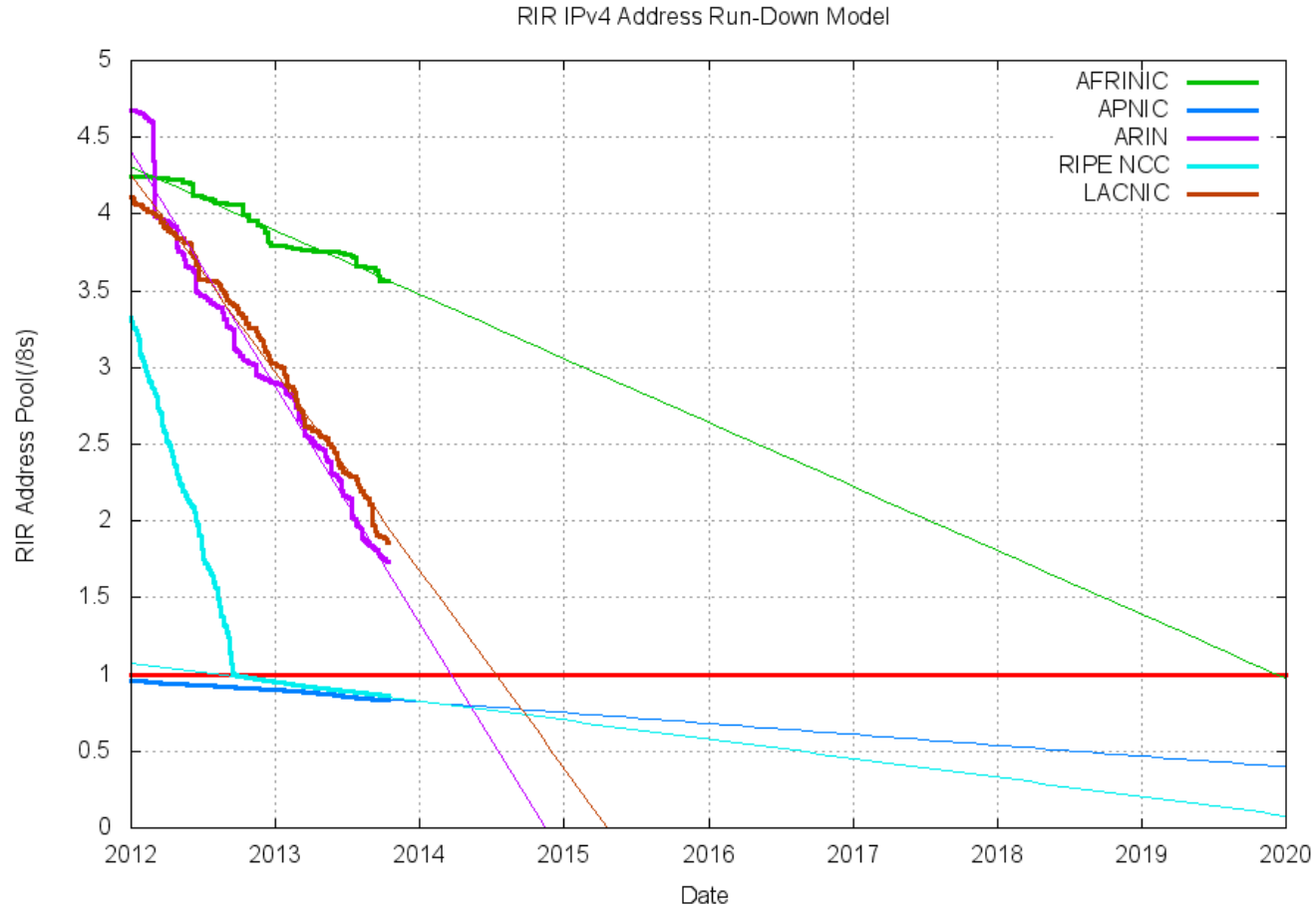
Outline

Update since CHEP2012

- General IPv4 address exhaustion and use of IPv6 growing
- IPv4 & IPv6 at CERN
- The HEPiX IPv6 working group
- HEPiX IPv6 testbed
 - And file transfer tests
- Software and tools survey
- Future plans
 - Use cases for testing
 - Highlighting problem areas
- Summary

General IPv4 address exhaustion and use of IPv6 ...

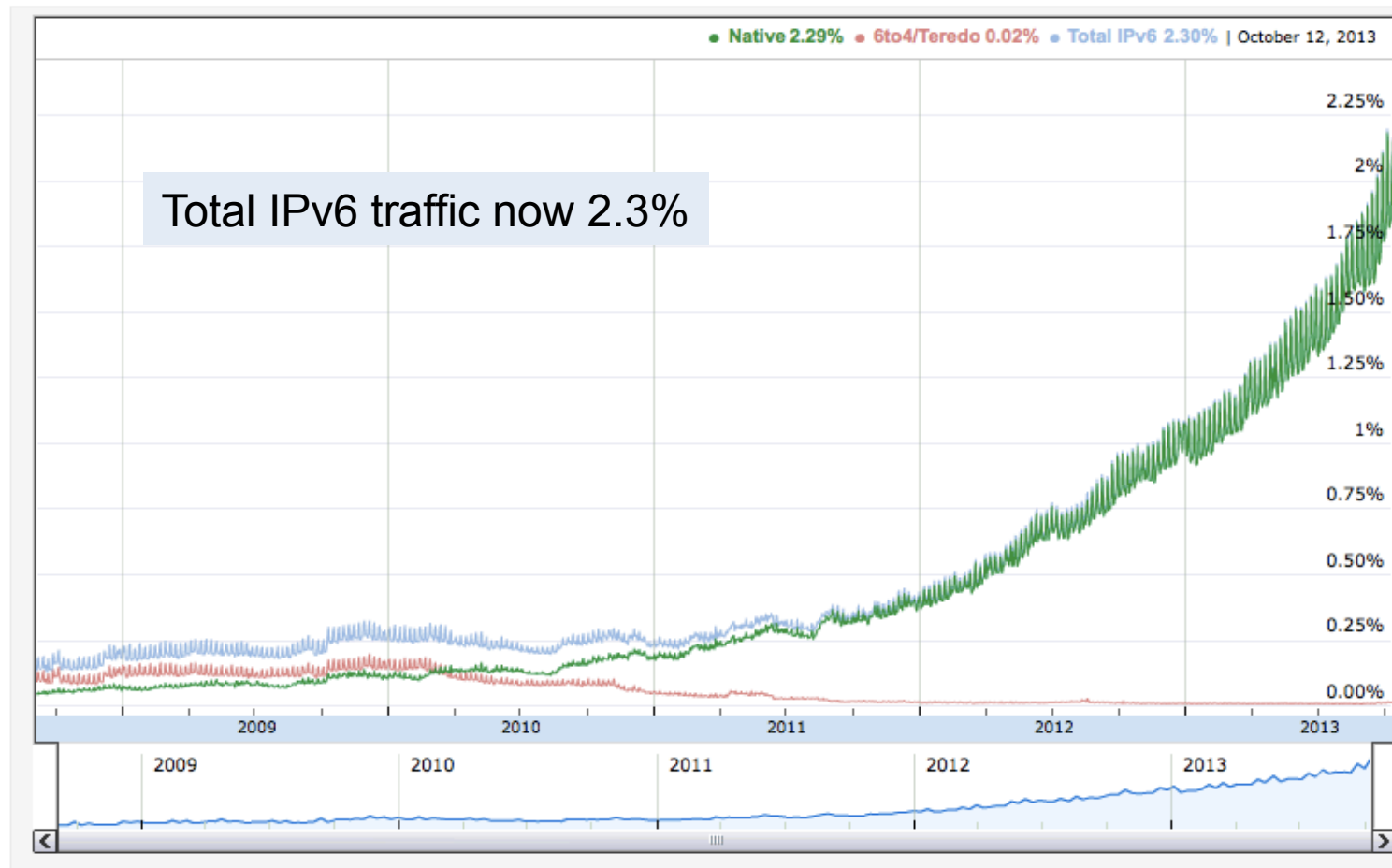
RIR IPv4 Available Addresses (/8 blocks)



<http://www.potaroo.net/tools/ipv4/index.html>

IPv6 clients accessing Google

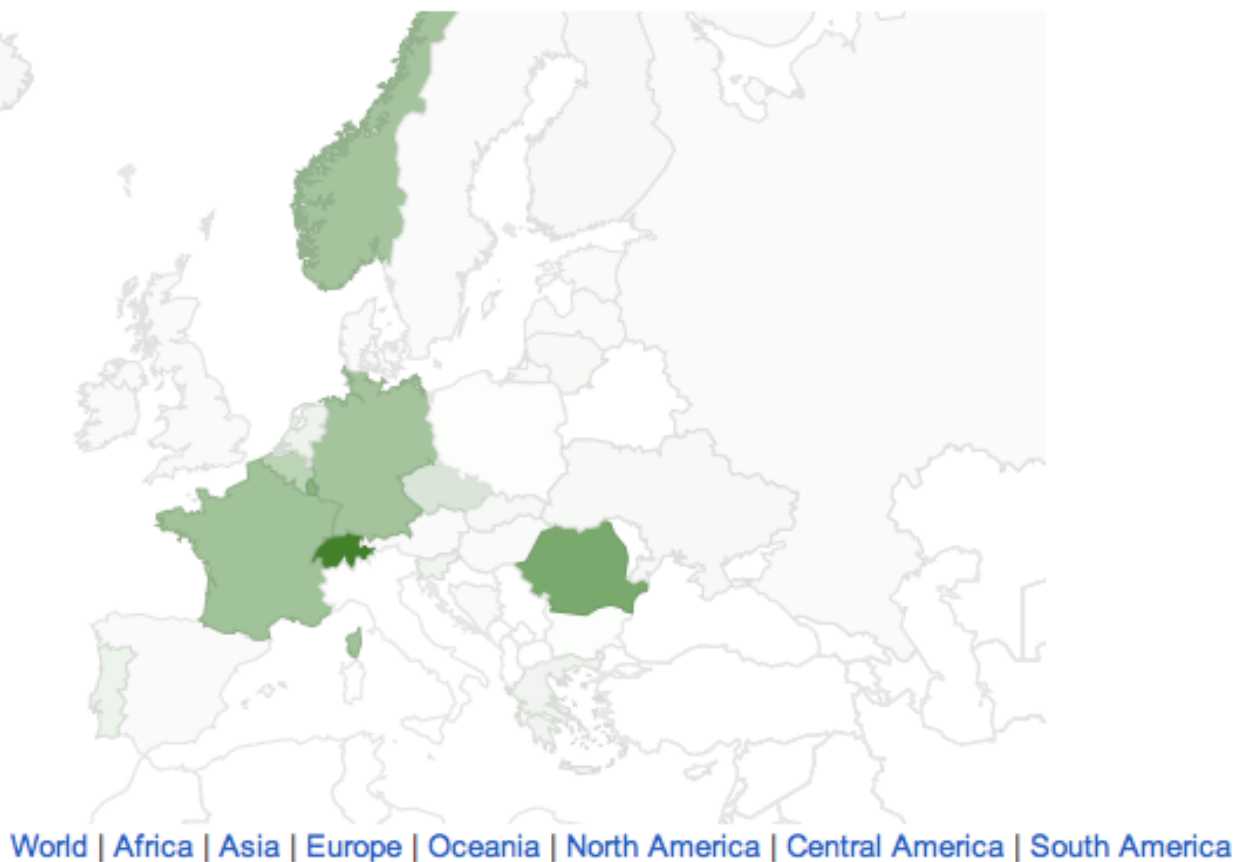
<http://www.google.nl/ipv6/statistics.html>



Google per-country IPv6 adoption

Per-Country IPv6 adoption

Switzerland
11.4%



IPv4 & IPv6 at CERN ...

CERN – “Current VMs adoption plan may cause IPv4 depletion during 2014” (E. Martelli, Jan 2013)

IPv6 deployment at CERN

CHEP2013 talk by David GUTIERREZ RUEDA – “Network architecture and IPv6 deployment at CERN”

2012

Network Database: Schema and Data IPv6 Ready

Admin Web: IPv6 integrated

Configuration Manager supports IPv6 routing

2013

Gradual deployment on the routing infrastructure starts

The Data Center is Dual-Stack

NTPv6 and DNSv6

DHCPv6

Today



- Infrastructure is Dual-Stack
- Firewallv6 automated configuration
- User Web and SOAP integrate IPv6
- Automatic DNS AAAA configuration



The HEPiX IPv6 Working Group ...

HEPiX IPv6 Working Group

Created in 2011 with aims:

- Consider whether/how IPv6 should be deployed in HEP
 - especially WLCG (Worldwide Large Hadron Collider Grid)
- Readiness and Gap analysis
- HEP applications, middleware, security issues, system management and monitoring tools, end to end network monitoring tools
- Run a distributed HEP testbed
 - to help explore all the above issues
- We meet face to face 3 or 4 times a year
 - And by video conference in between

News since CHEP2012

- IPv4 addresses running out at CERN!
- New testbed sites
 - IHEP (CN), Glasgow, Imperial London, PIC, USLHCNet Caltech (Chicago), QMUL & FNAL
- More engagement of LHC Experiments
 - CMS (in from the start)
 - LHCb next (although not so active right now)
 - ATLAS and ALICE now engaged
- Lots of testing (data transfer, monitoring, jumbo frames, ...)
- WLCG Ops Team IPv6 task force created – we work together
 - Defining use cases for testing and performing software survey
- Test of dual-stack services at Imperial London Tier 2

The IPv6 testbed ...

The HEPiX IPv6 Testbed

- A dedicated distributed testbed
- Connected to IPv6 and IPv4 networks
 - IPv6-only/IPv4-only names also registered in DNS
 - e.g. hepix-v6.desy.de & hepix-v4.desy.de
- Caltech, CERN, Chicago, DESY, FNAL, FZU, Glasgow, Imperial, IHEP, INFN, KIT, PIC
- <http://hepixonweb.cern.ch/content/hepixonweb-distributed-testbed>

IPv6 file transfer tests

- Tony Wildish (CMS)
- Simple mesh of simultaneous data transfers (IPv6)
- Transfers a 1 GB file using GridFTP (globus-url-copy)
 - Measures time to transfer
 - Records any errors
- Uses UberFTP to confirm arrival and then delete
- Have been running for last 6 months

GridFTP IPv6 data transfer mesh

- > 2 PB data transferred in last 6 months
- Success rate > 87%
- Very High!
 - Operating at risk

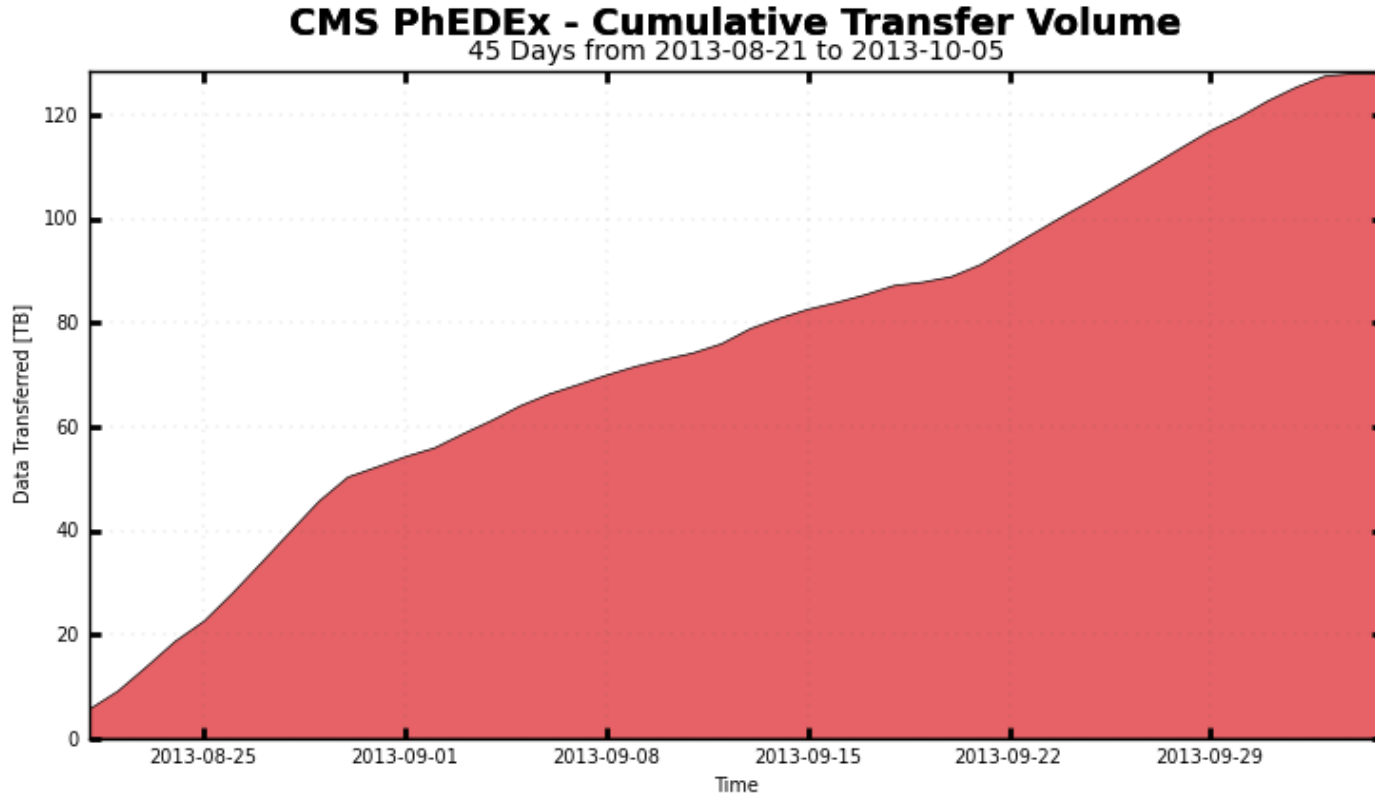


Time to transfer max
500 s

CMS PhEDEx/FTS3/DPM

- Dual-stack FTS3 server deployed at Imperial
- IPv6-only DPM SE's at Imperial and Glasgow
- CMS transfer data using PhEDEx (IPv6)
- > 120 TB data transferred in 2 months
- Conclusion – “PhEDEx can operate to CMS production standards with IPv6-enabled services”

CMS PhEDEx transfers



■ T2_IPv6_Glasgow

Total: 128.00 TB, Average Rate: 0.00 TB/s

Data transfer lessons learned

- CMS PhEDEx is production ready on IPv6
- The testing is very useful for finding site IPv6 problems
 - E.g. allows some level of stress testing
- Over the months we have experienced some breaks, performance problems, asymmetric routing, ...
 - E.g. cpu utilisation border router reaches 60% with IPv6 traffic flows of 200-300 Mbps
- The reason for ongoing data transfer tests
- It is useful for a site to be in the IPv6 testbed

Dual-Stack services at Imperial

- Tier 2 Imperial College, London
- Configured a subset of their service to be IPv4/IPv6 dual stack
 - DNS, SSH, NFS, EMI 2 and EMI 3 CREAM CEs, EMI 2 Worker Nodes, ARC CE and dCache (headnode, SRM component only), all BDII services
 - Stateless address autoconfiguration (SLAAC)
- No problems observed
 - Did not need to turn IPv6 off



Software and Tools survey ...

Software & Tools IPv6 Survey

- An “Asset” survey is underway
 - Need to reach out to all sites and the LHC experiments
 - Includes **all** applications, middleware and tools
- If IPv6-readiness is known, can be recorded
- Otherwise we will need to investigate further
 - Ask developer and/or supplier
 - Scan source code or look for network calls while running
 - Test the running application under dual stack conditions

ARGUS	Middleware	ALICE, ATLAS, CMS, LHCb		Unknown
BDII	Middleware	ATLAS, CMS, LHCb	EMI 2	Unknown
BestMAN	Middleware	ATLAS, CMS		Unknown
CASTOR	Middleware	ALICE, ATLAS, CMS, LHCb		Unknown
cfengine	Monitoring			Unknown
CMS Tag Collector	LHC Experiment Application	CMS		Unknown
CMSSW	LHC Experiment Application	CMS		Unknown
cmsweb	LHC Experiment Application	CMS		Unknown
Condor	Middleware	ATLAS, CMS		Unknown
Cream CE	Middleware	ALICE, ATLAS, CMS, LHCb	EMI 2	Unknown
CVMFS	Other Application	ALICE, ATLAS, CMS, LHCb		Unknown
Dashboard Google Earth	Monitoring	ALICE, ATLAS, CMS, LHCb		Claimed
DBS	LHC Experiment Application	CMS		Unknown
dCache	Middleware	ALICE, ATLAS, CMS, LHCb	2.6.6	Claimed
dCache	Middleware	ALICE, ATLAS, CMS, LHCb	1.9.12	NO
DIRAC	LHC Experiment Application	LHCb		Unknown
DPM	Middleware	ALICE, ATLAS, CMS, LHCb		Claimed
EGI Accounting Portal	Monitoring	ALICE, ATLAS, CMS, LHCb		Unknown

Software survey (2)

Experiment Dashboards	Monitoring	ALICE, ATLAS, CMS, LHCb		Unknown
Frontier	LHC Experiment Application	ATLAS, CMS		Unknown
FTS	Middleware	ATLAS, CMS, LHCb	FTS 3	Claimed
Ganglia	Monitoring			Unknown
GFAL/lcg_util	Middleware	ALICE, ATLAS, CMS, LHCb		Claimed
glideinWMS	Middleware	CMS		Unknown
GOCdb	System Management Tool	ALICE, ATLAS, CMS, LHCb		Unknown
Gratia Accounting	Monitoring			Unknown
Gridsite	Middleware			Claimed
GridView	Monitoring			Unknown
Gstat	Monitoring			Unknown
iCMS	LHC Experiment Application	CMS		Unknown
LFC	Middleware	ATLAS, LHCb	EMI 2	Unknown
MonALISA	Monitoring			Claimed
MyOSG	Monitoring			Unknown
MyProxy	Middleware		EMI 1	Unknown
MyWLCG	Monitoring	ALICE, ATLAS, CMS, LHCb		Unknown

Software survey (3)

Nagios	Monitoring	ALICE, ATLAS, CMS, LHCb		Claimed
OpenAFS	Other Application	ALICE, ATLAS, CMS, LHCb	Last	NO
PanDA	LHC Experiment Application	ATLAS, CMS		Unknown
perfSONAR	Monitoring	ATLAS, CMS		Claimed
PhEDEx	LHC Experiment Application	CMS		YES
REBUS	Monitoring	ALICE, ATLAS, CMS, LHCb		Unknown
SAM	Monitoring	ALICE, ATLAS, CMS, LHCb		Unknown
Scientific Linux	Operating System	ALICE, ATLAS, CMS, LHCb	5.7	YES
SDT IB and QA pages	LHC Experiment Application	CMS		Unknown
StoRM	Middleware	ALICE, ATLAS, CMS, LHCb		Unknown
Ticket system -- GGUS	System Management Tool	ALICE, ATLAS, CMS, LHCb		Unknown
various D web tools	LHC Experiment Application	CMS		Unknown
VOMS	Middleware	ALICE, ATLAS, CMS, LHCb	EMI-2	Unknown
WMAgent	LHC Experiment Application	CMS		Unknown
WMS	Middleware	ALICE, ATLAS, CMS, LHCb	EMI 2	Unknown
xroot	Middleware	ALICE, ATLAS, CMS	3.1.0	Unknown

Future plans and next steps ...

Use Cases for testing – dual stack services

- <http://hepixon.web.cern.ch/content/wlcg-ipv6-task-force-0>
- Basic job submission
 - The user submits a job using the native middleware clients (CREAM client, Condor-G, etc.) or intermediate services (gLite WMS, glideinWMS, PanDA, DIRAC, AliEN, etc.)
- Basic data transfer
 - The user copies a file from his node to a SE and back
- Third party data transfer
 - The user replicates a bunch of files between sites via FTS-3
- Production data transfer
 - The user replicates a dataset using experiment-level tools (PhEDEx, DDM, DIRAC, etc.)
- Conditions data
 - A job access conditions data from a batch node via Frontier/squid
- Experiment software
 - A job accesses experiment software in CVMFS from a batch node

Use cases (2)

- Experiment workflow
 - A user runs a real workflow (event generation, simulation, reprocessing, analysis)
 - This test combines all previous tests into one
- Information system
 - A user queries the information system
- Job monitoring
 - Monitoring information from jobs, coming either from central services or from batch nodes via messaging systems, is collected, stored and accessed by a user

TESTING plan

- all services must be tested in dual stack
- all user clients must be tested in IPv4 and dual stack
- all batch nodes must be tested in IPv4, IPv6 and dual stack

Testing 2013-14

- Systematic tests of the Use Cases
- Try more (a small number!) Tier 2 sites with dual stack services
- Glasgow, KIT, CERN are deploying larger scale test-clusters
 - Use these for more complete testing
- Turn on dual-stack services as soon as possible

Timetable WLCG IPv6 transition

- In 2012 we said:
 - Support for IPv6-only clients *not before* Jan 2014
- Still true!
 - And likely to be **much** later
 - Needs MANY sites to support IPv6
- Sysadmins, Security staff, Monitoring and Operations
 - Training required
 - New operational procedures
- The WG will continue to test Use Cases
- Aim for dual-stack on most/many services – when?
- Aiming for an IPv6 workshop at CERN in Spring/Summer 2014

Further info

- HEPiX IPv6 web

<http://hepixonweb.cern.ch>

- HEPiX IPv6 wiki

<https://w3.hepixon.org/ipv6-bis/>

- 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>

Summary

- Good progress has been made by the HEPiX IPv6 working group with proving that dual stack IPv4/IPv6 services work in WLCG
 - But there is lots more to do!
- **We are very much effort-limited**
 - Tier 1 and Tier 2 sites
 - Experiments
- We do need to encourage sites to deploy IPv6
 - The 2014 workshop should help here
- VOLUNTEERS to join the group are always welcome
 - please contact me!

Other IPv6 papers - CHEP2013

#30: Network architecture and IPv6 deployment at CERN (Talk)

- GUTIERREZ RUEDA, David et al

#356: Enabling IPv6 at FZU - WLCG Tier2 in Prague (Poster)

- KOUBA, Tomas et al

#1: A Voyage to Arcturus (Poster)

- MITCHELL, Mark et al

#306: Deploying an IPv6-enabled grid testbed at GridKa (Poster)

- HOEFT, Bruno Heinrich et al



Questions?