The Deployment of Dual-stack IPv6/IPv4 Services on WLCG

David Kelsey (STFC-RAL)
ISGC 2015 Taipei
18 Mar 2015
HEPiX IPv6 Working Group

• Members of the group

J. Bernier (IN2P3), S. Campana (CERN), K. Chadwick (FNAL), J. Chudoba (FZU), A. Dewhurst (RAL), M. Elias (FZU), S. Fayer (Imperial), T. Finnern (DESY), C. Grigoras (CERN), T. Hartmann (KIT), B. Hoeft (KIT), T. Idiculla (RAL), D. Kelsey (RAL), F. Lopez Munoz (PIC), E. MacMahon (Oxford), E. Martelli (CERN), R. Nandakumar (RAL), K. Ohrenberg (DESY), F. Prelz (INFN), D. Rand (Imperial), A. Sciaba (CERN), U. Tigerstedt (CSC), R. Voicu (Caltech), C. Walker (QMUL), T. Wildish (Princeton)

And many others in earlier times
Outline

- Use of IPv6 & status of IPv4 address space
- WLCG IPv6 site readiness survey
- IPv6 testing and data transfers
- Software readiness
- Experiment plans
- Monitoring – PerfSONAR
- Deployment of dual-stack services
IPv6 growth
(global Google clients)

Globally now > 5% use IPv6 address

Largest national use:
Belgium now > 30% use IPv6 address
% Networks with IPv6
IPv4 address exhaustion

IANA Unallocated Address Pool Exhaustion:
03-Feb-2011

Projected RIR Address Pool Exhaustion Dates:

<table>
<thead>
<tr>
<th>RIR</th>
<th>Projected Exhaustion Date</th>
<th>Remaining Addresses in RIR Pool (/32s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>APNIC</td>
<td>19-Apr-2011 (actual)</td>
<td>0.7576</td>
</tr>
<tr>
<td>RIPE NCC</td>
<td>14-Sep-2012 (actual)</td>
<td>1.0284</td>
</tr>
<tr>
<td>LACNIC</td>
<td>10-Jun-2014 (actual)</td>
<td>0.1865</td>
</tr>
<tr>
<td>ARIN</td>
<td>01-Jun-2015</td>
<td>0.3116</td>
</tr>
<tr>
<td>AFRINIC</td>
<td>10-Jan-2019</td>
<td>2.6332</td>
</tr>
</tbody>
</table>

Projection of consumption of Remaining RIR Address Pools

This report generated at 14-Mar-2015 08:07 UTC.
Outline

• Use of IPv6 & status of IPv4 address space
• WLCG IPv6 site readiness survey
• IPv6 testing and data transfers
• Software readiness
• Experiment plans
• Monitoring – PerfSONAR
• Deployment of dual-stack services
WLCG Sites - IPv6 readiness

- Survey in summer 2014
- Live wiki table at https://www.gridpp.ac.uk/wiki/2014_IPv6_WLCG_Site_Survey
- Received responses from most Tier 1s and 60% of Tier 2s
  - Assume the other 40% Tier 2s have not yet started with IPv6!
- Tier 1s: ~75% are IPv6 ready. Rest within 1 year
- Tier 2s: ~20% are IPv6 ready. ~20% within 1-2 year
  - ~ 60% have no plans yet
- ~10% of sites report lack of IPv4 addresses (in next 1-2 years)
  - Especially for Virtual Machines or Containers
- This is a challenge!
IPv6 at UK Universities.
Example - my home country

http://www.mrp.net/ipv6_survey/
Outline

• Use of IPv6 & status of IPv4 address space
• WLCG IPv6 site readiness survey
• IPv6 testing and data transfers
• Software readiness
• Experiment plans
• Monitoring – PerfSONAR
• Deployment of dual-stack services
IPv6 testbed – data transfers

- Caltech, CERN, DESY, FNAL, FZU, Glasgow, IHEP, KIT, INFN, Imperial, Lyon, NDGF, PIC, Wisconsin
- GridFTP tests over IPv6
  - Running for 706 days (turned off on 2 Feb 2015)
    - Total data transferred ~ 6.8 PB (10 TB per day)
  - Each site transfers a 1GB file to every other site using globus-url-copy over IPv6
    - 3rd-party transfer launched from CERN
  - Filesize is checked and file removed using uberftp/IPv4
- Useful for: Stability & Verifying network routes etc.

18/03/2015

WLCG IPv6
Data transfers – last 100 days

Time to transfer (seconds)
More data transfer tests

• Dual-stack FTS3 server at Imperial
  – DPM storage elements at Imperial and Glasgow
  – Transfers via CMS PhEDEx
  – Worked well

• Now deploying larger FTS3 testbed with more storage technologies
  – Not using PhEDEx agents
  – Add dCache endpoints
  – Other technologies later (StoRM, XRootD)
Dual-stack production services

• Several sites in the IPv6 working group have been testing dual-stack services in production
  – Including storage endpoints

• To check
  – Does the service over IPv4 continue to work?
  – Does it bind to both IPv4 and IPv6?
  – Is IPv6 preferred by default?
  – Can the choice be configured?

• We need more of this controlled testing
  – To date there have been no big problems
Outline

• Use of IPv6 & status of IPv4 address space
• WLCG IPv6 site readiness survey
• IPv6 testing and data transfers
• **Software readiness**
  • Experiment plans
  • Monitoring – PerfSONAR
• Deployment of dual-stack services
<table>
<thead>
<tr>
<th>Software Component</th>
<th>Type</th>
<th>Used by Experiment</th>
<th>Version</th>
<th>IPv6 Compliance</th>
</tr>
</thead>
<tbody>
<tr>
<td>AliEN</td>
<td>LHC Experiment Application</td>
<td>ALICE</td>
<td></td>
<td>YES</td>
</tr>
<tr>
<td>ARC CE</td>
<td>Middleware</td>
<td>ATLAS, CMS</td>
<td></td>
<td>YES</td>
</tr>
<tr>
<td>ARGUS</td>
<td>Middleware</td>
<td>ALICE, ATLAS, CMS, LHCb</td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td>BDII</td>
<td>Middleware</td>
<td>ATLAS, CMS, LHCb</td>
<td>EMI 2</td>
<td>YES</td>
</tr>
<tr>
<td>BeetMAN</td>
<td>Middleware</td>
<td>ATLAS, CMS</td>
<td></td>
<td>YES</td>
</tr>
<tr>
<td>CASTOR</td>
<td>Middleware</td>
<td>ALICE, ATLAS, CMS, LHCb</td>
<td>NO</td>
<td></td>
</tr>
<tr>
<td>cLengine</td>
<td>Monitoring</td>
<td></td>
<td></td>
<td>Unknown</td>
</tr>
<tr>
<td>CMS Tag Collector</td>
<td>LHC Experiment Application</td>
<td>CMS</td>
<td></td>
<td>Unknown</td>
</tr>
<tr>
<td>CMSSW</td>
<td>LHC Experiment Application</td>
<td>CMS</td>
<td></td>
<td>Unknown</td>
</tr>
<tr>
<td>cmsweb</td>
<td>LHC Experiment Application</td>
<td>CMS</td>
<td></td>
<td>YES</td>
</tr>
<tr>
<td>CRAB 2</td>
<td>LHC Experiment Application</td>
<td>CMS</td>
<td></td>
<td>Unknown</td>
</tr>
<tr>
<td>Cream CE</td>
<td>Middleware</td>
<td>ALICE, ATLAS, CMS, LHCb</td>
<td>YES</td>
<td></td>
</tr>
<tr>
<td>CVMFS</td>
<td>Other Application</td>
<td>ALICE, ATLAS, CMS, LHCb</td>
<td>YES</td>
<td></td>
</tr>
<tr>
<td>Dashboard Google Earth</td>
<td>Monitoring</td>
<td>ALICE, ATLAS, CMS, LHCb</td>
<td>Claimed</td>
<td></td>
</tr>
<tr>
<td>dCache</td>
<td>Middleware</td>
<td>ALICE, ATLAS, CMS, LHCb</td>
<td>NO</td>
<td></td>
</tr>
<tr>
<td>dCache</td>
<td>Middleware</td>
<td>ALICE, ATLAS, CMS, LHCb</td>
<td>YES with caveats</td>
<td></td>
</tr>
<tr>
<td>dCache</td>
<td>Middleware</td>
<td>ALICE, ATLAS, CMS, LHCb</td>
<td>NO</td>
<td></td>
</tr>
</tbody>
</table>
## Software readiness (2)

<table>
<thead>
<tr>
<th>Service</th>
<th>Type</th>
<th>Applications</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>DIrAC</td>
<td>LHC Experiment App.</td>
<td>LHCb</td>
<td>Unknown</td>
</tr>
<tr>
<td>DPM</td>
<td>Middleware</td>
<td>ALICE, ATLAS, CMS, LHCb</td>
<td>YES with caveats</td>
</tr>
<tr>
<td>EGI Accounting</td>
<td>Monitoring</td>
<td>ALICE, ATLAS, CMS, LHCb</td>
<td>Unknown</td>
</tr>
<tr>
<td>Accounting Portal</td>
<td>Monitoring</td>
<td>ALICE, ATLAS, CMS, LHCb</td>
<td>NO</td>
</tr>
<tr>
<td>EOS</td>
<td>Middleware</td>
<td>ALICE, ATLAS, CMS, LHCb</td>
<td>Unknown</td>
</tr>
<tr>
<td>Experiment Dashboards</td>
<td>Monitoring</td>
<td>ALICE, ATLAS, CMS, LHCb</td>
<td>Unknown</td>
</tr>
<tr>
<td>Frontier</td>
<td>LHC Experiment App.</td>
<td>ATLAS, CMS</td>
<td>NO</td>
</tr>
<tr>
<td>FTS</td>
<td>Middleware</td>
<td>ATLAS, CMS, LHCb</td>
<td>YES</td>
</tr>
<tr>
<td>Ganglia</td>
<td>Monitoring</td>
<td></td>
<td>YES</td>
</tr>
<tr>
<td>QFAL/cg_util</td>
<td>Middleware</td>
<td>ALICE, ATLAS, CMS, LHCb</td>
<td>YES with caveats</td>
</tr>
<tr>
<td>glideinWMS</td>
<td>Middleware</td>
<td>CMS</td>
<td>Claimed</td>
</tr>
<tr>
<td>globus toolkit</td>
<td>Middleware</td>
<td>ALICE, ATLAS, CMS, LHCb</td>
<td>YES</td>
</tr>
<tr>
<td>GOCdb</td>
<td>System Management Tool</td>
<td>ALICE, ATLAS, CMS, LHCb</td>
<td>Unknown</td>
</tr>
<tr>
<td>Gratia Accounting</td>
<td>Monitoring</td>
<td></td>
<td>Unknown</td>
</tr>
<tr>
<td>Gridsite</td>
<td>Middleware</td>
<td></td>
<td>YES with caveats</td>
</tr>
<tr>
<td>Gstat</td>
<td>Monitoring</td>
<td></td>
<td>Unknown</td>
</tr>
<tr>
<td>GlueMS</td>
<td>Middleware</td>
<td>ATLAS, CMS</td>
<td>Unknown</td>
</tr>
<tr>
<td>Hadoop</td>
<td>Middleware</td>
<td>ATLAS, CMS</td>
<td>NO</td>
</tr>
<tr>
<td>HammerCloud</td>
<td>LHC Experiment App.</td>
<td>ATLAS, CMS, LHCb</td>
<td>Unknown</td>
</tr>
</tbody>
</table>
## Software readiness (3)

<table>
<thead>
<tr>
<th>Software</th>
<th>Type</th>
<th>Applications</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>HTCCondor</td>
<td>Middleware</td>
<td>ATLAS, CMS</td>
<td>YES with caveats</td>
</tr>
<tr>
<td>ICMS</td>
<td>LHC Experiment Application</td>
<td>CMS</td>
<td>Unknown</td>
</tr>
<tr>
<td>LFC</td>
<td>Middleware</td>
<td>ATLAS, LHCb</td>
<td>YES</td>
</tr>
<tr>
<td>MonALISA</td>
<td>Monitoring</td>
<td></td>
<td>Claimed</td>
</tr>
<tr>
<td>MyOSG</td>
<td>Monitoring</td>
<td></td>
<td>Unknown</td>
</tr>
<tr>
<td>MyProxy</td>
<td>Middleware</td>
<td>EMI 1</td>
<td>Unknown</td>
</tr>
<tr>
<td>MySQL / MariaDB</td>
<td>Other Application</td>
<td>ALICE, ATLAS, CMS, LHCb</td>
<td>YES with caveats</td>
</tr>
<tr>
<td>MyWLCG</td>
<td>Monitoring</td>
<td>ALICE, ATLAS, CMS, LHCb</td>
<td>Unknown</td>
</tr>
<tr>
<td>Nagios</td>
<td>Monitoring</td>
<td>ALICE, ATLAS, CMS, LHCb</td>
<td>Claimed</td>
</tr>
<tr>
<td>OpenAFS</td>
<td>Other Application</td>
<td>ALICE, ATLAS, CMS, LHCb</td>
<td>NO</td>
</tr>
<tr>
<td>PanDA</td>
<td>LHC Experiment Application</td>
<td>ATLAS, CMS</td>
<td>Unknown</td>
</tr>
<tr>
<td>perfSONAR</td>
<td>Monitoring</td>
<td>ATLAS, CMS</td>
<td>YES</td>
</tr>
<tr>
<td>PhEDEx agents</td>
<td>LHC Experiment Application</td>
<td>CMS</td>
<td>YES</td>
</tr>
<tr>
<td>Puppet</td>
<td>System Management Tool</td>
<td>ALICE, ATLAS, CMS, LHCb</td>
<td>YES</td>
</tr>
<tr>
<td>REBUS</td>
<td>Monitoring</td>
<td>ALICE, ATLAS, CMS, LHCb</td>
<td>Unknown</td>
</tr>
<tr>
<td>SAM</td>
<td>Monitoring</td>
<td>ALICE, ATLAS, CMS, LHCb</td>
<td>Unknown</td>
</tr>
<tr>
<td>Scientific Linux</td>
<td>Operating System</td>
<td>ALICE, ATLAS, CMS, LHCb</td>
<td>YES</td>
</tr>
<tr>
<td>SDT IB and QA pages</td>
<td>LHC Experiment Application</td>
<td>CMS</td>
<td>Unknown</td>
</tr>
</tbody>
</table>
## Software readiness (4)

<table>
<thead>
<tr>
<th>Software</th>
<th>Description</th>
<th>Applications</th>
<th>Version</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Squid</td>
<td>Other Application</td>
<td>ALICE, ATLAS, CMS, LHCb</td>
<td>2.8</td>
<td>NO</td>
</tr>
<tr>
<td>Squid</td>
<td>Other Application</td>
<td>ALICE, ATLAS, CMS, LHCb</td>
<td>3.1</td>
<td>YES</td>
</tr>
<tr>
<td>StoRM</td>
<td>Middleware</td>
<td>ALICE, ATLAS, CMS, LHCb</td>
<td>1.11.2</td>
<td>YES with caveats</td>
</tr>
<tr>
<td>Ticket system -- GGUS</td>
<td>System Management Tool</td>
<td>ALICE, ATLAS, CMS, LHCb</td>
<td></td>
<td>Unknown</td>
</tr>
<tr>
<td>various D web tools</td>
<td>LHC Experiment Application</td>
<td>CMS</td>
<td></td>
<td>Unknown</td>
</tr>
<tr>
<td>VOMS</td>
<td>Middleware</td>
<td>ALICE, ATLAS, CMS, LHCb</td>
<td>2.0.11</td>
<td>YES</td>
</tr>
<tr>
<td>VOMS-Admin</td>
<td>Middleware</td>
<td>ALICE, ATLAS, CMS, LHCb</td>
<td>3.3.1</td>
<td>YES</td>
</tr>
<tr>
<td>WLCG squid monitoring</td>
<td>Monitoring</td>
<td>ATLAS, CMS</td>
<td></td>
<td>Claimed</td>
</tr>
<tr>
<td>WMAgent</td>
<td>LHC Experiment Application</td>
<td>CMS</td>
<td></td>
<td>Unknown</td>
</tr>
<tr>
<td>WMS</td>
<td>Middleware</td>
<td>ALICE, ATLAS, CMS, LHCb</td>
<td>EMI-2, EMI-3</td>
<td>YES with caveats</td>
</tr>
<tr>
<td>xrootd</td>
<td>Middleware</td>
<td>ALICE, ATLAS, CMS, LHCb</td>
<td>4.0.0</td>
<td>YES with caveats</td>
</tr>
<tr>
<td>xrootd</td>
<td>Middleware</td>
<td>ALICE, ATLAS, CMS</td>
<td>3.1.0</td>
<td>NO</td>
</tr>
</tbody>
</table>
IPv6 & “Data”

- **DPM** – Yes – need config changes & MySQL 5.5+
- **dCache** – Yes – V2.9.x (2.9.4 several fixes)
- **StoRM** – Yes – need config changes – V1.11.2 tested
- **XRootD** – Yes – V4.0.0 (dual-stack), V4.1+ (IPv6 only)
- **FTS** – Yes – Version 3 and beyond
- **CASTOR** - No
- **EOS** – No – but will when built on xRootD 4

Many IPv6 improvements during last year
Outline

• Use of IPv6 & status of IPv4 address space
• WLCG IPv6 site readiness survey
• IPv6 testing and data transfers
• Software readiness
• Experiment plans
• Monitoring – PerfSONAR
• Deployment of dual-stack services
LHC Experiments & IPv6

• Good engagement with the Experiments in 2014
  – Very important
  – Testing should be driven by their use cases
  – But they are now busy preparing for LHC Run 2

• Important IPv6-only use case to address
  – Virtual Machines/Worker Nodes – IPv6-only (IPv4 gone!)
  – Must be able to access federated storage services
  – Storage services must be dual-stack
  – And central services too (VOMS, catalogues, info etc)
IPv6 Requirements

ATLAS (GDB Oct 2014)

• Request that all Tier 1s provide a dual stack PerfSONAR machine by April 2015
• Request that T2Ds provide a dual stack PerfSONAR machine by August 2015
• All to get sites ready to host dual-stack data services (FAX)

CMS (GDB Sep 2014)

• Request to have a substantial fraction of CMS data accessible via AAA through IPv6 by end 2015 - to use from IPv6-only WN
Experiments (2)

• ALICE runs dual-stack central services (works)
  – testing all their frameworks
  – Waiting to test XRootD V4.1

• ATLAS
  – Dual-stack pilot to Dual-stack CE (works)
  – Will test BigPanda instance
  – Testing dual-stack Squid 3
Experiments (3)

• CMS
  – Test dual-stack glideinWMS
    • Does not prefer IPv6 at moment – takes first in DNS
  – test AAA with IPv6
    • Nebraska enabling IPv6 access to their XRootD

• LHCb job submission
  – All DIRAC communication needs to be tested
  – Tests started on lxplus-ipv6 (dual-stack)
    • Works for IPv4
  – About to test on DIRAC v6r13 (should work on dual-stack)
Outline

• Use of IPv6 & status of IPv4 address space
• WLCG IPv6 site readiness survey
• IPv6 testing and data transfers
• Software readiness
• Experiment plans
• Monitoring – PerfSONAR
• Deployment of dual-stack services
IPV6 Network Monitoring

- **PerfSONAR** monitoring
- As we turn on production dual-stack
  - We need to collect monitoring data
- A good way for a new site to get started
  - add to PerfSONAR dashboard when ready
- Since V3.4 Dual-stack PerfSONAR is recommended
  - Add AAAA record to DNS
  - If both ends are dual-stack
    - IPv4 & IPv6 tests will be run
- **ATLAS**: Tier 1 by April 2015, T2Ds by August 2015
PerfSONAR dashboard

http://maddash.aglt2.org/
IPv4
6 Gbps

Oxford -> QMUL

IPv6 connection at Oxford has less bandwidth

IPv6
0.6 Gbps
Outline

- Use of IPv6 & status of IPv4 address space
- WLCG IPv6 site readiness survey
- IPv6 testing and data transfers
- Software readiness
- Experiment plans
- Monitoring – PerfSONAR
- Deployment of dual-stack services
Deployment - dual-stack services

• PerfSONAR – dual-stack monitoring
• Gradually deploy dual-stack data services during 2015
  – When testing shows functionality and performance is sufficient
• Important central services to dual-stack
• To support worker nodes/virtual machines which are IPv6-only (by end 2015)
Lessons to other communities

• Moving to IPv6 takes time
  – Need to consider all software, tools and management/operations
  – Staff need training

• Starting with a distributed testbed allows work to start
  – Participants can share experiences
  – And learn together
Links

• HEPiX IPv6 web
  http://hepix-ipv6.web.cern.ch

• HEPiX IPv6 wiki
  https://w3.hepix.org/ipv6-bis/

• 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

• Paper published in proceedings of CHEP2013
Pre-GDB IPv6 workshop

• 10 June 2014 at CERN
• https://indico.cern.ch/event/313194/
• IPv6 technical background
• IPv6 at CERN
• File transfer testing – the IPv6 testbed activities
• Status and configuration of some services
• Experiment testing and plans
• Monitoring
• Site status and experiences
• Next steps – discussion
Questions?