

IPv6 only CPU Deployment Status and Plans

Alastair Dewhurst,

IPv6 only CPU plan

WLCG IPv6 only CPU deployment strategy

HEPiX IPv6 Working group

January 21, 2016

Executive Summary

This document describes the Worldwide LHC Computing Grid's (WLCG) strategy to allow sites to provide IPv6 only CPU resources to the VOs.

- Sites can provide IPv6 only CPU resources from April 2017 onwards.
- Sites will not be required to provide dual stack interfaces (e.g. CEs and Squids) for their CPU resources.
- The VO infrastructure (e.g. central services provided by VOs) must provide an equal quality of service to both IPv4 and IPv6 resources.
- An IPv6 only CPU resource will need access to (dual stack) storage. Several VOs have federated data access models. We need to provide incentives for sites to provide dual stack storage access.

<https://github.com/stfc/IPv6>

Input from CMS and LHCb expected soon!

Don't expect large uptake, <1% pledged resources but does allow sites to plan

Entire services should be made dual stack

TI will need to play leading role



Tier I requirements

- CERN + Tier I will need to play a crucial role interfacing between current IPv4 only sites and new IPv6 only sites.
- Have been asked for a long time to provide dual stack perfonar box.
 - Ticket them now?
 - Regular review of performance at meeting?
- Those providing CVMFS Stratum I to make dual stack by April 2017.
- New: Assume all other external services (e.g. Bdi, GGUS, GOCDDB) should be made dual stack by April 2017.
 - If not possible, there should be a decommission plan for service.
- New: Should provide functional dual stack access to storage:
 - 1GB/s and 90% availability by April 2017.
 - 10Gb/s and 95% availability by April 2018.



Other requirements

- We need to have an IPv6 only SAM box which will submit the CE style tests to sites.
- We need CA to provide IPv6 access.
 - Or we require sites to run their own dual stack squid.
- GOCDB / OIM needs to be dual stack.
 - Dual stack GOCDB being tested at RAL.
- Accounting.
 - Talked to APEL, need to make message brokers dual stack.
- RPM repositories



ATLAS status



ATLAS IPv6 only CPU

6

- Assuming site has access to a dual stack storage. WN will talk to the following central nodes:
- For Panda:
 - Production Panda Servers: `aipanda03[0-7].cern.ch`
- For Rucio:
 - Auth nodes : `rucio-auth-prod-0[1,2].cern.ch`
 - Prod nodes through 3 HA proxy frontends `rucio-lb-prod-0[1-3].cern.ch`
- Pilot factories talk to CEs.
- All use `http(s)`.



What do we need to test?

7

- Can a pilot come from APF to WN via IPv6 only?
 - APF → CE → Batch head → WN
- Can a pilot download payload from PanDA via IPv6?
 - WN → PanDA server (http ⇒ should work)
- Can pilot query Rucio and register the output via IPv6?
 - WN → Rucio servers (http ⇒ should work)
- Can pilot upload from WN to SE via IPv6?
- Can a WN get CVMFS data via IPv6?
 - Stratum 1s are currently IPv4 only (http ⇒ should work)
 - What about Frontier queries?
- Can FTS3 transfer to/from RSE via IPv6?



Pilot Testing

- Pilot Factories have been made IPv4 only, which is allowing some sites to deploy dual stack CEs.
- <https://its.cern.ch/jira/browse/CSOPS-796>
- Dev Pilot Factory is dual stack
- <https://its.cern.ch/jira/browse/ASPDA-56>
- Test pilots appear to be working over IPv6 between dev pilot factory and dual stack CEs.
- http://apfmon.lancs.ac.uk/aipanda020-ipv6dev/UKI-LT2-QMUL_TEST-ce04ipv6

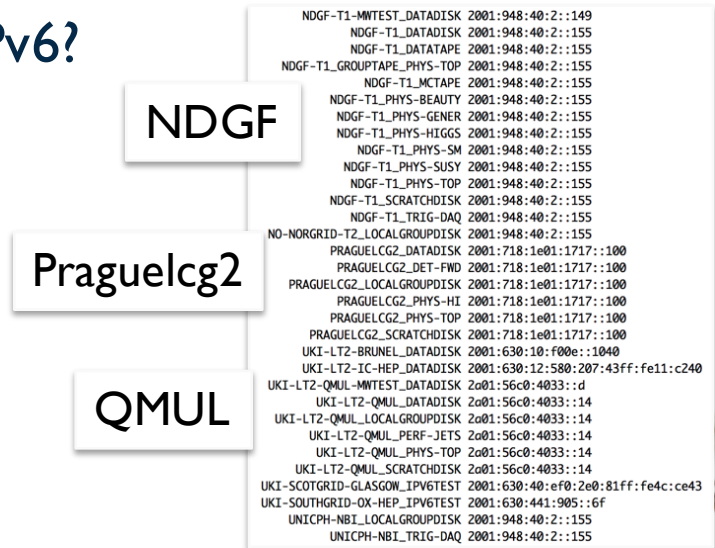
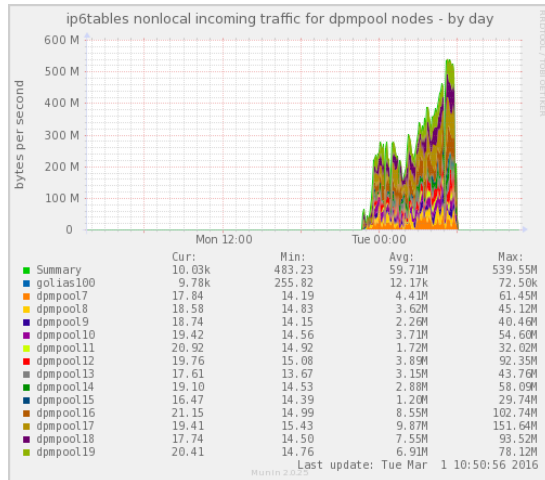


Usage!

- Can I transfer ATLAS data (with Rucio) over IPv6 network?
- Need at least two sites with dual stack storage
- FTS3 pilot service at CERN supports IPv6

```
rucio add-rule --lifetime 286400 --source-replica-expression NDGF-TI_DATADISK
data15_I3TeV.00267360.physics_MinBias.recon.ESD.r7566_tid07740236_00 I PRAGUELCG2_SCRATCHDISK
```

- Did the data really go over IPv6?



Backup



Broken dual stack

11

- The problem was first spotted during Pilot Factory testing:
 - <https://its.cern.ch/jira/browse/ASPDA-177>
- CVMFS Stratum 0:
 - DNS for cvmfs-stratum-zero.cern.ch had IPv4 and 6 addresses but hosts had no IPv6 interface.
 - CC7 AI image was broken
- Not seen any recent problems so hopefully fixed now.



DNS Round Robin Alias

12

- DNS round robin alias provide a quick and dirty way to spread load across multiple machines.
 - `gethostbyname()` – Used by many IPv4 programs will return host in the order provided (i.e. random).
 - `getaddrinfo()` – Replacement for `gethostbyname()`, can return hosts in a sorted order.
- Depending on how software was written there is a good chance that it will “break” DNS round robin and just use one machine.
 - Observed behaviour in FTS testing.
- Solutions include proper load balancing (HA Proxy, Power DNS) or updating software.

