



GridPP

UK Computing for Particle Physics

IPv6 WAN testing

Duncan Rand and Steve Lloyd

- A fraction of Imperial jobs currently go to the Imperial IPv6-only CE (v6ce00)
- However, they hang whilst trying to contact the LFC at CERN which is IPv4 only
- Could use the Imperial LFC which is now dual-stack
- Jobs would then hang whilst trying to download from any IPv6 only SE's
- Should work with dual-stack SE's however
- Could be interesting to use both IPv4 and IPv6 protocols where they exist and have two tables like the one below one for IPv4 and one for IPv6, much like the perfSONAR testbed
- Do we need a way to indicate when a CE has IPv6-only WN?

GridPP33
Ambleside

Ordered by averages

		Storage Element																					
CE	T1-SE	SE-WN	Cam	Man	Lan	Ox	PPD	RHUL	Birm	Gla	Brun	Shef	Liv	UCL	Bris	Susx	JET	T1	Imp	QMUL	Ed	Ave	CE
QMUL			375.1	130.3	107.6	53.8	431.7	70.5	28.3	22.4	67.1	26.7	15.1	20.7		20.8	4.4					83.1	QMUL
Gla			117.6	126.3	113.8	96.8	56.3	104.5	99.2	119.7	40.9	38.5	36.6	69.0		18.4	4.1					75.0	Gla
Man			95.2	133.8	94.3	100.2	142.6	79.8	91.7	73.5	49.0	46.0	43.9	25.4		16.4	3.4					70.7	Man
Ox			129.7	105.6	118.0	114.8	93.4	79.2	84.9	94.5	59.1	20.3	25.0	35.8		19.5	7.7					69.9	Ox
PPD			95.7	104.9	82.4	101.0	62.4	75.5	105.5	68.5	42.9	64.9	27.2	21.6		20.5						65.2	PPD
RHUL			125.0	87.2	76.5	104.2	32.2	118.9	63.7	83.0	51.5	10.5	31.9	18.3	21.8	20.0	5.2					62.1	RHUL
Cam			101.5	108.0	84.7	93.6	36.8	76.4	57.3	27.4	61.7	53.5	32.0	8.7		14.4						61.8	Cam
Lan			24.5	50.2	43.3	23.5	62.4	22.4	23.8	11.9	19.6	31.2	17.4	11.3		12.9						26.3	Lan
Brun			24.6	33.2	21.8	24.5	26.4	27.7	21.9	6.7	90.0	7.9	4.6	21.0		10.1						23.9	Brun
Liv			10.1		4.3	4.1	11.2	9.8	19.7	6.7	11.3	4.5	132.4			7.9	4.1					22.9	Liv
JET			21.5	26.1	13.3	21.9	27.7	28.0	16.1	19.4	27.3	16.8	19.7	24.8		15.3						21.4	JET
Shef			12.5	28.4	20.7	12.5	16.7	9.1	13.7	18.7	10.4	98.5	10.1	10.9		6.5						20.9	Shef
T1																							T1
Imp																							Imp
UCL																							UCL
Ed																							Ed
Birm																							Birm
Bris																							Bris
Susx																							Susx
Ave			89.5	82.6	64.9	62.8	62.6	58.3	50.5	46.9	44.8	34.7	30.1	25.5	21.8	15.7	4.9					50.5	Ave

Code: 0.0 5.0 10.0 15.0 20.0 25.0 30.0 35.0 40.0 45.0 50.0 MB/s

- In the lead up to LHC Run2 the middleware is changing
- lcg-utils are being phased out
 - to be replaced by gfal2
- SRM likely to be phased out
 - http is a possible replacement
 - ATLAS interested in running an SRM-less Tier-2 and in submitting FTS transfers over http
 - LHCb and ATLAS developing http storage federations
- Xrootd becoming more prominent
 - xrootd federations are now an integral part of ATLAS and CMS computing models
- Also, IPv6 is being rolled out across WLCG and GridPP
- These new clients and protocols need testing and monitoring at every GridPP site over both IPv4 and IPv6

- Decided to test lcg-cp (gsiftp), gfal-copy (using gsiftp), xrdcp (xrootd) and curl (http) over IPv4 and IPv6
- Cloned existing network tests producing a new web page to display the results
- Then implemented a new separate set of tests examining IPv6 and IPv4 explicitly, trying to use gfal2 wherever possible
- Jobs now submitted to IC dual-stack WMS (to be able to submit jobs to IPv6-only CE)
- Files now registered in IC dual-stack LFC (so that IPv6-only WN can communicate with it)
- First (IPv6) problem: question from Steve 'how to upload the files'
- Answer: use the dual-stack UI: lxplus-ipv6.cern.ch

- Would like to be confident we are actually testing the use of IPv4 or IPv6
- Many linux clients offer flags to specify IPv4 or IPv6 explicitly, usually -4 and -6
 - for example curl does offer these flags
 - if nothing set often clients fall back to the alternative
 - but if set -4 or -6 and there's a problem need to fail not fall back
- lcg-cp will use IPv4 unless appropriate environment variables are set
 - If the variables are set but there is a problem may fall back to IPv4
- xrootd does offers e.g. -DSNetworkStack IPv6
- gfal-copy doesn't currently have this option
- Submitted a GGUS ticket requesting this:
https://ggus.eu/index.php?mode=ticket_info&ticket_id=108200
- “So I am afraid this does not seem doable, since only xrootd would actually benefit from it, and other would silently ignored, confusing things even more.

- Replacing lcg-util with gfal2
- Upload files and register in LFC. lcg-lr lists multiple replicas of an lfn:

```
$ lcg-lr lfn://grid/dteam/user/lloyd/data_100mb  
srm://dc2-grid-23.brunel.ac.uk/dpm/brunel.ac.uk/home/dteam/generated/2014-08-29/file49b1ac6d-  
e6de-4d1e-89b8-66defede7734  
srm://t2dpm1-v6.physics.ox.ac.uk/dpm/physics.ox.ac.uk/home/dteam/generated/2014-08-29/  
file53159295-b603-4741-a3f3-535edf360ce8
```

- But gfal-xattr currently only lists one of these replicas:

```
$ gfal-xattr lfn://grid/dteam/user/lloyd/data_100mb user.replicas  
srm://dc2-grid-23.brunel.ac.uk/dpm/brunel.ac.uk/home/dteam/generated/2014-08-29/file49b1ac6d-  
e6de-4d1e-89b8-66defede7734
```

- GGUS: https://ggus.eu/index.php?mode=ticket_info&ticket_id=108121
- “Solution: Solved in gfal2.7”
- Gfal2.7 in epel-testing from around 2014-11-10

- Currently use lcg-gt to get the turls
- Can use gfal-xattr to get these:

```
$ gfal-xattr --key user.replicas srm://gfe02.grid.hep.ph.ic.ac.uk/pnfs/hep.ph.ic.ac.uk/data/dteam/generated/2014-10-09/filedb9a5d7d-d3ad-4576-8d16-9dd0284a5cae
```

```
user.replicas = gsiftp://sedsk53.grid.hep.ph.ic.ac.uk:2811//pnfs/hep.ph.ic.ac.uk/data/dteam/generated/2014-10-09/filedb9a5d7d-d3ad-4576-8d16-9dd0284a5cae
```

```
user.status = ONLINE
```

```
$ gfal-xattr --key user.replicas -D "SRM_PLUGIN:TURL_PROTOCOLS=https" srm://v6se00.grid.hep.ph.ic.ac.uk/dpm/grid.hep.ph.ic.ac.uk/home/dteam/generated/2014-08-28/filef6103227-4947-420f-b004-dcf3be5bdeb0
```

```
user.replicas = https://v6se00.grid.hep.ph.ic.ac.uk//dpm/grid.hep.ph.ic.ac.uk/home/dteam/generated/2014-08-28/filef6103227-4947-420f-b004-dcf3be5bdeb0
```

```
user.status = ONLINE
```

- From 1.8.9, DPM advertises webdav which causes some issues one of which is it doesn't respond to a request for a webdav TURL:

```
[drand1@v6ui00 ~]$ gfal-xattr --key user.replicas -D "SRM_PLUGIN:TURL_PROTOCOLS=webdav" srm://v6se00.grid.hep.ph.ic.ac.uk/dpm/grid.hep.ph.ic.ac.uk/home/dteam/generated/2014-08-28/filef6103227-4947-420f-b004-dcf3be5bdeb0
```

```
gfal-xattr: error: Resource temporarily unavailable
```

- GGUS ticket: https://ggus.eu/index.php?mode=ticket_info&ticket_id=109168
- At time of writing IC test DPM not advertising https or webdav, don't know why 7

- Initial sites
 - Brunel: Dual-stack DPM and production dual-stack WN
 - Imp_v6: IC testbed - CE, SE, WN - all IPv6 only
 - QMUL: production IPv4-only WN and dual-stack Storm (se02 then se04)
 - Oxford: IPv6-only test DPM and production dual-stack WN
- Later when IC dCache upgraded to 2.10, thereby adding support for IPv6, we added:
 - Imp: IC dual-stack production services: CE, SE, WN

Protocol: lcg

	Storage Element					
CE	Brun	Imp	Imp_V6	QMUL	Ox	CE
Brun	0	15	15	15	15	Brun
Imp	0	23	23	23	23	Imp
Imp_V6	0	24	24	24	24	Imp_V6
QMUL	0	0	0	0	0	QMUL
Ox	0	2	2	2	2	Ox
	Brun	Imp	Imp_V6	QMUL	Ox	

- Current status: Running
Most recent job submitted on Tue Nov 11 2014 at 16:00
- Test jobs are submitted every 1 hours to all UK CEs which read data from all SEs to the WN on which the job is running. The transfers are timed to calculate the throughput. The averages in MB/s over the last 24 hours are shown below.
- SE to Local WN Transfers
- The files are copied to the local WN from all SEs.

Site	Latest Jobs										Time of Last Job	
	10	9	8	7	6	5	4	3	2	1	Submitted	Successful
UKI-LT2-Brunel	X	C	E	S	S	S	C	S	S	C	11/11/14 17:00	11/11/14 16:00
UKI-LT2-IC-HEP	S	S	S	S	S	S	S	S	S	C	11/11/14 17:00	11/11/14 16:00
UKI-LT2-IC-HEP-V6-TESTBED	S	S	S	S	S	S	S	S	S	C	11/11/14 17:00	11/11/14 16:00
UKI-LT2-QMUL	F	E	E	E	F	F	F	E	E	C	11/11/14 17:00	14/09/14 14:20
UKI-SOUTHGRID-OX-HEP	A	A	A	A	A	A	A	C	C	C	11/11/14 17:00	11/11/14 06:00

- Not unexpectedly: most SEs and CEs support lcg-cp
 - (Brunel SE undergoing maintenance at time of writing)
- Slightly confusing test, e.g. false positive: QMUL was working over IPv4 so implemented a job kill if not IPv6 connectivity to www.google.com not found
 - On reflection not sure this is the best idea as we want 'true' IPv4 tests to succeed, perhaps we should just not run the lcg-cp test

Protocol: lcg

Ordered by sites

	Storage Element						
CE	Brun	Imp	Imp_V6	QMUL	Ox	Ave	CE
Brun		57.3	30.2	55.0	30.6	43.3	Brun
Imp		90.0	72.7	64.5	42.7	67.4	Imp
Imp_V6		53.0	23.4	47.3	42.3	41.5	Imp_V6
QMUL							QMUL
Ox		69.6	50.9	70.3	96.3	71.8	Ox
Ave		68.0	43.3	56.0	41.6	52.3	Ave
	Brun	Imp	Imp_V6	QMUL	Ox	Ave	

- Default in TURL_PROTOCOLS section in /etc/gfal2.d/srm_plugin.conf is

TURL_PROTOCOLS=rpio;gsidcap;dcap;kdcap;gsiftp

i.e. for a DPM site that advertises it, to use rpio

Many DPM sites advertise rpio but the port is blocked from outside

- Can over-ride with -D “SRM_PLUGIN:TURL_PROTOCOLS=gsiftp”

- But might it be better to have a fall-back to the next protocol?
- Similarly in /etc/gfal2.d/gsiftp_plugin.conf IPV6=false by default
 - Over-ride with -D “GRIDFTP_PLUGIN:IPV6=true”

- “Gfal / lcg-util 1.16.0 have been already pushed into EPEL with the remaining functionality that was still pending. Since there are no open tickets against these components, it seems it is a good moment to stop the development. There will still be support for critical bug-fixes until the 31st October 2014. After that, there will be no support of any kind, and the packages will be retired from the Fedora / EPEL repositories. Help and support for migration into gfal2 / gfal-util will be kept indefinitely.”
- But at ‘Date of problem: 2014-09-03 11:49:00’ gfal2-util not on WN or UI by default
- GGUS ticket requesting these be installed:
https://ggus.eu/index.php?mode=ticket_info&ticket_id=108201
- Rapidly solved and EMI-WN/EMI-UI metapackages now include gfal2-util
- As at 11 Nov 2014 a number of the participating sites have not upgraded to the latest version of the WN so gfal2-util still not available

- CEs: Currently only Imperial V6 testbed and Brunel have gfal-copy client installed
- SEs: all respond
- Aside: snoplus having problems copying out of RAL CASTOR with gfal-copy
 - GGUS: https://ggus.eu/index.php?mode=ticket_info&ticket_id=109694

Protocol: gfal

Ordered by sites

	Storage Element						
CE	Brun	Imp	Imp_V6	QMUL	Ox	Ave	CE
Brun		42.7	37.9	59.3	29.3	42.5	Brun
Imp							Imp
Imp_V6		35.2	28.2	33.2	30.7	31.8	Imp_V6
QMUL							QMUL
Ox							Ox
Ave		38.0	31.8	42.8	30.2	35.7	Ave
	Brun	Imp	Imp_V6	QMUL	Ox	Ave	

- WMS creates proxies that curl is seemingly unable to use to authenticate with DPM
- DPM uses gridsite which uses 'caNL' - interesting debugging
- For more details see GGUS ticket:
https://ggus.eu/index.php?mode=ticket_info&ticket_id=108269
- dCache and Storm not affected

- Noticed relatively poor http performance at QMUL: only ~15-20MB/s
- Storm http implementation called Milton copies files to /tmp prior to making them available
- Have also seen complaints on dCache user list
- GGUS ticket: ‘/tmp/milton_partial files filling up our SE’
https://ggus.eu/index.php?mode=ticket_info&ticket_id=105361
- At time of writing Chris has updated webdav interface on se02:
<https://se02.esc.qmul.ac.uk:8443/webdav/>
...and partially se04 which is the SE we’re testing:
<https://se02.esc.qmul.ac.uk:8443/dteam/>

- CEs:
 - QMUL jobs would work but are killed off
 - Imperial_V6 doesn't work which is as expected!

- SEs:
 - Only IC dCache works
 - DPMs (Brun, Imp_V6, Ox) have WMS proxy issue
 - QMUL: At time of writing se04 had no webdav interface

Protocol: http_ipv4

Ordered by sites

	Storage Element						
CE	Brun	Imp	Imp_V6	QMUL	Ox	Ave	CE
Brun		96.6				96.6	Brun
Imp		250.4				250.4	Imp
Imp_V6							Imp_V6
QMUL							QMUL
Ox		204.7				204.7	Ox
Ave		180.2				180.2	Ave
	Brun	Imp	Imp_V6	QMUL	Ox	Ave	

- CEs:
 - Dual-stack WN (Brun, Imp, Ox) all work - good!
 - QMUL (IPv4 only) are killed off - as expected
 - Imperial_V6 works - good!
- SEs:
 - Only Imp dCache works
 - DPMs (Brun, Imp_V6, Ox) have WMS proxy issue
 - QMUL: At time of writing se04 had no webdav interface

Protocol: http_ipv6

Ordered by sites

	Storage Element						
CE	Brun	Imp	Imp_V6	QMUL	Ox	Ave	CE
Brun		85.7				85.7	Brun
Imp		258.6				258.6	Imp
Imp_V6		46.5				46.5	Imp_V6
QMUL							QMUL
Ox		76.8				76.8	Ox
Ave		133.6				133.6	Ave
	Brun	Imp	Imp_V6	QMUL	Ox	Ave	

- Xrootd is now a popular client used by ATLAS, CMS and ALICE
 - Using versions installed on cvmfs
- We are doing our own tests, we don't want to use an ATLAS or CMS version buried in cvmfs
- What about small VOs?
- Suggests we need a 'middleware' version of xrootd client.
 - Useful for others, e.g. cloud work
- GGUS ticket: https://ggus.eu/index.php?mode=ticket_info&ticket_id=108653
 - Response: "I'll try to add it in the next version, after checking that it is ok also from the dependencies p.o.v"
- Of course for IPv6 compatibility this needs to be xrootd

- CEs:
 - Only Imperial IPv6 WN has installed xrootd4 client but as expected it doesn't work over IPv4 - good!
- SEs:
 - Imp dCache should work but no clients testing it
 - Dual-stack DPM (Brun): should work but no clients testing it
 - IPv6-only DPMs (Imp_v6 and Ox): shouldn't work over IPv4
 - QMUL: Only ATLAS supported over xrootd

Protocol: xrootd_ipv4

Ordered by sites

	Storage Element						
CE	Brun	Imp	Imp_V6	QMUL	Ox	Ave	CE
Brun							Brun
Imp							Imp
Imp_V6							Imp_V6
QMUL							QMUL
Ox							Ox
Ave							Ave
	Brun	Imp	Imp_V6	QMUL	Ox	Ave	

- CEs:
 - Only Imperial IPv6 has installed (xrootd4) client. It works - good!

- SEs:
 - Imperial dCache indeed works
 - DPMs (Brun, Imp_V6, Ox): don't support IPv6 yet: dpm-xrootd available but not yet in epel/epel-testing
 - QMUL: Only ATLAS supported over xrootd

Protocol: xrootd_ipv6

Ordered by sites

	Storage Element						
CE	Brun	Imp	Imp_V6	QMUL	Ox	Ave	CE
Brun							Brun
Imp							Imp
Imp_V6		133.3				133.3	Imp_V6
QMUL							QMUL
Ox							Ox
Ave		133.3				133.3	Ave
	Brun	Imp	Imp_V6	QMUL	Ox	Ave	

- LFC and WMS being deprecated, need to test Dirac as alternative
- Plan to either migrate or clone tests to use GridPP Dirac instance
- Obviously needs an IPv6-enabled Dirac
- Brief attempt but GridPP dirac instance doesn't support dteam and IPv6 test VOs don't support gridpp VO
- Have managed to submit of a job to our IPv6 only CE and WN via test Dirac using the dteam VO
- Job used curl to download a file from our IPv6-only DPM SE
- Work in progress...

- Imperial production services are all dual-stack so less need for our IC IPv6 testbed
- Would like to decommission IPv6 testbed if possible
- But it would be still be nice to be able to simulate an IPv6-only WN
- <https://github.com/ic-hep/ipv6only/tree/master>
- Library emulates an IPv6 only worker node (in a fairly poor way)
- We can add it to LD_PRELOAD for all grid jobs and then any tests simply set the IPV6_ONLY environment variable to disable IPv4
- Early days, needs more testing...

- We're trying to test the many storage access protocols over IPv4 and IPv6 in a systematic way
- Currently only using a small set of IPv6 related CEs, WNs and SEs
- We could also broaden it out to include all GridPP sites
- Many issues discovered on the way, see GGUS tickets
- Future work include porting to the GridPP Dirac once IPv6 is enabled and simulating IPv6-only worker nodes on our production cluster