

dCache at Tier3

Joe Urbanski
University of Chicago

US ATLAS Tier3/Tier2 Meeting, Bloomington
June 20, 2007

What is dCache?

From dCache.org:

"A system for storing and retrieving huge amounts of data, distributed among a large number of heterogenous server nodes, under a single virtual filesystem tree"

Features & Advantages

- What can dCache do?
 - Capable of combining hundreds of commodity disk servers to get a huge petabyte scale data store
 - Allows several copies of a single file for distributed data access
 - Has internal load balancing using cost metrics and transfers between the site's pools
 - Has automatic file replication on high load ("hotspot detection")

What does dCache look like?

- The single virtual filesystem is provided by pnfs (Pretty Normal File System).
- Mounting and unmounting through *mount* and *umount* and */etc/fstab*, much like traditional *nfs*.
- *pnfs* is “POSIX-like”:
 - can use: *ls*, *mkdir*, *find*
 - cannot use: *cp*, *md5sum*

Clients

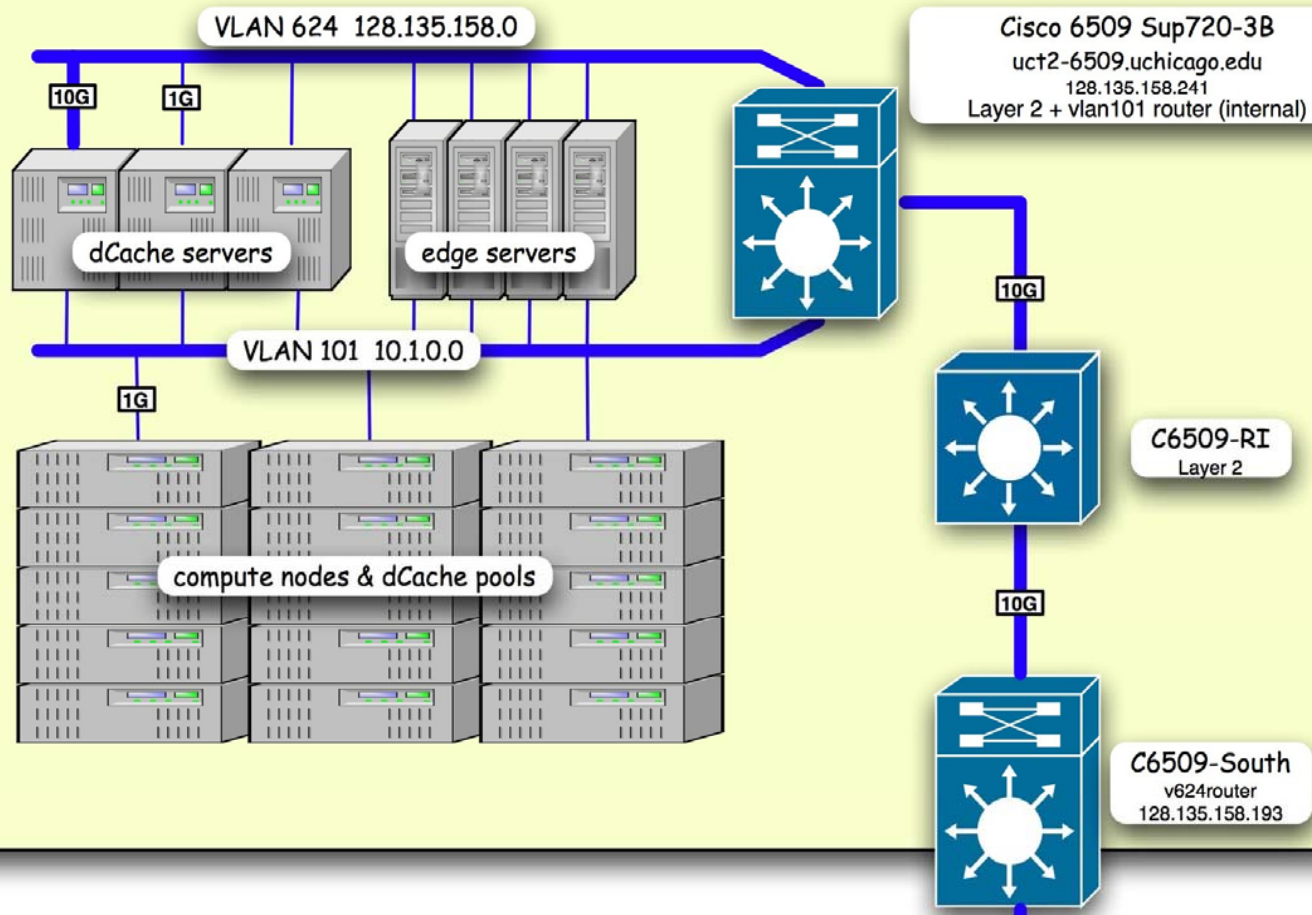
- **How do I access files in dCache?**
 - **dCap**: dCache's native method, uses *dccp*. Easiest for local transfers.
 - **gridFTP**: via *globus-url-copy*
 - **SRM**: via *srmcp*

Architecture

- *What about the backend?*
 - **Admin nodes:** Provide basic admin services. One or more of these.
 - **Door nodes:** Provide I/O access via SRM or GridFTP. One or more, may reside on admin node on small installs.
 - **pnfs node:** Provides the unified namespace that masks complexity. Only one per install.
 - **Pool nodes:** Provides storage. Can be installed on other nodes.

UC Tier3

ATLAS MWT2 Network: UC Campus



UC Tier3 (cont'd)

- 3 Admin nodes:
 - uct3-edge1: gridFTP, dCap
 - uct3-edge2: pnfs
 - uct3-edge3: admin, SRM
- 25 Pool nodes:
 - Currently: 22 compute nodes x 1.9TB + 3 admin nodes x 1.9TB = **47TB**

Installing dCache

- The VDT Installer
 - Greatly simplifies and quickens the install process
 - Automatically configures which services to run, and on what nodes to run them.
 - Installs needed rpms.
 - Configures dCache and its postgresql databases.
 - Latest vdt version is v1.1.8
 - Not to be confused with the dCache version, latest is v1.7.0.

Running the VDT Installer

- Download the latest vdt tarball, untar and cd into the install directory.
- Run node-config.pl to generate site-info.def file:
 - <http://twiki.mwt2.org/bin/view/Admins/DCacheUpgradeTov114>

config-node.pl

```
[root@uct3-edge2 install] ./config-node.pl
```

How many admin nodes (non-pool and non-door nodes) do you have? 2

The recommended services for node 1 are:

ImDomain poolManager adminDoor httpDomain utilityDomain gplazmaService infoProvider
srm

The recommended services for node 2 are:

pnfsManager dirDomain

Enter the FQDN for the node 1: uct3-edge3.uchicago.edu

Which services do you wish to run on node uct2-dc3.uchicago.edu (Enter for defaults)?

Enter the FQDN for the node 2: uct3-edge2.uchicago.edu

Which services do you wish to run on node uct2-dc2.uchicago.edu (Enter for defaults)?

How many door nodes do you have? 1

Enter the FQDN of door number 1: uct3-edge1.uchicago.edu

node-config.pl (cont'd)

Enter the private network that the pools are in.

If this does not apply, just press enter to skip: 10.2.0.0

Enter the number of dcap doors to run on each door node [default 1]: 1

Enter a pool FQDN name(Press Enter when all are done): uct3-c001.mwt2.org

Enter the first storage location (Press Enter when all are done)): /dcache

Enter another storage location (Press Enter when all are done)):

--SNIP--

Enter another pool FQDN name(Press Enter when all are done): uct3-edge3.mwt2.org

Enter the first storage location (Press Enter when all are done)): /dcache

Enter another storage location (Press Enter when all are done)):

Enter another pool FQDN name(Press Enter when all are done):

Created site-info.def file.

[root@uct2-mgt install]

Running the VDT Installer (cont'd)

- Copy vdt tarball, and site-info.def file to all nodes.
- Run './install.sh -d' for a dryrun. (This will be very verbose). If successful, run the actual install.
- Start dCache services in the following order:
 - pnfs node core services
 - other admin nodes core services
 - all dcache pool services

Verifying the install

- Check the status webpage:
 - <http://uct3-edge3.uchicago.edu:2288/cellInfo>

Services

CellName	DomainName	RP	TH	Ping	Creation Time	Version
DCap-gsi-uct3-edge1	gsidcap-uct3-edge1Domain	0	4	55 msec	06/15 10:22:50	production-1-7-0(1.13.6.2)
DCap-uct3-edge1	dcap-uct3-edge1Domain	0	4	55 msec	06/15 10:22:36	production-1-7-0(1.13.6.2)
GFTP-uct3-edge1	gridftp-uct3-edge1Domain	0	4	55 msec	06/15 10:22:43	production-1-7-0(1.16)
LoginBroker	httpdDomain	0	2	1 msec	06/15 10:16:14	cells(1.3.2.2)
PnfsManager	pnfsDomain	0	5	52 msec	06/15 10:43:38	production-1-7-0(1.24.2.4)
PoolManager	dCacheDomain	0	4	5 msec	06/15 10:16:00	production-1-7-0(1.32)
srm>LoginBroker	httpdDomain	0	2	1 msec	06/15 10:16:14	cells(1.3.2.2)
uct3-c001_1	uct3-c001Domain	0	9	52 msec	06/14 11:32:31	production-1-7-0(1.130.2.4)
uct3-c002_1	uct3-c002Domain	0	9	53 msec	06/14 11:32:48	production-1-7-0(1.130.2.4)
uct3-c003_1	uct3-c003Domain	0	9	51 msec	06/14 11:58:46	production-1-7-0(1.130.2.4)
uct3-c004_1	uct3-c004Domain	0	9	50 msec	06/14 11:50:52	production-1-7-0(1.130.2.4)
uct3-c005_1	uct3-c005Domain	0	9	55 msec	06/14 11:33:40	production-1-7-0(1.130.2.4)
uct3-c006_1	uct3-c006Domain	0	9	53 msec	06/14 11:33:57	production-1-7-0(1.130.2.4)
uct3-c007_1	uct3-c007Domain	0	9	54 msec	06/14 11:34:14	production-1-7-0(1.130.2.4)
uct3-c008_1	uct3-c008Domain	0	9	53 msec	06/14 12:02:13	production-1-7-0(1.130.2.4)
uct3-c009_1	uct3-c009Domain	0	9	53 msec	06/14 11:34:49	production-1-7-0(1.130.2.4)
uct3-c010_1	uct3-c010Domain	0	9	55 msec	06/14 11:35:05	production-1-7-0(1.130.2.4)
uct3-c011_1	uct3-c011Domain	0	9	51 msec	06/14 11:35:22	production-1-7-0(1.130.2.4)
uct3-c012_1	uct3-c012Domain	0	9	48 msec	06/14 11:35:39	production-1-7-0(1.130.2.4)
uct3-c013_1	uct3-c013Domain	0	9	54 msec	06/14 11:35:56	production-1-7-0(1.130.2.4)
uct3-c014_1	uct3-c014Domain	0	9	45 msec	06/14 11:36:12	production-1-7-0(1.130.2.4)

Verifying the install (cont'd)

- Test the doors:
 - dCap: use *dccp*

```
dccp -d999 /tmp/test.file \ /pnfs/uchicago.edu/data/ddm1/test.file
```

- gridFTP: use *globus-url-copy*

```
globus-url-copy -dbg \  
file:///tmp/test.file \  
gsiftp://uct3-edge1.uchicago.edu/pnfs/uchicago.edu/data/ddm1/test.file
```

- SRM: use *srmcp*

```
srmcp -debug file:///tmp/test.file \  
srm://uct3-edge3.uchicago.edu:8443/pnfs/uchicago.edu/data/ddm1/test.file
```

Troubleshooting

- Always check the status page!

CellName	DomainName	RP	TH	Ping	Creation Time	Version
DCap-gsi-uct2-dc1	gsidcap-uct2-dc1Domain	0	4	48 msec	06/15 16:44:38	production-1-7-0(1.13.6.2)
DCap-uct2-dc1	dcap-uct2-dc1Domain	0	4	52 msec	06/15 16:44:24	production-1-7-0(1.13.6.2)
GFTP-uct2-dc1	gridftp-uct2-dc1Domain	0	9	51 msec	06/15 16:44:31	production-1-7-0(1.16)
LoginBroker	httpdDomain	0	2	2 msec	06/15 16:43:38	cells(1.3.2.2)
PnfsManager	pnfsDomain	0	5	50 msec	06/15 16:42:53	production-1-7-0(1.24.2.5)
PoolManager	dCacheDomain	0	4	8 msec	06/15 16:43:23	production-1-7-0(1.32)
SRM-uct2-dc3	srm-uct2-dc3Domain	0	3	57 msec	06/18 11:14:18	production-1-7-0(1.88.2.11)
srm-LoginBroker	httpdDomain	0	2	2 msec	06/15 16:43:38	cells(1.3.2.2)
uct2-c001_1	uct2-c001Domain	0	10	50 msec	06/18 10:51:17	production-1-7-0(1.130.2.3)
uct2-c002_1	uct2-c002Domain	0	10	51 msec	06/18 10:51:27	production-1-7-0(1.130.2.3)
uct2-c003_1	uct2-c003Domain	0	10	49 msec	06/18 10:51:35	production-1-7-0(1.130.2.3)
uct2-c004_1	uct2-c004Domain	0	10	54 msec	06/18 10:51:44	production-1-7-0(1.130.2.3)
uct2-c005_1	uct2-c005Domain	0	10	53 msec	06/18 10:51:52	production-1-7-0(1.130.2.3)
uct2-c006_1	uct2-c006Domain	0	10	56 msec	06/18 10:52:00	production-1-7-0(1.130.2.3)
uct2-c007_1	uct2-c007Domain	0	10	55 msec	06/18 10:52:09	production-1-7-0(1.130.2.3)
uct2-c008_1	uct2-c008Domain	OFFLINE				
uct2-c009_1	uct2-c009Domain	0	10	49 msec	06/18 10:52:24	production-1-7-0(1.130.2.3)
uct2-c011_1	uct2-c011Domain	0	10	55 msec	06/18 10:52:38	production-1-7-0(1.130.2.3)
uct2-c012_1	uct2-c012Domain	0	10	52 msec	06/18 10:52:46	production-1-7-0(1.130.2.3)
uct2-c013_1	uct2-c013Domain	0	10	55 msec	06/18 10:52:54	production-1-7-0(1.130.2.3)
uct2-c014_1	uct2-c014Domain	0	10	52 msec	06/18 10:53:03	production-1-7-0(1.130.2.3)
uct2-c015_1	uct2-c015Domain	0	9	49 msec	06/18 10:51:47	production-1-7-0(1.130.2.3)
uct2-dc1_1	uct2-dc1Domain	0	10	52 msec	06/18 10:49:49	production-1-7-0(1.130.2.3)
uct2-dc2_1	uct2-dc2Domain	0	10	47 msec	06/15 16:44:08	production-1-7-0(1.130.2.3)

Troubleshooting (cont'd)

- Check the logs:
 - Most dCache cells: /var/log/*Domain.log. Each cell, or service, will generate a log with the appropriate name
 - SRM: /opt/d-cache/libexec/apache-tomcat-5.5.20/logs/catalina.out
 - pnfs: /var/log/pnfsd.log
- dCache, The Book:
 - <http://www.dcache.org/manuals/Book/>
- Ask for help:
 - dCache.org support: support@dcache.org
 - OSG-storage: osg-storage@opensciencegrid.org

UC Tier3 Pool Usage

Disk Space Usage

CellName	DomainName	Total Space/MB	Free Space/MB	Precious Space/MB	Layout (precious/free)
uct3-c001_1	uct3-c001Domain	1906688	538991	1361244	
uct3-c002_1	uct3-c002Domain	1906688	541105	1358074	
uct3-c003_1	uct3-c003Domain	1906688	541254	1357982	
uct3-c004_1	uct3-c004Domain	1906688	539367	1360066	
uct3-c005_1	uct3-c005Domain	1906688	538994	1362731	
uct3-c006_1	uct3-c006Domain	1906688	542617	1359764	
uct3-c007_1	uct3-c007Domain	1906688	538596	1362295	
uct3-c008_1	uct3-c008Domain	1906688	541149	1358144	
uct3-c009_1	uct3-c009Domain	1906688	538958	1361369	
uct3-c010_1	uct3-c010Domain	1906688	539266	1359654	
uct3-c011_1	uct3-c011Domain	1906688	538992	1361517	
uct3-c012_1	uct3-c012Domain	1906688	539501	1361119	
uct3-c013_1	uct3-c013Domain	1906688	540059	1360811	
uct3-c014_1	uct3-c014Domain	1906688	539466	1361626	
uct3-c015_1	uct3-c015Domain	1906688	538979	1360937	
uct3-c016_1	uct3-c016Domain	1906688	538993	1361335	
uct3-c017_1	uct3-c017Domain	1906688	539326	1361125	
uct3-c018_1	uct3-c018Domain	1906688	544007	1362680	
uct3-c019_1	uct3-c019Domain	1906688	538985	1360268	
uct3-c020_1	uct3-c020Domain	1906688	537648	1362958	
uct3-c021_1	uct3-c021Domain	1906688	538602	1360515	

Useful URL's

- dCache homepage
 - <http://www.dcache.org/>
- VDT Installer homepage
 - <http://vdt.cs.wisc.edu/software/dcache/dcache/>
- VDT Installer HOWTO
 - <http://vdt.cs.wisc.edu/extras/InstallingDcacheForOSG.html>
- OSG dCache overview
 - <http://vdt.cs.wisc.edu/software/dcache/dcache/>
- USATLAS dCache at BNL
 - <http://vdt.cs.wisc.edu/extras/InstallingDcacheForOSG.html>