

Experience with dCache at MWT2

Charles G Waldman
USATLAS/University of Chicago
cgw@hep.uchicago.edu

OSG Storage Forum, June 30, Fermi National Accelerator Laboratory

Overall summary

- First install: dCache 1.6.6, Oct. 2006: 30 TB in 17 pools (x2), 28 job slots
- Current: dCache 1.9.2-5, 295TB in 54 pools, 1328 job slots (with faster CPUs)
- Storage increased by ~10x, CPU by ~50x
- Increased complexity of workflow & network
- With scaling, discovery of new failure modes
- Lots of time spent debugging, applying “band-aids”, etc.
- Situation is improving, esp. with 1.9 release

Support model

- Configuring dCache is somewhat complex, hence VDT package largely preconfigured
- Multiple avenues for support:
 - osg-storage@opensciencegrid.org
 - support@dcache.org
 - user-forum@dcache.org (often extremely helpful)
- Support requires good understanding of config.
- Cross-posting: good or bad?

Support model – 2

- OSG-Storage weekly meeting: should site representatives attend?
- Documentation issues:
 - Overlap between dCache.org and OSG docs
 - User input needs to get incorporated into documentation

Case studies - 1

- “hot file” problem: solved with pcache
- dccp is most reliable client, but space token support only came recently.
- “no write pools available” error caused by network packet loss between pools and pool manager.
- Error rate decreased further with 1.9.2-5
- Excessive memory consumption due to `vm.overcommit_memory=2` and Java 1.5
 - Solved by installing Java 1.6

Case studies - 2

- Tuning of Postgres database backend
- Hanging transfers due to non-uniform access patterns (copy vs. random-access dcap I/O)
 - Solved via multiple mover queues, implemented locally. Required understanding of workflow
- Public vs private pools
 - Gridftp2 available in recent dCache versions
- SRM scalability & performance
 - ITB tests functionality only, no load or stress testing. (Sites could help with this)

Case studies - 3

- gPlazma authorization timeout failures
 - Currently using gridmap/gridvorolemap files
 - Continuing to debug GUMS callout with help of OSG-Storage

How can support be made easier?

- Simpler, annotated, cleaned-up dCacheSetup?
- Most frequent complaint, “single most important improvement” - format and contents of log files
- Many support issues are just a matter of understanding/interpreting logs
- Log files often make it difficult to track down real error

Logging - 2

- Gridftp door logs: filenames no longer logged
- Line format hard to read (error message starts in column >40)
- catalina.out:
 - Why no proper SRM logfile?
 - Misleading messages (this is a successful transfer)

```
2009-06-16 04:03:19.304 (SRM-uct2-dc1) []  
dmg.cells.nucleus.CellNucleus.log(CellNucleus.java:847) ERROR -  
PutRequestHandler error: copy request state changed to Done
```

Logging - 3:

- Simple “directory exists” results in:

```
2009-06-16 04:03:18.710 (SRM-uct2-dc1) [] dmg.cells.nucleus.CellNucleus.log(CellNucleus.java:873) ERROR - org.dcache.srm.SRMDuplicationException: already exists
org.dcache.srm.SRMDuplicationException: already exists
  at diskCacheV111.srm.dcache.Storage.createDirectory(Storage.java:3126)
  at org.dcache.srm.handler.SrmMkdir.srmRmdir(SrmMkdir.java:122)
  at org.dcache.srm.handler.SrmMkdir.getResponse(SrmMkdir.java:75)
  at sun.reflect.GeneratedMethodAccessor155.invoke(Unknown Source)
  at sun.reflect.DelegatingMethodAccessorImpl.invoke(DelegatingMethodAccessorImpl.java:25)
  at java.lang.reflect.Method.invoke(Method.java:597)
  at org.dcache.srm.server.SRMServerV2.handleRequest(SRMServerV2.java:290)
  at org.dcache.srm.server.SRMServerV2.srmMkdir(SRMServerV2.java:400)
  at sun.reflect.GeneratedMethodAccessor153.invoke(Unknown Source)
  at sun.reflect.DelegatingMethodAccessorImpl.invoke(DelegatingMethodAccessorImpl.java:25)
  at java.lang.reflect.Method.invoke(Method.java:597)
  at org.apache.axis.providers.java.RPCProvider.invokeMethod(RPCProvider.java:397)
  at org.apache.axis.providers.java.RPCProvider.processMessage(RPCProvider.java:186)
  at org.apache.axis.providers.java.JavaProvider.invoke(JavaProvider.java:323)
  at org.apache.axis.strategies.InvocationStrategy.visit(InvocationStrategy.java:32)
  at org.apache.axis.SimpleChain.doVisiting(SimpleChain.java:118)
  at org.apache.axis.SimpleChain.invoke(SimpleChain.java:83)
  at org.apache.axis.handlers.soap.SOAPService.invoke(SOAPService.java:453)
  at org.apache.axis.server.AxisServer.invoke(AxisServer.java:281)
  at org.apache.axis.transport.http.AxisServlet.doPost(AxisServlet.java:699)
  at javax.servlet.http.HttpServlet.service(HttpServlet.java:709)
  at org.apache.axis.transport.http.AxisServletBase.service(AxisServletBase.java:327)
  at javax.servlet.http.HttpServlet.service(HttpServlet.java:802)
  at org.apache.catalina.core.ApplicationFilterChain.internalDoFilter(ApplicationFilterChain.java:252)
  at org.apache.catalina.core.ApplicationFilterChain.doFilter(ApplicationFilterChain.java:173)
  at org.apache.catalina.core.StandardWrapperValve.invoke(StandardWrapperValve.java:213)
  at org.apache.catalina.core.StandardContextValve.invoke(StandardContextValve.java:178)
  at org.apache.catalina.core.StandardHostValve.invoke(StandardHostValve.java:126)
  at org.apache.catalina.valves.ErrorReportValve.invoke(ErrorReportValve.java:105)
  at org.apache.catalina.core.StandardEngineValve.invoke(StandardEngineValve.java:107)
  at org.apache.tomcat.coyote.valves.HTTPSValve55.invoke(HTTPSValve55.java:45)
  at org.apache.catalina.connector.CoyoteAdapter.service(CoyoteAdapter.java:148)
  at org.apache.coyote.http11.Http11Processor.process(Http11Processor.java:869)
  at org.apache.coyote.http11.Http11BaseProtocol$Http11ConnectionHandler.processConnection(Http11BaseProtocol.java:664)
  at org.apache.tomcat.util.net.PoolTcpEndpoint.processSocket(PoolTcpEndpoint.java:527)
  at org.apache.tomcat.util.net.LeaderFollowerWorkerThread.runIt(LeaderFollowerWorkerThread.java:80)
  at org.apache.tomcat.util.threads.ThreadPool$ControlRunnable.run(ThreadPool.java:684)
  at java.lang.Thread.run(Thread.java:619)
```