



# Bits and Pieces

Christian Bernardt (DESY)

EMI rep dCache, dCache Developer

# New features in dCache 1.9.12 (emi-1)

- GLUE 2.0
- JMS
- P2p transfers
- SRM
- NFS, Non-ROOT, New Pin Manager



EMI INFO-RI-261611

3/17/11

Bits and pieces - Christian Bernardt - dCache (DESY HH)



# GLUE 2.0 (EMI Global effort)

- Motivation – GLUE 2.0 standard
  - No duplication of entries (SE, Service, SRM)
  - More flexibility in publishing data (future proof) → e.g. which endpoints are allowed to use what kind of resource
  - Hierarchical structures → Inheritance
  - Validation service for published data (external script, querying all grid info)
- Migration:
  - both are published simultaneously in different LDAP root



# GLUE 2.0 - Configuration

- Site-info.def concept still there
  - /opt/d-cache/share/defaults/info-provider.properties  
info-provider.publish=[1.3 | 2.0 | both]
- glue-1.3.xml
  - Sufficiently descriptive to alter directly in the file
  - Problem: template file gets updated at instance update, but not the glue-1.3.xml
  - Solution: separating the untouched xml in the file from the site specific information
    - site specific file + template file = published data
      - /opt/d-cache/etc/infoprovider.xml (site specific)
      - /opt/d-cache/etc/info-provider/glue... (static data)
    - Please let us know if YOU need change in the static data



# JMS



EMI INFO-RI-261611

3/17/11

Bits and pieces - Christian Bernardt - dCache (DESY HH)

5

# JMS – Why!

- Beta status → DO NOT use in production
- Why should you switch to JMS?
  - Better setup of messaging (SSL)
  - Auto discovery (no central message broker)
  - Redundant communication infrastructure (no central point of failure - dCacheDomain)
  - NDGF – coping better with DNS mapping changes (connection to old IPs)
  - Future benefits:
    - Redundant service deployment





# JMS – How!

- How to switch?
  - broker.scheme=hybrid (in dcache.conf for all nodes)
  - Mixed installation possible:
    - dCache domain need to have hybrid defined (serves as hub, gateway for old and new transport)



# JMS – Scale out

- How to scale out? – Deploy redundant setup with multiple brokers
  - as soon as message broker is outside of dCacheDomain
    - Then just configure:  
broker.scheme=jms  
broker.amq.url=<ActiveMQ connection URI>





# Log System

- Easier
  - More flexible
- Mattias Wadenstein dedicated talk



# p2p Transfer



EMI INFO-RI-261611

3/17/11

Bits and pieces - Christian Bernardt - dCache (DESY HH)

10

# p2p Transfer

- Cost cut calculation (Jon's algorithm)

- Situation before the change

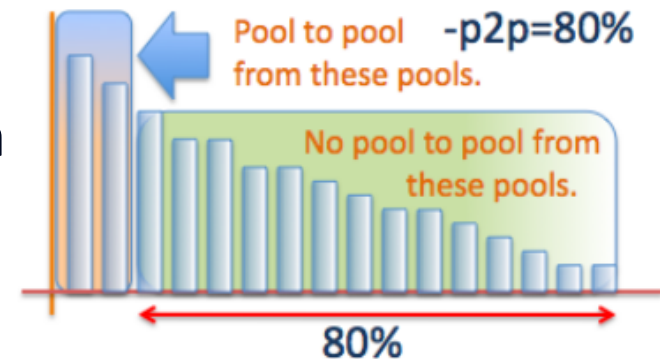
- Costcut depends queue length

- Situation after the change

- Dependency now resolved: Look at x busiest pools → only do a transfer when request lands on them

- Disadvantages:

- If you have a lightly loaded machine, you always have p2p transfers → only works good on heavily loaded machines (this needs fixing)



# SRM



EMI INFO-RI-261611

3/17/11

Bits and pieces - Christian Bernardt - dCache (DESY HH)

12

# SRM

- Jetty instead of Tomcat – Why?
  - Jetty can be a component of dCache (Tomcat can only run a cell)
  - Better logging (NO CATALINA.OUT!!!)
  - Better memory footprint
    - better scalability for connection with long idle time
    - Less CPU cache for servlet container, more for execution



# SRM – SSL (EMI Global Effort)

- SSL instead of GSI
  - Why are we doing this?
    - No *dependency* on globus
    - GSI not well specified only one implementation
    - SSL well tested lib code → move towards standards
    - Dedicated hardware acceleration
    - Client support: one switch
  - How to configure?
    - Endpoint config to GSI or SSL
    - manual switch: /opt/d-cache/share/srm.properties,  
srmGssMode=GSI|SSL
  - What is missing?
    - Prototype stage, but available
    - Not currently possible to run both on same port



# SRM

- Synchronous getTURL
  - Advantages
    - No reconnect
    - Asynchronous: polling until succeeded
    - Synchronous: one connection open per query
  - How to configure?
    - Default switch from synchronous to asynchronous after 1 second  
( /opt/d-cache/share/defaults/srm.properties,  
...SwitchToAsynchronousModeDelay=?)





# SRM - Spacemanager

- Multiple Link Groups
  - Granularity for space management in link group level
  - SpaceManager need to communicate to PoolManager (Links for writing) → no write link problem



# Other dCache news

NFS, Non-ROOT, Non-ROOT, New  
Pin Manager, ...



# Other dCache News

- No NFS mount point needed anymore
  - Srm needed to interact with namespace ( srmls)
- How to run as non-root
  - Domain that hosts ChimeraNFS server needs to run as root (run the server on a separate node)
- Single port per protocol per pool
  - For xrootd, nfs, webDAV
  - Firewall friendly
- Progress bar for dccp (-H option)



# Other dCache News

- Important fixes since 1.9.5
  - Pin Manager
    - Faster
    - More robust
    - Important bug fixes (stuck file transfer)
  - Pools – Performance and meta data
    - Removing stale meta data on startup
    - Message processing done by pool of threads
    - Single lock on metadata replaced by fine grained lock
      - Read only mode of pool during meta data recovery



# Conclusions

- dCache is moving towards:
  - More future proof technologies (SRM Jetty, GLUE 2.0, JMS)
  - Increased flexibility (Messaging, JMS)
  - Increased security (SRM SSL, one port per protocol per pool)
  - Easier maintainability, debugging (new logging system)





**Thank you**

**EMI is partially funded by the European Commission under Grant Agreement INFSO-RI-261611**