dCache Storage Resource Manager Configuration and Tuning

Timur Perelmutov

For the dCache Team

People

Head of dCache.org project

Patrick Fuhrman

Core Team (DESY and Fermilab)

Jon Bakken
Forrest Christian
Ted Hesselroth
Robert Kennedy
Alex Kulyavtsev
Birgit Lewendel
Dmitry Litvintsev
Dirk Pleiter
David Melkumyan
Vladimir Podstavkov
Martin Radicke
Owen Synge
Elena Tews

Leader of Fermilab Development

Timur Perelmutov

Leader of DESY Development

Tigran Mrktchyan

External Development

Gerd Behrmann, NGDF
Abhishek Sigh Rana, SDSC
BNL in preparation

Support and Help

Greig Cowan, gridPP (monAmi)

Stijn De Weirdt (Quatfor)

Maarten Lithmaath, CERN

Flavia Donno, CERN

N. N. CERN

SRM V2.2 Progress

- Worked on V2.2 implementation since May 06

- WLCG demanded functionality is ready

  - Functionality Tested

  - Compatibility Tested

  - Stress Tests are in progress
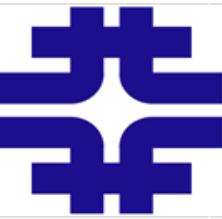
- Will be available in dCache 1.8

SRM V2.2 concepts

- **Storage Service properties**
  - AccessLatency := NEARLINE|ONLINE
  - RetentionPolicy := REPLICA|OUTPUT|CUSTODIAL
  - SpaceReservation is a vehicle for the above attributes
- **The mapping to labels 'TapeXDiskY' is given by:**
  - Tape1Disk0: NEARLINE + CUSTODIAL
  - Tape1Disk1: ONLINE + CUSTODIAL
  - Tape0Disk1: ONLINE + REPLICA

SRM V2.2

- Most functionality will work out of the box
- Will run on the same port as srm v1.1
- dCache SrmCopy will be compatible with both V1.1 and V2.2 servers
- dCache new component "SrmSpaceManager" performs space reservations on behalf of SRM
- All write requests are rerouted though SrmSpaceManager to guarantee consistency.
- SrmSpaceManager relies of LinkGroup functionality of PoolManager.

LinkGroups

- LinkGroup is a New Pool Manager abstraction
- LinkGroup is a collection of Links
- LinkGroups properties
  - Available Size
  - HSM Type
  - VO attributes (optionally)
- Space Reservations are made against LinkGroup sizes

LinkGroup Creation

To create a linkGroup that contains link "LFSOnly-link" which has no hsm backend (for storing Replica and Output files) that belongs to the "alice" VO and allows roles "/alice/production" and "/alice/test", add the following to the PoolManager.conf
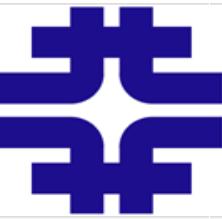
create linkGroup

psu create linkGroup ReplicaLinkGroup

add link to the linkGroup

psu addto linkGroup ReplicaLinkGroup LFSOnly-link

Specify the hsm type, "None" for Replica or Output, anything other

than "None" (say "enstore" or "OSM)is interpreted as Custodial

psu set linkGroup attribute –r ReplicaLinkGroup HSM=None

#optionally restrict access to VO and VO Roles (syntax might change)

psu set linkGroup attribute ReplicaLinkGroup VO=alice

psu set linkGroup attribute ReplicaLinkGroup aliceRole=/alice/production

psu set linkGroup attribute ReplicaLinkGroup aliceRole=/alice/test

Tuning

- After latest optimizations GSI authentication is a limiting factor
- Use separate powerful node
- Configure postgres DB well
  - Dedicated disk
  - Set kernel parameters
  - Enable autovacuuming
- Configure FTS to poll less frequently
- Detailed recommendations (dCache 1.7) at
  https://srm.fnal.gov/twiki/bin/view/SrmProject/DcacheSrmInstallation

Resources

- dCache http://www.dcache.org
- dCache SRM http://srm.fnal.gov
- Fermilab dCache project http://dcache.fnal.gov
- SRM WG http://sdm.lbl.gov/srm-wg/
- SRM V2.2 spec http://sdm.lbl.gov/srm-wg/doc/SRM.v2.2.html
- WLCG SRM V2.2 MoU https://srm.fnal.gov/twiki/pub/WorkshopsAndConferences/GridStorageInterfacesWSAge
- CERN SRM Tests and Issues (Flavia Donno) https://twiki.cern.ch/twiki/bin/view/SRMDev/WebHome
- LBNL SRM V2.2 tester http://sdm.lbl.gov/srm-tester/v22daily.html