

Data Management System

SRM/dCache

A Talk by
Owen Synge, DESY

Structure Of This Talk

- Introduction
 - What is a storage Cluster?
 - dCache storage cluster components.
 - SRM and its progress
 - Near future for dCache
 - Summary.
-

Introduction to Storage

- Requirements
 - Persistence of data (Cant just reinstall).
 - High performance / Bandwidth.
 - Security.
 - Scales as requirements grow.
 - Proportional to computing cluster.
 - Conclusion
 - Must be a cluster
 - Will be complex to manage
-

Introduction to a dCache

- What is a dCache Storage Cluster?
 - Mass storage system cache management
 - Provides a site based storage cluster
 - Manages cluster
 - Written primarily HEP community
 - Low security requirements
 - Performance a priority
 - Long production history
 - Expanding to suit disk only services
 - Pool hopping, data checksumming.
-

Not Just for DESY/HEP any more

- Mass auto deployed
 - RPM, APT/YUM, YAUM and Quattor
 - Authentication
 - rsh, kerberos, GSI, VOMS, SAML
 - By default uses gPlamsa to provide this (SL4)
 - Large multi-national user community
 - Multi Grid
 - wLCG, OSG, NorduGrid, D-Grid, GridPP and more
-

What is a dCache Storage Cluster?

- Many storage servers
 - One name space
 - One administrative interface
 - Many sub service's
 - Each door is a service
 - Information system
 - Tape cache
 - Cluster management
-

dCache Cluster Anatomy

- Administration Node
 - Administration and central services
 - Name space Node
 - Provides a single directory structured view.
 - Pool Nodes
 - Store the data.
 - Door Nodes
 - Allow data access.
-

Administration Node

- Runs central services
 - Billing database
 - Legal requirement for global deployment
 - Pool management
 - assigning data to pool, pools in pool groups etc
 - Cluster management (LM domain)
 - Discovery service (BDII)
 - Administration Interface (ssh)
 - Monitoring/Status interface (http)
-

Name Space Node : PNFS

- dCache decouples name service and authorization from data storage.
 - The current implementation (PNFS)
 - Unix UID/GID based
 - Requires host UID/GID available
 - Not scalable to Grid world
 - Files can have only one GID
 - No way to support privileges within a VO
 - production/user distinction
-

Name Space Node: Chimera

- Chimera Name Space service
 - Uses Virtual UID/GID model
 - Mapped 1:1 from unique DN/FQN or UID GIG
 - All resources have vUID/vGID
 - Action is extended (Not just RWX as in UNIX)
 - Built in hierarchy support.
 - All directories in a tree tested.
 - NFSv4 model
 - Will Support ACL's (POSIX Style)
-

Name Space Node: Chimera

- Database based
 - UID/GID will not be needed on host
 - Each resource will have single vUID
 - Each resource will have multiple vGID
 - But a primary vGID to enhance UNIX interoperability
 - Interoperability with NFSv4.1 clients has been shown for development versions.
 - for windows, sun and linux
-

Pool Nodes

- Store data
 - So need RAID storage
 - With Unique ID (Name Space gives name)
 - Checksum Data (New system)
 - Store meta data
 - So Name Space node can be regenerated
 - Only used in emergency data recovery
 - Move Data between Nodes
 - Movers can also transfer data to 3' storage.
-

Door Nodes

- GridFtp
 - Defacto WAN transfer protocol
 - Dcap
 - Available in GSI/rsh modes
 - POSIX style local file access
 - XrootD
 - POSIX style local file access
 - http
 - Currently supports only read access.
-

Door Nodes: SRM

- Load balancing protocol neutral interface
 - Extra management features.
 - SRM v1
 - Service Orientated Architecture Service
 - SRM v2
 - Space management
 - Directory management
 - Undergoing interoperability testing.
-

Near Future of dCache

- New 1.8.0 release coming very soon
 - Completed SRM 2.2 support
 - Fully validated by wLCG with test clients
 - Better logging
 - All new code uses Log4j
 - Supports Chimera or PNFS
 - Beyond 1.8.0 release
 - Improvements will remain driven by user communities.
-

Authentication, Authorization Near Future

- Support for SAZ
 - Acts as a client to Site AuthZ server
 - When
 - Release 1.8.0
 - Dynamic UID,GID output with Chimera
 - Currently DCache UID matches host UID
 - Virtual UID/GID allows for richer Authorization
 - Enables ACL's to be used.
 - When
 - Release 1.8.1/1.8.2
-

Summary

- Dcache is progressing in a use case driven way.
 - Authentication is still evolving.
 - ACL support is coming soon
 - Module for Chimera Name service
 - SRM 2.2 support
 - Validated against multiple clients
 - NFS v4.1
 - Providing a new grid interface soon after Chimera release.
-