



WLCG



Middleware Status

Markus Schulz, Oliver Keeble, Flavia Donno

Oxana Smirnova (ARC)

Alain Roy (OSG)

~ ~ ~

LCG-LHCC Mini Review

1st July 2008

Overview



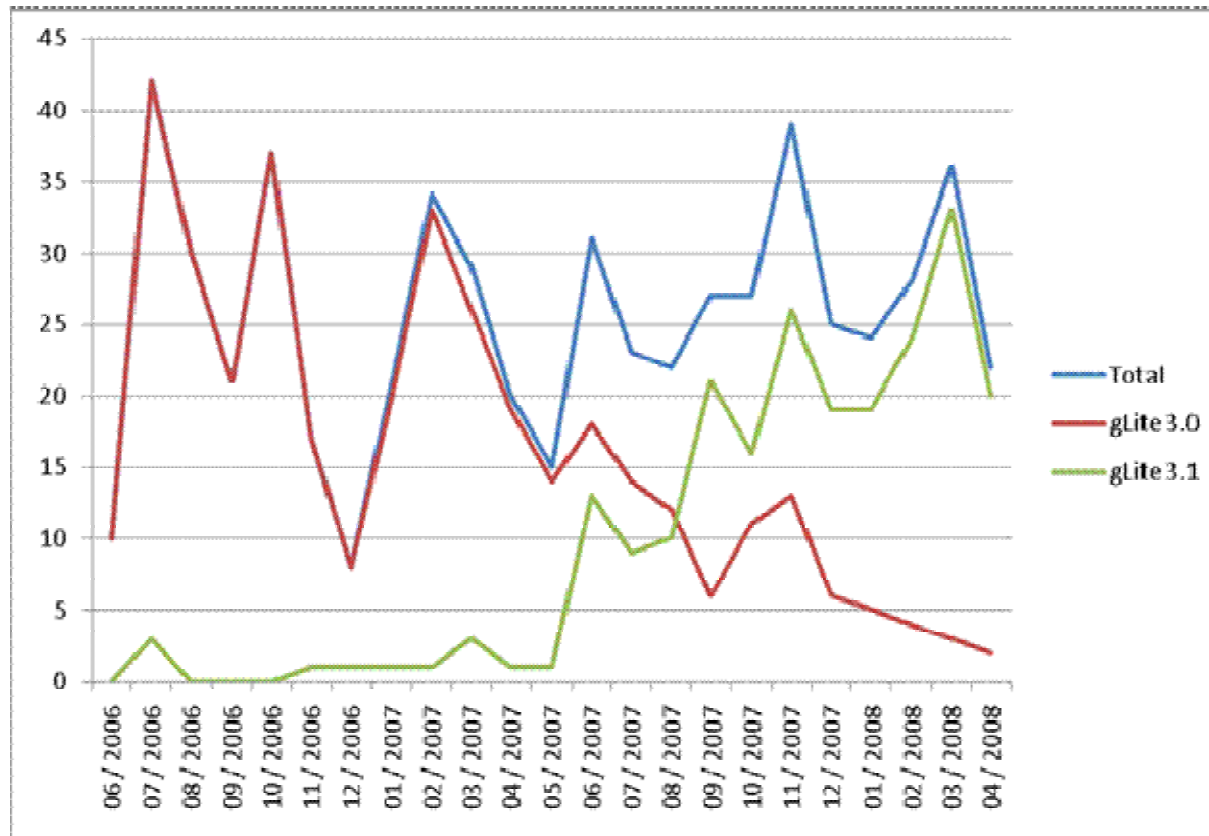
- Discussion on the status of
 - gLite
 - Status summary
 - CCRC and issues raised
 - Future prospects
 - OSG
 - Alain Roy
 - ARC
 - Oxana Smirnova
- Storage and SRM
 - Flavia Donno

gLite: Current Status



- Current middleware stack is gLite 3.1
 - Available on SL4 32bit
 - Clients and selected services available also on 64bit
 - Represents ~15 services
 - Only awaiting FTS to be available on SL4 (in certification)
- Updates are released every week
 - Updates are sets of patches to components
- Components can evolve independently
- Release process includes full certification phase
- Includes
 - DPM, dCache and FTS for storage management
 - LFC and AMGA for catalogues
 - WMS/LB and lcg-CE for workload
 - Clients for WN and UI
 - BDII for the Information System
 - Various other services (eg VOMS)

gLite patch statistics



New service types are introduced as patches

CCRC08 May Summary



- During CCRC we
 - Introduced new services
 - Handled security issues
 - Produced the regular stream of updates
 - Responded to CCRC specific issues
- The software process operated as usual
 - No special treatment for CCRC
 - Priorities were reviewed twice a week in the EMT
- 4 Updates to gLite 3.1 on 32bit
 - 18 Patches
- 2 Updates to gLite 3.1 on 64bit
 - 7 Patches
- 1 Update to gLite 3.0 on SL3
 - 1 Patch

Normal Operation

CCRC08 May Highlights



- CCRC demonstrated that the software **process** was functioning
- It did not reveal any fundamental problems with the stack
 - Various minor issues were promptly addressed
 - Scalability is always a concern
 - But didn't affect CCRC08
- The **WMS/LB** for gLite3.1/SL4 was released
- An implementation of **Job Priorities** was released
 - Not picked up by many sites
- First gLite release of **dCache 1.8** was made
- **WN** for x86_64 was released
- An improved **LCG-CE** had been released the week before
 - Significant performance/scalability improvements
 - Covers our resource access needs until the CREAM-CE reaches production



Middleware Baseline at start of CCRC



Component	Patch #	Status
LCG CE	Patch #1752	Released gLite 3.1 Update 20
FTS (T0)	Patch #1740	Released gLite 3.0 Update 42
FTS (T1)	Patch #1671	Released gLite 3.0 Update 41
gFAL/lcg_utils	Patch #1738	Released gLite 3.1 Update 20
DPM 1.6.7-4	Patch #1706	Released gLite 3.1 Update 18

- All other packages were assumed to be at the latest version
 - This led to some confusion.....
- An explicit list will have to be maintained from now on

Outstanding middleware issues



- **Scalability of lcg-CE**
 - Some improvements were made, but there are architectural limits
 - Additional improvements are being tested at the moment
 - Acceptable performance for the next year
- **Multiplatform support**
 - Work ongoing for full support of SL5 and Debian 4
 - SuSE 9 available with limited support
- **Information system scalability**
 - Incremental improvements continue
 - With recent client improvements the system performs well
 - Glue 2 will allow major steps forward
- **Main storage issues**
 - SRM 2.2 issues
 - Consistency between SRM implementations
 - Authorisation
 - Covered in more detail in the storage part

Forthcoming attractions



- **CREAM CE**
 - In the last steps of certification (weeks)
 - Addresses scalability problems
 - Standards based --> eases interoperability
 - Allows parameter passing
 - No GRAM interface
 - Redundant deployment may be possible within a year
- **WMS/ICE**
 - Scalability tests are starting now
 - ICE is required on the WMS to submit to CREAM
- **glexec**
 - Entering PPS after some security fixes
 - To be deployed on WNs to allow secure multi-user pilot jobs
 - Requires SCAS for use on large sites

Forthcoming attractions



■ SCAS

- Under development, stress tests started at NIKHEF
 - Can be expected to reach production in a few months
- With glexec it ensures site wide user ID synchronization
- This is a critical new services and requires deep certification

■ FTS

- New version in certification
 - SL4 support
 - Handling of space tokens
 - Error classification
- Within 6 months
 - Split of SRM negotiations and gridFTP
 - Improved throughput
 - Improved logging
 - Syslog target
 - Streamlined format for correlation with other logs
 - Full VOMS support
 - Short term changes for "Addendum to the SRM v2.2 WLCG Usage Agreement"
 - Python client libraries

Forthcoming attractions



- **Monitoring**
 - Nagios / ActiveMQ
 - Has an impact on middleware and operations, but is not middleware
- **Distribution of middleware clients**
 - Cooperation with Application Area
 - Currently via tar balls within the AA repository
 - Active distribution of concurrent versions
 - Technology for this is (almost) ready
 - Large scale tests have to be run and policies defined
- **SL5/VDT1.10 release**
 - Execution of the plan has started
 - WN (clients) expected to be ready by end of September
- **Glue2**
 - Standardization within OGF is in the final stages
 - Non backward compatible, but much improved
 - Work on implementation and phase-in plan started

OSG



- Slides provided by [Alain Roy](#)
- OSG Software
- OSG Facility
- CCRC and last 6 months
- Interoperability and Interoperation
- Issues and next steps



OSG for the LHCC Mini Review

Created by Alain Roy
Edited and reviewed by others
(including Markus Schulz)

- OSG Software
 - Goal: Provide software needed by VOs within the OSG, and by other users of the VDT, including EGEE and LCG.
 - Work is mainly software integration and packaging: we do not do software development
 - Work closely with external software providers
 - OSG Software is component of the OSG Facility, and works closely with other facility groups

- Engagement
 - Reach out to new users
- Integration
 - Quality assurance of OSG releases
- Operations
 - Day-to-day running of OSG Facility
- Security
 - Operational security
- Software
 - Software integration and packaging
- Troubleshooting
 - Debug the hard end-to-end problems

- Responded to ATLAS/CMS/LCG/EGEE update requests and fixes,
- MyProxy fixed for LCG(threading problem)
 - in certification now
- Globus proxy chain length problem fixed for LCG
 - in certification now
- Fixed security problem (rpath) reported by LCG
- **OSG 1.0 released**—major new release
 - Added lcg-utils for OSG software installations
 - Now use system version of OpenSSL instead of Globus's OpenSSL (more secure)
 - Introduced and supported SRM V2 based storage software based on dCache, Bestman, and xrootd.

- Big improvements in reporting
 - New RSV software collects information about functioning of sites, reports it to GridView. Much improved software, better information reported.
 - Generic Information Provider
 - OSG-specific version
 - greatly improved to provide more accurate information about a site to the BDII.



- US OSG sites need more LCG client tools for data management (LFC tools, etc.) .
- Working to improve interoperability via testing between OSG's ITB and EGEE's PPS.
- gLite plans to move to new version of VDT:
 - We'll help with the transition.
- We have regular communication
 - via Alain Roy's attendance at the EMT meetings.

ARC



- Slides provided by Oxana Smirnova
- ARC introduction and highlights
- ARC-classic status
- ARC-1 Status

Advanced Resource Connector in a nutshell

Concept

- General purpose Open Source European Grid middleware
 - One of the major production grid middlewares
 - Developed & maintained by the NorduGrid Collaboration
 - Deployment support, extensive documentation, available on most of the popular Linux distributions
- Lightweight architecture for a dynamic heterogeneous system following Scandinavian design principles
 - start with something simple that works for users and add functionality gradually
 - non-intrusive on the server side
 - Flexible & powerful on the client side
- User- & performance-driven development
 - Production quality software since May 2002
 - First middleware ever to contribute to HEP data challenge
- Strong commitment to standards & interoperability
 - JSDL, GLUE, Active OGF player
- Middleware of choice by many national grid infrastructures due to its technical merits
 - SweGrid, SWISS Grid(s), Finnish M-Grid, NDGF, etc...
 - Majority of ARC users are NOT from the HEP community



Illustrations: "Scandinavian Design beyond the Myth"

www.scandesign.org

ARC Classic: overview

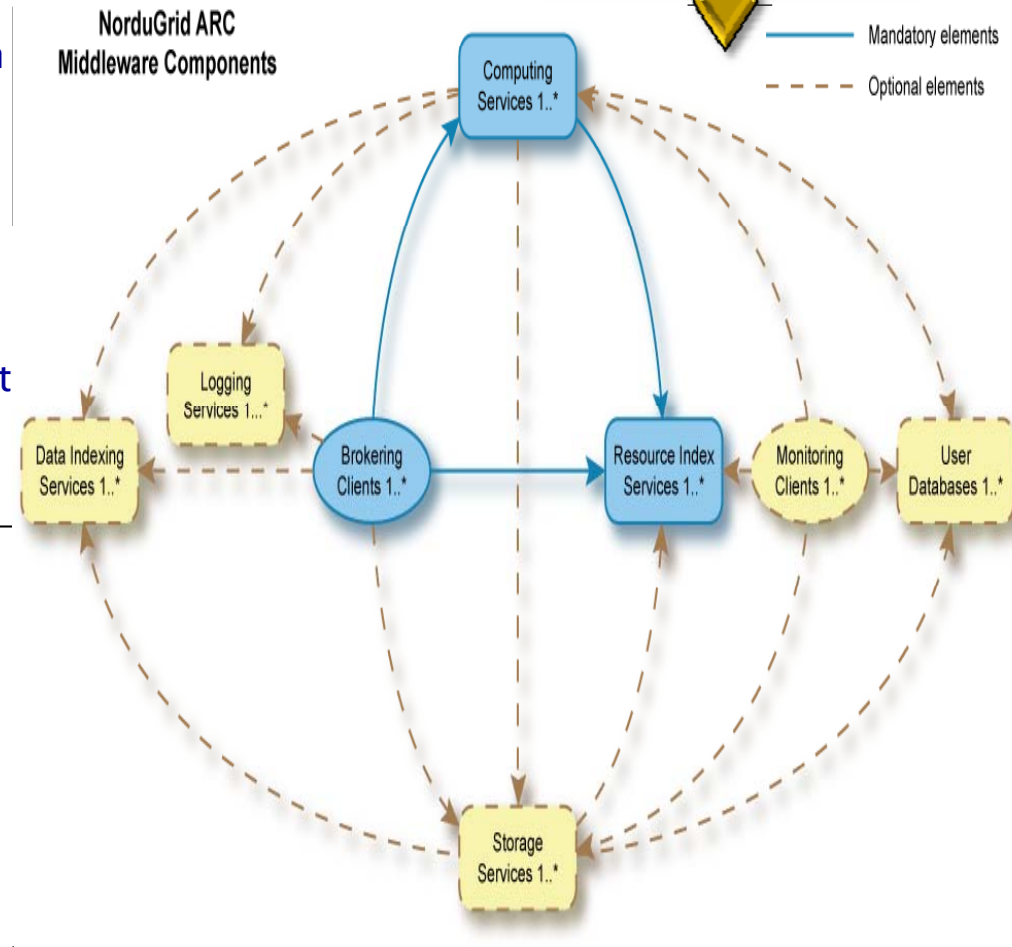
Concept

Provides reliable implementation of fundamental Grid services:

- The usual grid security: single sign on, Grid ACLs (GACL), VOs (VOMS)
- Job submission: direct or via matchmaking and brokering
- Job monitoring & management
- Information services: resource aggregation, representation, discovery and monitoring
- Implements core data management functionality
 - Automated seamless input/output data movement
 - Interfacing to Data Indexing, client-side data movement
 - Storage Elements
- Logging service

Builds upon standard open source solutions and protocols

- Globus Toolkit® pre-WS API and libraries (no services!)
- OpenLDAP, OpenSSL, SASL, SOAP, GridFTP, GSI



ARC Classic: status and plans

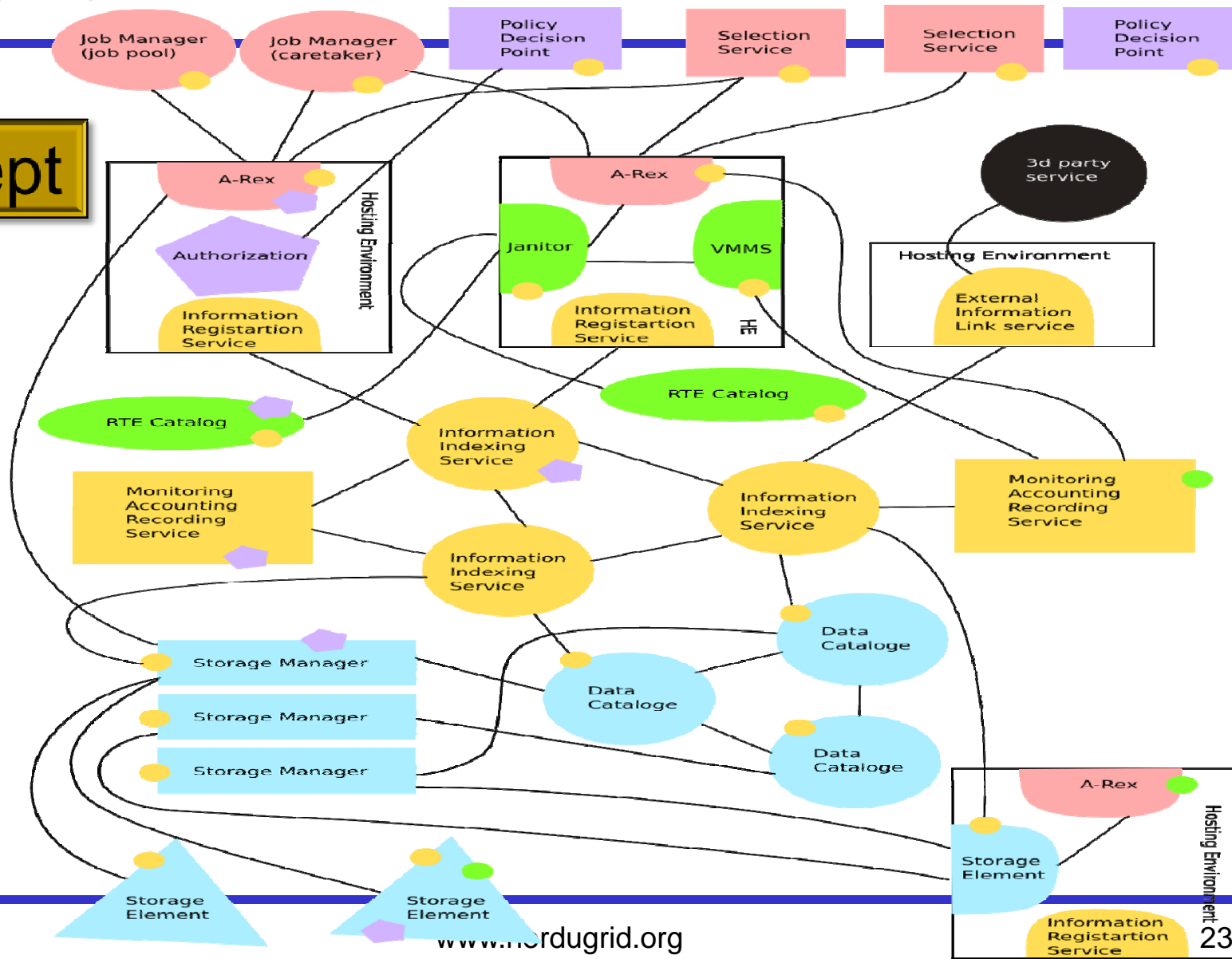
- ARC is a mature software which has proven its strenghts in numerous areas
- Production release 0.6.3 is out
 - Stability improvements, bug fixes
 - LFC (file catalogue) support
- ARC faces a scalability challenge posed by "ten-thousand-core" clusters
 - File cache redesign is necessary (ongoing)
 - Uploaders/downloaders load on frontends needs to be optimized
 - Local information system needs to be improved (BDII is being tested)
- Release plans based on ARC Classic
 - 0.6.x stable releases will be periodically released
 - Preliminary planning for an 0.8 release incorporating new major features
- Future: ARC1
 - A Web-service based solution (working prototypes exist)
 - Main goal: achieve interoperability via standards conformance (e.g. BES, JSDL, GLUE, SRM, GridFTP, X509, SAML)
 - Migration plan: gradually replace components with ARC1 modules, initially co-deploying both versions



Status

ARC1 Design: service decomposition

Concept



Design document: www.knowarc.eu/documents/knowarc_D1.1

SRM



- Slides provided by Flavia Donno
- Goal of the SRM v2.2 Addendum
- Short term solution
 - CASTOR, dCache, STORM, DPM



The Addendum to the WLCG SRM v2.2 Usage Agreement



Flavia Donno
CERN/IT

LCG-LHCC Mini review
CERN 1 July 2008



The requirements



- The **goal of the SRM v2.2 Addendum** is to provide answers to the following (requirements and priorities given by the experiments)
- **Main requirements :**
 - Space protection (VOMS-awareness)
 - Space selection
 - Supported space types (T1D0, T1D1, T0D1)
 - Pre-staging, pinning
 - Space usage statistics
 - Tape optimization (reducing the number of mount operations, etc.)



The document



- The most recent version is v1.4 available on CCRC/SSWG twiki:

https://twiki.cern.ch/twiki/pub/LCG/WLCGCommonComputingReadinessChallenges/WLCG_SRMv22_Memo-14.pdf

- 2 main parts:
 - An implementation-specific with limited capabilities short-term solution that can be made available by the end of 2008
 - A detailed description of an implementation-independent full solution
- The document has been agreed by *storage developers, clients developers, experiments* (ATLAS, CMS, LHCb)



The recommendations and priorities



- **24 June 2004** : WLCG Management Board approval with the following recommendations

- **Top priority is services functionality, stability, reliability and performance**

- The implementation plan for the “**short term solution**” can start

- it introduces the minimal new features needed to guarantee protection from resources abuse and support for data reprocessing

- It offers limited functionality and is implementation specific with clients hiding differences

- The long term solution is a solid starting point

- It is what is technically missing/needed

- Details and functionalities can be re-discussed with acquired experience



The short-term solution



■ **CASTOR**

- **Space Protection** based on UID/GID.
 - Administrative interface ready by 3rd quarter of 2008.
- **Space selection** already available.
 - srmPurgeFromSpace available in 3rd quarter of 2008.
- **Space types**
 - **T1D1** provided.
 - pinLifeTime parameter negotiated to be always equal to a system defined default.
- **Space statistics and tape optimization**
 - Already addressed
- **Implementation plan**
 - Available by the end of 2008



The short-term solution



■ dCache

■ Space Protection

- Protected creation and usage of “write” space tokens
- Controlling access to the tape system by DN's or FQAN's.

■ Space selection

- Based on the IP number of the client, the requested transfer protocol or the path of the file.
- Use of SRM special structures for more refined selection of read buffers.

■ Space types

- T1D0 + pinning provided. Releasing pins will be possible for a specific DN or FQAN.

■ Space statistics and tape optimization

- Already addressed

■ Implementation plan

- Available by the end of 2008



The short-term solution



■ DPM

■ Space Protection

- Support a list of VOMS FQANs for the space write permission check, rather than just the current single FQAN

■ Space selection

- Not available, not necessary at the moment

■ Space types

- Only T0D1

■ Space statistics and tape optimization

- Space statistics already available. Tape optimization not needed

■ Implementation plan

- Available by the end of 2008



The short-term solution



■ **StoRM**

■ Space Protection

- Spaces in StoRM will be protected via DN or FQAN based ACLs. StoRM is already VOMS-aware.

■ Space selection

- Not available

■ Space types

- T0D1 and T1D1 (no tape transitions allowed in WLCG)

■ Space statistics and tape optimization

- Space statistics already available. Tape optimization will be addressed

■ Implementation plan

- Available by November 2008



The short-term solution



■ Client tools: FTS, lcg-utils/gfal

■ Space selection

- The client tools will pass both the SPACE token and SRM special structures for refined selection of read/write pools.

■ Pinning

- Client tools will internally extend the pinlifetime of newly created copies.

■ Space Types

- The same SPACE Token might be implemented as T1D1 for CASTOR, or T1D0 + pinning in dCache. The clients will perform the needed operations to release copies in both types of spaces transparently.

■ Implementation plan

- Available by the end of 2008



Summary



- gLite handled the CCRC load
 - Release process was not affected by CCRC
 - Improved lcg-CE can bridge the gap until we move to CREAM
 - Scaling for core services is understood

- OSG
 - Several minor fixes and improvements
 - User driven
 - released OSG-1.0
 - SRM-2 support
 - Lcg data management clients
 - Improved interoperation and interoperability
 - Middleware ready for show time

Summary



■ ARC

- ARC 0.6.3 has been released
 - Stability improvements, bug fixes
 - LFC (file catalogue) support
 - Will be maintained
- ARC 0.8 release planning started
 - Will address several scaling issues for very large sites (>10k)
- ARC-1 prototype exists

■ SRM-2.2

- SRM v2.2 Addendum
 - Agreement has been reached
- Short term plan implementation can start
- Affects Castor, dCache, Storm, DPM and clients
 - Clients will hide differences
 - Expected to be in production by the end of the year
- Long term plan can still reflect new experiences