LHCC referees meeting, 29 June 2004

"**ARDA status**"
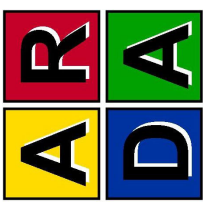
Massimo Lamanna / CERN

http://cern.ch/arda

Enabling Grids for
E-science in Europe

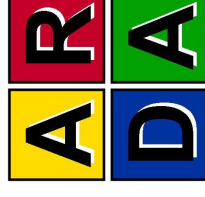www.eu-egee.org   cern.ch/lcg

Contents

- The project in a nutshell

- Highlights from the 4 experiment prototypes

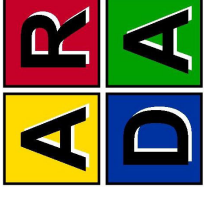- ARDA-related workshops

- Conclusions

ARDA in a nutshell

- ARDA is an LCG project whose main activity is to enable LHC analysis on the grid
- ARDA is coherently contributing to EGEE NA4 (using the entire CERN NA4-HEP resource)
- Use the grid software as it matures (EGEE project)
  - ARDA should be the key player in the evolution from LCG2 to the EGEE infrastructure
  - Provide early and continuous feedback (guarantee the software is what experiments expect/need)
- Use the last years experience/components both from Grid projects (LCG, VDT, EDG) and experiments middleware/tools (Alien, Dirac, GAE, Octopus, Ganga, Dial,…)
  - Help in adapting/interfacing (direct help within the experiments)
  - Every experiment has different implementations of the standard services, but:
    - Used mainly in production environments
      - **Few expert users**
      - **Coordinated update and read actions**
    - ARDA
      - Interface with the EGEE middleware
      - Verify (help to evolve to) such components to analysis environments
        - Many users (Robustness might be an issue)
        - Concurrent "read" actions (Performance will be more and more an
- One prototype per experiment
  - A Common Application Layer might emerge in future
  - ARDA emphasis is to enable each of the experiment to do its job
- Provide a forum for discussion
  - Comparison on results/experience/ideas
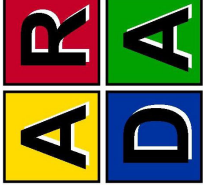 - Interaction with other projects
  - …

The experiment interfaces agree with the ARDA project leader the work plan and coordinate the activity on the experiment side (users)

LHCb

- Grid activity:
  - Use of the Glite testbed (since May 18th)
  - Test jobs from Ganga to Glite ☺
- Other contributions:
  - GANGA interface to Condor (Job submission) and Condor DAGMAN for splitting/merging and error recovery
  - GANGA Release management and software process
  - LHCb Metadata catalogue tests
    - Performance tests
      - Collaborators in Taiwan (ARDA + local DB know-how on Oracle)

- The LHCb system within ARDA uses GANGA as main component.
- The LHCb/GANGA plans:
  - enable physicists (via GANGA) to analyse the data being produced during 2004 for their studies
  - It naturally matches the ARDA mandate
  - Deploy the prototype where the LHCb data will be the essential (CERN, RAL, …)
- At the beginning, the emphasis will be to validate the tool focusing on usability, validation of the splitting and merging functionality for users jobs
- DIRAC (LHCb production grid): convergence with GANGA / components / experience

CMS

RefDB in CMS DC04

The CMS system within ARDA is still under discussion

- Provide eas... sharing) of ... key issue (D...
  - RefDB is... and stee... phases ... some de...
    - This...
  - It contai... except f... info rela... system (...
  - The actu... data to a...
  - Measuri... (similar ... Metadat...
- Exploratory...
  - ORCA j...
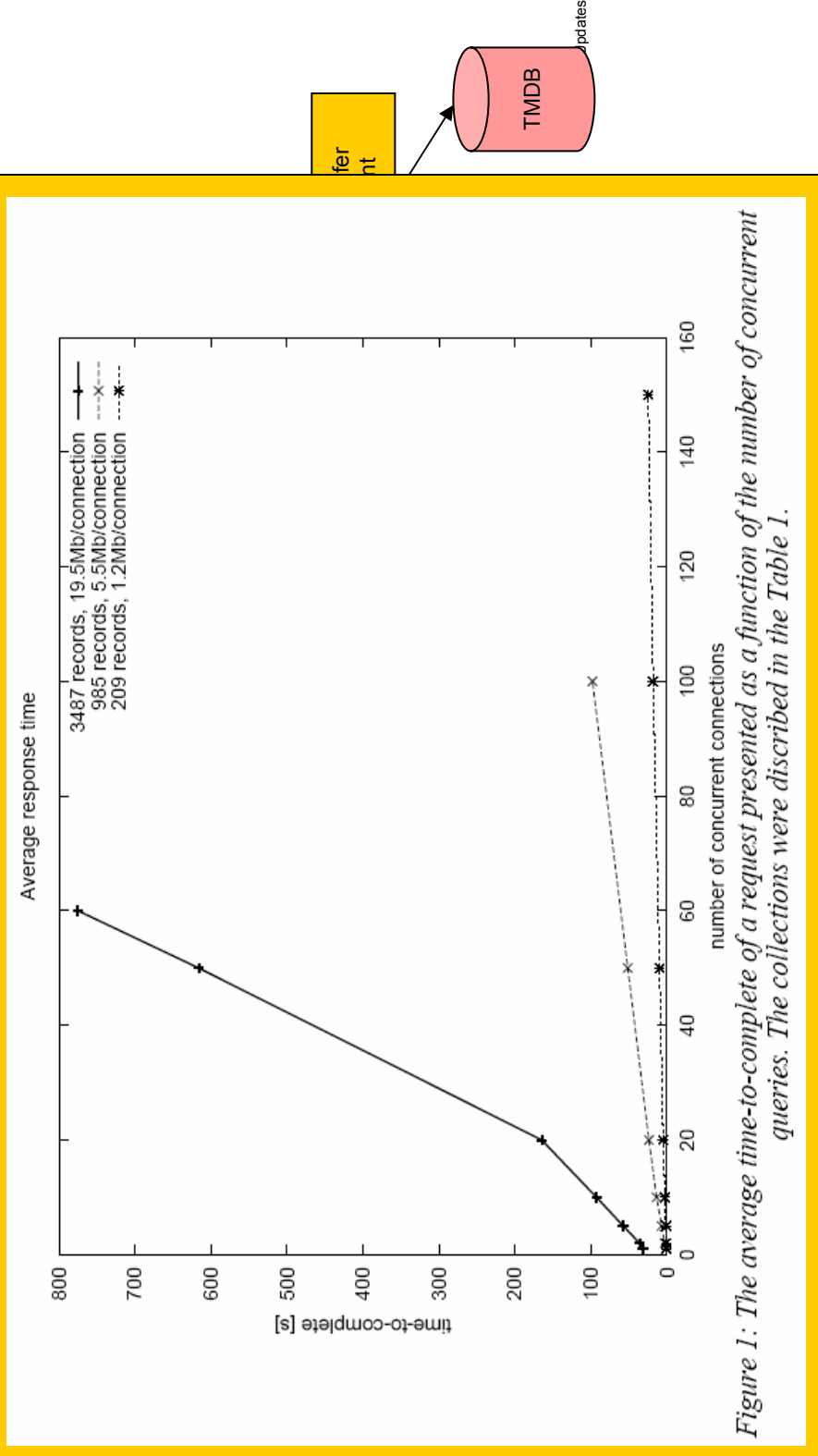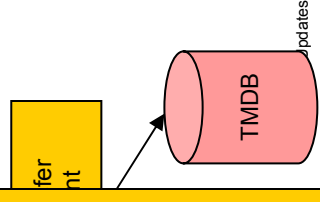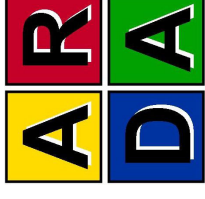 - Glite file

TMDB

fer
nt

pdates

Figure 1: The average time-to-complete of a request presented as a function of the number of concurrent queries. The collections were discribed in the Table 1.

Average response time

Legend:
- 3487 records, 19.5Mb/connection
- 985 records, 5.5Mb/connection
- 209 records, 1.2Mb/connection

x-axis: number of concurrent connections
y-axis: time-to-complete [s]

ATLAS

- The ATLAS system within ARDA has been agreed
  - ATLAS has a complex strategy for distributed analysis, addressing different area with specific projects (www.usatlas.bnl.gov/ADA)
  - Starting point is: DIAL analysis model (high level web services)
- The AMI metadata catalog is a key component
  - Robustness and performance tests from ARDA
  - Very good relationship with the ATLAS Grenoble group
  - Discussions on technology (EGEE JRA1 in the loop)
- In the start up phase, ARDA provided help in developing ATLAS production tools

- Submission to Glite (via simplified DIAL system now possible ☺)
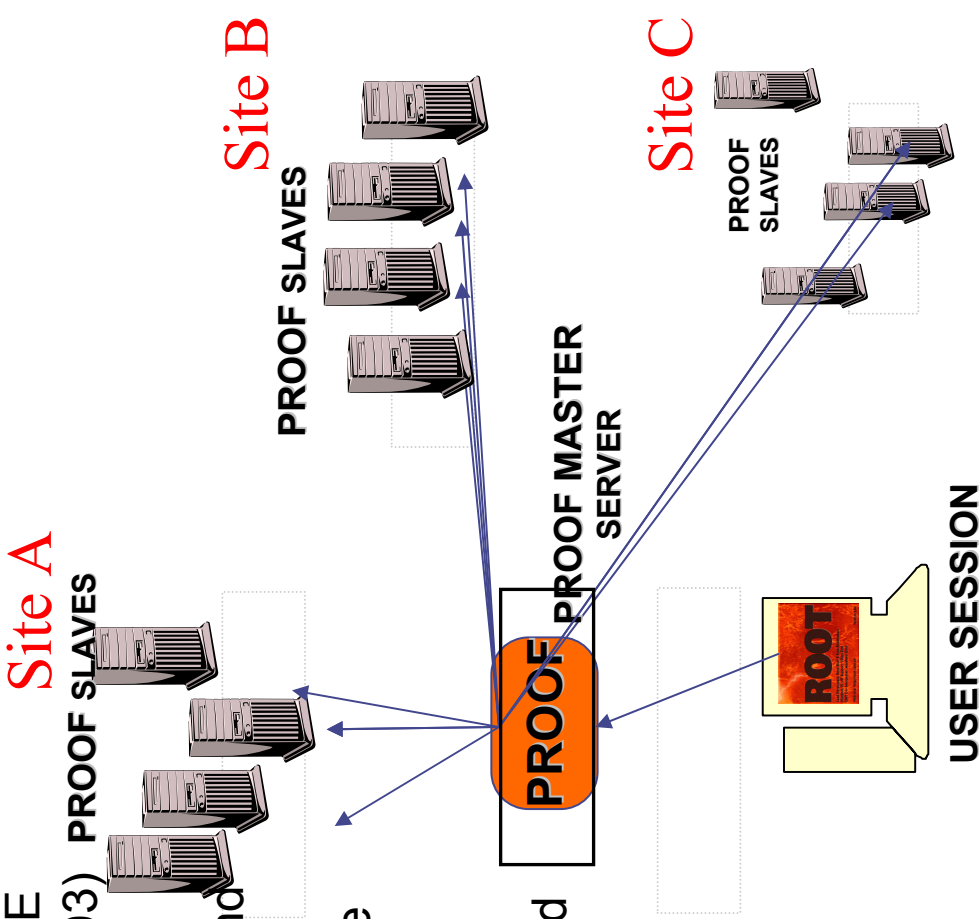- First skeleton of high level services

AMI Tests

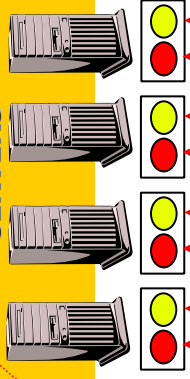| Clients				Rows in Response			
	5	10	20	50	100	150
1	0.22	0.27	0.35	0.87	2.49	5.26
5	0.40	0.48	0.74	2.94	10.99	27.98
10	0.67	0.75	1.74	4.77	21.99	56.17
20	1.02	1.34	2.46	9.51	41.79	timeout
30	1.42	2.36	3.10	14.21	66.61	timeout
40	1.80	2.33	4.84	19.94	timeout	timeout
50	2.32	6.43	5.02	21.43	timeout	timeout
100	9.94	9.82	SOAP-Err	SOAP-Err		
150	16.51	SOAP-Err				

ALICE
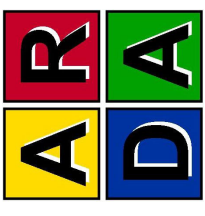
- Strategy:
  - The ALICE/ARDA will evolve the ALICE analysis system (SuperComputing 2003)
- Where to improve:
  - Strong requests on networking (inbound connectivity)
  - Heavily connected with the middleware services
  - "Inflexible" configuration
  - No chance to use PROOF on federated grids like LCG in AliEn
  - User libraries distribution
- Activity on PROOF
  - Robustness and Error recovery
- Grid activity:
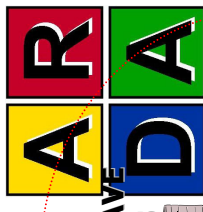 - First contact with the Glite testbed
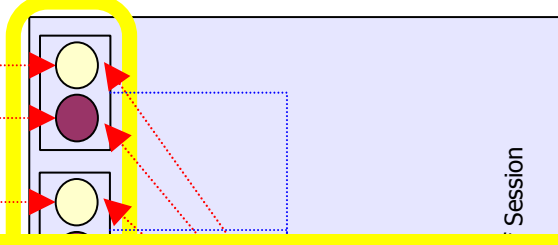 - C++ access library on Glite ☺

Site A

Site B

Site C

PROOF SLAVES

PROOF SLAVES

PROOF SLAVES

PROOF

PROOF MASTER SERVER

ROOT

USER SESSION

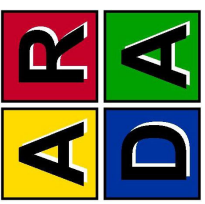Status report an as a demo during the workshop
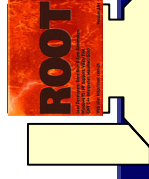
"The first 30 days of the EGEE middleware" ARDA workshop

- 1st ARDA workshop (January 2004 at CERN; open)

  - •Closed meeting
  - •Focus on the experiments (more than on the MW)
  - •Lot of time for technical discussion (never enough...)
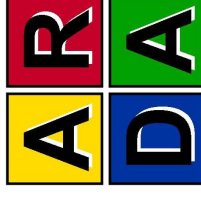
- 3rd **ARDA workshop (September 2004; open)**
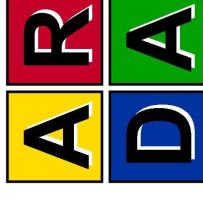 - "The ARDA end-to-end prototypes"

ARDA Workshop

- Programme

http://agenda.cern.ch/fullAgenda.php?ida=a042197&stylesheet=tools/printable&dl=&dd=

"The first 30 days of the EGEE middleware" ARDA workshop

- Effectively, this is the 2nd workshop (January '04 workshop)
- Given the new situation:
  - Glite middleware becoming available
  - LCG ARDA project started
  - Experience + need of technical discussions
- New format:
  - "Small" (30 participants vs 150 in January)
  - To have it small, by invitation only....
  - ARDA team + experiments interfaces
  - EGEE Glite team (selected persons)
  - Experiments technical key persons
  - Technology experts
  - NA4/EGEE links (4 persons)
  - EGEE PTF chair
- Info on the web:
 - URL:http://lcg.web.cern.ch/LCG/peb/arda/LCG_ARDA_Workshops.htm

Prototype info (F. Hemmer)
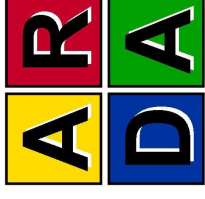
- A *First* Prototype Middleware on a testbed at CERN and Wisconsin delivered to ARDA on May 18, 2004 and to NA4/Biomed on June 15, 2004
  - Being integrated in SCM
  - Being used by Testing Cluster
  - Prototype GAS service
  - Using Integration tools

- Significant contribution from University of Wisconsin Madison on
  - Adapting the Grid Manager for interfacing to PBS/LSF
  - Supporting and debugging the prototype
  - Contributing to the overall design
  - Interfacing with Globus and ISI

- DJRA1.2: preliminary work performed in the MW working document

Prototype info (F. Hemmer)
Status of gLite Prototype
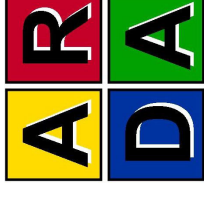
- A *Initial* Prototype Middleware on a testbed consisting of
  - AliEn "shell"
  - Job submission:
    - Alien CE->Condor-G->blahp->PBS/Condor
    - Globus Gatekeeper
  - Data Management
    - File catalog
      – Service factored out of Alien, Web Service interface; WSDL to be done
    - Castor & D-Cache SE with SRM
    - gridFTP for transfers
    - AliEn FTD
    - Aiod/GFal investigations
    - RLS (EDG)
      – Perl RLS Soap interface for File Catalog integration
      – Not used yet
  - Security
    - VOMS for certificate handling/SE gridmap files (NIKHEF)
    - MyProxy for certificate delegation in GAS
  - GAS (Grid Access Service)
    - Prototype with a few file cataloging/RLS functions
  - R-GMA
    - With new API; not used yet
- Being integrated in SCM

Prototype info (F. Hemmer)

## Next set of components to be added or changed
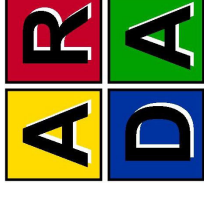
- Workload Management
  - Initial prototype WMS components supporting job submission and control, the handling of data requirements via RLS and POOL catalog queries, the ability for CEs to 'request' jobs, all while keeping LCG-2 compatibility.

- Information Services
  - R-GMA with new API
  - Redesign of Service/ServiceStatus tables and publishing mechanism

- SE
  - Finish File I/O design, integrate of AIO and GFAL with the security libraries, first prototype I/O system

- File Transfer Service
  - Integrate Condor Stork soon (needs to implement SRM interface)

- File Catalog
  - WSDL interface, clients in other languages

- Replica Catalog
  - Re-factored RLS, integrated with File Access Service

- Metadata Catalog
  - Initial implementation ready to be integrated in two weeks

- Grid Access service security model

ARDA and the prototype
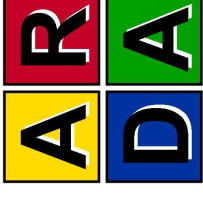
| May 18	First presentation of prototype to ARDA members
...	Obtaining certificates from CERN CA and registering for EGEE VO
June 2	most ARDA members cannot log in due to a registration/login name issue of an organizational nature.
June 9	Username problems solved for ARDA members. Learning to use the system.
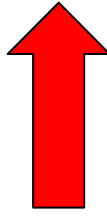
Actual testing took place during ~ 2 weeks
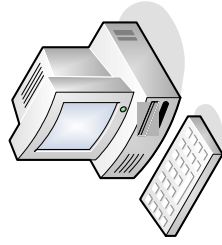
D. Feichtinger presentation

Testbed

WN (lxn5211)

CE (lxn5210)
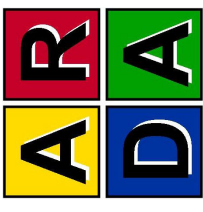
More resources (boxes and sites) absolutely needed. "Strategic" sites (Key Tier1s for the experiments) should start planning to join
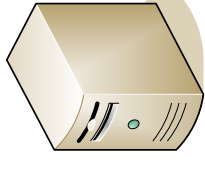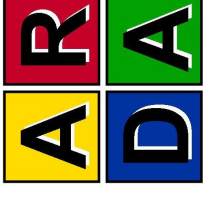
Issues:
- Synchronise with the preparation work in EGEE SA1 (Operations)
- Licensing scheme!

Source: http://egee-jra1.web.cern.ch/egee-jra1/Prototype/testbed.html

Glite: Tests done

- GAS was only cursorily tested due to its current limited functionality (simple catalogue interactions)

- Trivial tests exercising the basic catalogue and job execution functionality

- Running a few more complex tests using applications installed via AFS as a workaround for the not yet available package management.

- Some first mass registration and metadata tests
  - following same procedure as has been done for the ATLAS AMI and LHCb metadata catalogues
    - Writing of large number of files (~ 100'000) shows that storage time per files increases

- First simple GANGA plugin for job submission and monitoring of job status

- ARDA group members associated with experiments trying to get more complicated, experiment related applications to run (target is four end to end prototypes).
  - ATHENA jobs via DIAL
  - Work in progress for ORCA and DaVinci

- Developing a new C++ API and plugin for ROOT, which interacts efficiently with the present system (commands are executed via a service situated close to the core system. All transfers use a single custom encoded SOAP string)

Glite: General Problems (1)
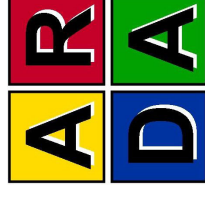
Since the newest update was installed only last Friday some issues may already have been resolved.

- Stability of the services
- Tutorial documentation good, but does not provide much help for more than basic usage.
- Documentation does not cover the commands extensively (options to commands missing and also Syntax for some arguments)
- High initial latency for job execution even if queue is empty (up to 10 min.)

Glite: General Problems (2)
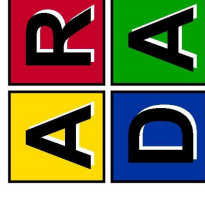
Minor issues:

- Online help behavior of commands not useful for unary commands. There should be a standard help flag like –h.

- Some messages returned to the user are cryptic and have no obvious connection with the action (~ more like debug messages).

- Easier retrieval of a job's output sandbox

- Question: should users use the "debug #" command to have a better error description when they submit bugs?

Infrastructure/prototype issues
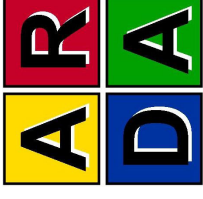
- First contact with Glite positive

- Prototype approach positive

  - Incremental changes, fast feedback cycle
  - Workshop useful to start iterating on priorities
  - See Derek's presentation

- Wish list on the current prototype installation:

  - **More resources!**
  - **More sites!**
  - **Improve on infrastructure**
    - Procedures should be improved
    - Service stability
    - *Move to a pre-production service as soon as possible*

GAS and shell
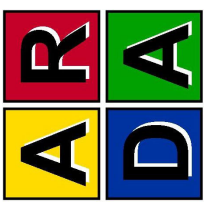
- Useful concepts/tools
- GAS as "Template" service for the experiments to provide their orchestration layer
  - **Individual services accessible as such**
  - **They should expose "atomic" operations**
  - **Custom GAS can be provided by the experiment**
    - **API needed**
- Shell extension preferred to dedicated shell emulation
  - Cfr. ALICE/ARDA demo presentation

General/Architectural concerns
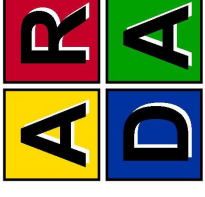
- Next month home work:
  - **List of LHC experiments priorities put together by ARDA**

- Connections with the architecture team
  - Strong need to see and interact at the architectural level
  - *At least agree on a mailing list....*

- **Priority is on developing the new system functionality**
  - **Whenever possible backward compatibility is a nice feature**
  - **Investment on the experiment side**

- Long standing issues:
  - Outbound connectivity
    - Functionality needed
    - Mechanism to provided needed (not necessarily direct connectivity)

  - Identity of user jobs running in a remote site
    - "gridmapfile / setuid" is an implementation
    - What is really needed is traceability

Database related
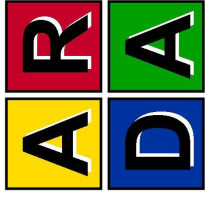
- Pervasive infrastructure and/or functionality
  - File catalogue →
  - Metadata catalogue →
  - Other catalogues (cfr. ATLAS talk)
- Plan for the infrastructure
  - Probably we are already late… hopefully not too late
- Different views on how to achieve
  - Load balance // remote access // replication
  - Disconnected operations
  - Level of consistency
  - Uniform "access"
    - GSI authentication
   - VO role mapping (authorisation)

File catalog related
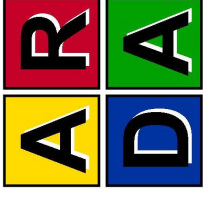
- Experimenting in ARDA has started
- Looks realistic that the File Catalog interface "validated" via independent experience
  - Cfr. Andrei Tsaregorodtsev's LHCb talk
- LFN vs (GU)ID discussion
  - Not conclusive but interesting
  - Bulk of the experiment data store might be built with files with a GUID (and metadata) but without LFN
    - POOL/ROOT compatible (GUID used in navigation, not LFN)
    - Need further exploration!!!
  - **Bulk of the experiment files: WORM**
    - **Working decision**
    - **Other experiment-dependent mechanisms to cover modifiable files needed**
   - **N.B. modifiable file = same GUID but different content**

Metadata database related
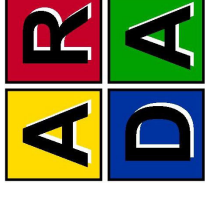
- Very many parallel initiatives!
  - in the LHC experiments
  - across the experiments (gridPP MDC)
  - in the MW providers community (LCG, OGSA-DAI, Glite, …)
  - Experiments interested in having input on technology
  - No clear plan of convergence
- **Web services is the way to go?**
  - **Or is it just a data base problem?**
  - **"any" GUID-Value pair system could do it**
    - **Maybe GUID-GUID pairs**
- Anatomy vs. physiology
  - The client API does not map 1:1 onto WSDL
  - Access method != data transfer protocol
  - This could trigger requests for specific client bindings

Data management related
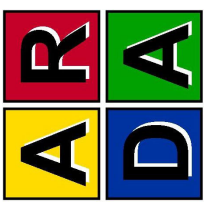
- Experimenting in ARDA has started
  - Not all components

- **Data management services needed**
  - **Data management transfer services**
    - Needed!
    - **ML: Can Glite provide the TM(DB) service?**
      - Reliable file transfer is a key ingredient
      - This functionality was implemented to fill a gap in DC04
      - CMS is entering a "continuous" production mode: room for constructive collaboration
      - Misuse protection (cfr. ALICE talk)

Application software "installation" (package management)
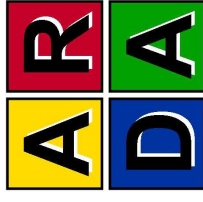
- All experiments have/use one or more solutions

- **Interest is high for a common solution, but the priority is not very clear**
  - **Service to manage local installation**

- Some features:
  - Lazy installation
    - The job expresses what it needs, the missing part is installed
  - Triggered for all resources
    - Overwrite the default installation for one experiment
  - Installed software is a resource
    - To be advertised: could be used in the matchmaking process

Conclusions
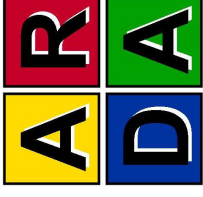
- Up and running
  - ▪ Since April the 1st (actually before of that) preparing the ground for the experiments prototype
    - Definition of the detailed programme of work
    - Contributions in the experiment-specific domain
  - ▪ 3 out of 4 prototype activity started
  - ▪ CMS prototype definition late by 1 month ("preliminary" activity going on)
- Playing with the Glite middleware since 30 days…
- Fruitful workshop 21-23 June
- Stay tuned ☺