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 issue)**
- 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
    - About 2 FTEs per prototype
- Provide a forum for discussion
  - Comparison on results/experience/ideas
  - Interaction with other projects
  - …

Experiment interfaces:
Piergiorgio Cerello (ALICE)
David Adams (ATLAS)
Lucia Silvestris (CMS)
Ulrik Egede (LHCb)

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

Milestones (Level 1)

LCG ARDA End-To-End Prototype activity

Milestone	Date	Description
1.6.18	Dec 2004	E2E prototype for each experiments (4 prototypes), capable of analysis (or advanced production)
1.6.19	Dec 2005	E2E prototype for each experiments (4 prototypes), capable of analysis and production

Prototype

- Available for us since May 18th
- In the first month, many problems connected with the stability of the service and procedures
- A second site (Madison) available at the end of June
- At that point just a few worker nodes available
- Now the no. of CPU is increasing (50 as a target for CERN, hardware available) as well as the no. of sites
- CASTOR access to the actual data store being delivered recently (essential)
- Look for a larger installation to be able to attract users (few more sites, some ~100 CPUs)

LHCb

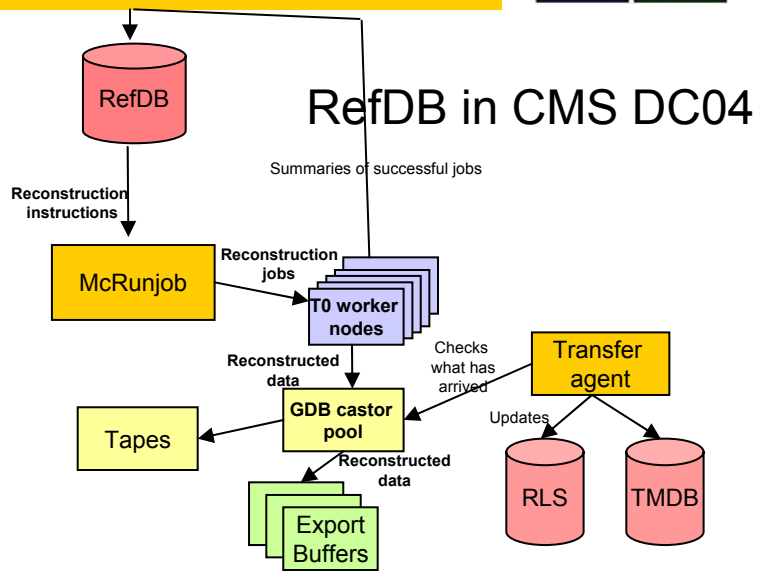
- The LHCb system within ARDA uses GANGA as main component.
- The LHCb/GANGA plans:
  - enable physicists (via GANGA) to analyse the data being produced during 2004 for their studies
  - It naturally matches the ARDA mandate
  - Deploy the prototype where the LHCb data will be the essential (CERN, RAL, …)
- At the beginning, the emphasis is 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

- Grid activity:
  - Use of the Glite testbed (since May 18th)
  - "Regular" DaVinci jobs onto Glite Published as an example and used by a LHCb test users (outside the ARDA team)
  - DaVinci 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 (CVS, Savannah,…)
  - Contributions to DIRAC
  - LHCb Metadata catalogue tests
    - Performance tests
    - Collaborators in Taiwan (ARDA + local DB know-how on Oracle)

CMS

- The CMS system within ARDA is still under discussion (Milestone 1.6.4 late by 3 months?)
- Provide easy access (and possibly sharing) of data for the CMS users is a key issue (Data management):
  - RefDB is the bookkeeping engine to plan and steer the production across different phases (simulation, reconstruction, to some degree into the analysis phase).
    - This service is under test
  - It contained all necessary information except file physical location (RLS) and info related to the transfer management system (TMDB)
  - The actual mechanism to provide these data to analysis users is under discussion
  - Measuring performances underway (similar philosophy as for the LHCb Metadata catalog measurements)
- Exploratory/preparatory activity
  - Successful ORCA job submission to Glite ☺. Published as an example.
  - Investigating with the package manager
  - Access to files directly from CASTOR
  - Glite file catalog tests



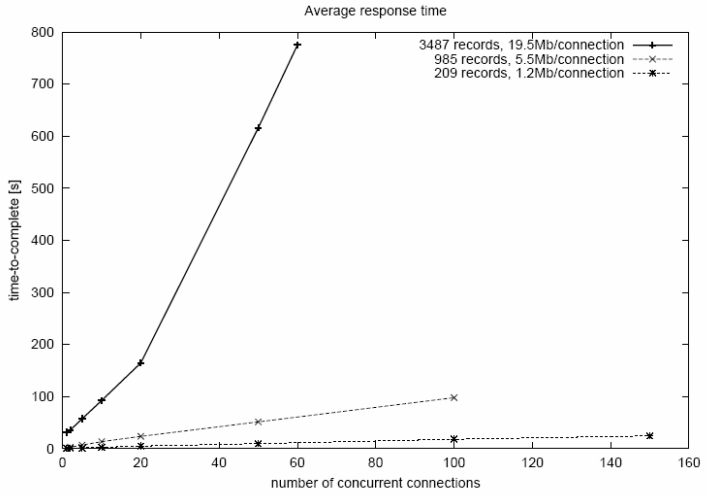


Figure 1: The average time-to-complete of a request presented as a function of the number of concurrent queries. The collections were discribed in the Table 1.

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 tools (ATCOM and CTB)

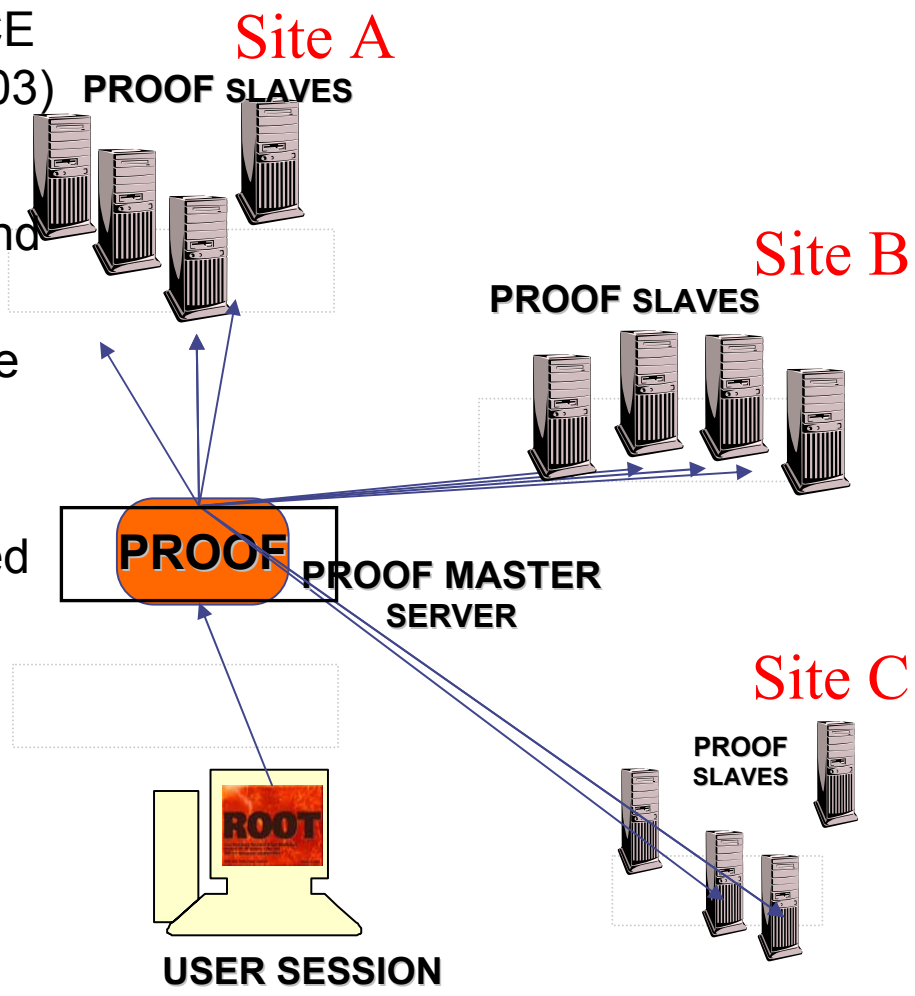
- DIAL on gLite OK (Evolution of the DIAL demo)
- ATHENA to gLite OK. Ready to be exposed to test users.
- First skeleton of high level services
- Detailed studies of the Don Quichote system (ATLAS data management)

AMI Tests

	Rows in Response					
Clients	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:
  - C++ access library on Glite ☺
  - IO library contributions



Status

ALICE

- Grid activity
  - Use of the Glite testbed
  - Access system to gLite services being developed (Demo available in June 2004)
    This is the key layer to allow ALICE software to be effectively use their prototype (evolution of their 2003 system, using gLite and PROOF)
- Other contributions
  - The access system is a generic piece of software (plans to be used in ATLAS and metadata access)
  - Tests of the metadata capabilities of the gLite file catalogue

ATLAS

- Grid Activity
  - Use of the Glite testbed
  - DIAL on gLite OK (Evolution of the DIAL demo)
  - ATHENA to gLite OK. Ready to expose to test users.
  - First skeleton of high level services
  - Detailed studies of the Don Quijote system (ATLAS data management system)
- Other contributions:
  - Detailed collaboration on AMI database (performance, support Oracle implementation,…)
  - Production and Combined Test Beam contributions

CMS

- The CMS system within ARDA not 100% defined
- Grid and other contributions:
  - Use of the Glite testbed
  - Successful ORCA job submission to Glite ☺.
  - Investigating with the package manager
  - Access to files directly from CASTOR
  - Glite file catalog tests
  - RefDB studies and evolution

LHCb

- Grid activity:
  - Use of the Glite testbed
  - “Regular” DaVinci jobs onto Glite exposed to test users (outside the ARDA team)
  - DaVinci jobs from Ganga to Glite
- Other contributions:
  - GANGA release mgt and software process (CVS,…)
  - Contributions to DIRAC
  - Metadata catalogue (Perform. test also in Taiwan)

ARDA workshops and liaison activity

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

- **2nd ARDA workshop (June 21-23 at CERN; by invitation)**
  - **"The first 30 days of EGEE middleware"**
  - **Main focus on LHC experiments and EGEE JRA1 (Glite)**

- EGEE NA4 (Applications) meeting mid July
  - NA4/JRA1 and NA4/SA1 sessions organised by M. Lamanna and F. Harris
  - EGEE/LCG operations new ingredient!

- **3rd ARDA workshop (October 20-22 2004; → registration open ←)**
  - "The LCG ARDA prototypes"
  - → http://lcg.web.cern.ch/LCG/peb/arda/LCG_ARDA_Workshops.htm

- EGEE Conference meeting mid November
  - NA4/JRA1 and NA4/SA1 sessions organised by M. Lamanna and F. Harris