http://cern.ch/arda

EGEE NA4 Meeting, 14-16 July 2004

"The ARDA project"

Massimo Lamanna / CERN

on behalf of the LCG-ARDA project

egee
Enabling Grids for
E-science in Europe

LCG

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

Contents

- "Historical" remarks
- The project in a nutshell
- ARDA-related workshops
- Highlights from the 4 experiment prototypes
- Conclusion

LCG ARDA working group recommendations
circa end 2003

- New service decomposition for a grid system
  - Strong influence of Alien system
    - the Grid system developed by the ALICE experiments and used by a wide scientific community (not only HEP)
- Role of experience, existing technology…
  - Web service framework

gLite Middleware

- Interfacing to existing middleware to enable their use in the experiment frameworks
- Early deployment of (a series of) end-to-end prototypes to ensure functionality and coherence

ARDA project

End-to-end prototypes: why?

- Provide a fast feedback to the EGEE MW development team
  - **Hit the road running!**
    - Avoid uncoordinated evolution of the middleware
    - Coherence between users expectations and final product
- Experiments ready to benefit from the new MW as soon as possible
  - **Hit the road running!!**
    - Frequent snapshots of the middleware available
    - Expose the experiments (and the community in charge of the deployment) to the current evolution of the whole system
    - Experiments system are very complex and still evolving
- Move forward towards new-generation real systems (analysis!)
  - **Hit the road running!!!**
    - Prototypes should be exercised with realistic workload and conditions
      - No academic exercises or synthetic demonstrations
      - LHC experiments users absolutely required here!!!
    - A lot of work (and useful software) is involved in current experiments data challenges: this will be used as a starting point
      - Adapt/complete/refactorise the existing: we do not need another system!

ARDA project 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
  - Complementary information to the LCG2 experience (e.g. Data challenges)
  - Provide early and continuous feedback (guarantee the software is what experiments expect/need)
  - Exposed to gLite since May 18th ☺
- 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 activities)
  - ARDA
    - Interface with the gLite 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
  - The emphasis is to enable each experiment to be ready to benefit from the new EGEE infrastructure!
- Provide a forum for discussion
  - Comparison on results/experience/ideas
  - Interaction with other projects
  - Organise ARDA workshops
  - Contribute to EGEE events

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

(ARDA) workshops

- 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"
- NA4 meeting mid July
  - NA4/JRA1 and NA4/SA1 joint sessions organised by M. Lamanna and F. Harris
- d ARDA workshop (beginning of October 2004; open)
  - 2-day meeting

New: Operations

Emphasis: gLite

Cross fertilisation between ARDA workshops (HEP focused) and NA4 events (broader scope and larger community)

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

- CERN: 21-23 of June 2004

- Monday, June 21
  - ARDA team / JRA1 team
  - ATLAS (Metadata database services for HEP experiments)

- Tuesday, June 22
  - LHCb (Experience in building Web Services for the Grid)
  - CMS (Data management)

- Wednesday, June 23
  - ALICE (Interactivity on the Grid)
  - Close out

"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" (30-35 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
  - <u>NA4/EGEE links (4 persons)</u>
  - <u>EGEE PTF chair</u>
- Info on the web:
  - URL:http://lcg.web.cern.ch/LCG/peb/arda/LCG_ARDA_Workshops.htm

Activity with the experiments

- ALICE
  - Interactive access to the gLite grid services
  - Good progress (also shown as a demo in the ARDA workshop)
- ATLAS
  - High level analysis services for the experiments (DIAL) on top of the gLite infrastructure
  - Collaboration
- CMS
  - Detailed plan
  - Preliminary analysis softw
- LHCb
  - Adapt their "uevolution of th
  - Collaboration on the LHCb metadata catalogue

See D. Feichtinger talk in the JRA1/NA4 session

Present status

- Usable software made available as promised ☺
  - We understand (and strongly support) the prototype model ("release early, release often" paradigm)
    - We understand that "early" might imply "incomplete"
  - What the pre-production system should look like?
    - See next slide
- Responsiveness of the JRA1 team
  - "High level"
    - Prototype access is an extra load
      - we believe it pays off for the common goal (i.e. high quality gLite software)
  - Bug fixes / workarounds / discussions
    - Good relationship, trust (developers working around the clock, positive reactions to input, constructive attitude!)

Where do we want to go from here (in the next ~ 3 months)?

- Present critical issues:

    JRA1 session

    - Stability of the gLite prototype

    SA1 session

    - Availability of significant resources (o(100) CPU sites, at least 3 sites required) to attract serious users

    JRA1 and SA1 sessions

    - Effective data access to significant fraction of the experiments data store
        - CASTOR storage element not yet available
        - "Key" POCs must join (some LCG Tier1s)

    - Coherent software distribution mechanism for the experiment software

    JRA1 session

Look forward to useful discussions in Catania!

Conclusions

- Up and running
- Main tool: end-to-end prototype
  - Definition of the detailed programme of work
  - Contributions in the experiment-specific domain
- Playing with the Glite middleware
- Experience discussed (and augmented) within the workshop activity
- Catania workshop
  - Discussions NA4/JRA1
  - Discussions NA4/SA1
- Stay tuned ☺