WP8 HEP Applications

Final Project evaluation of EDG middleware, and summary of workpackage achievements

F HARRIS
(OXFORD/CERN)
Email: f.harris@cern.ch

DataGrid is a project funded by the European Commission under contract IST-2000-25182

Outline

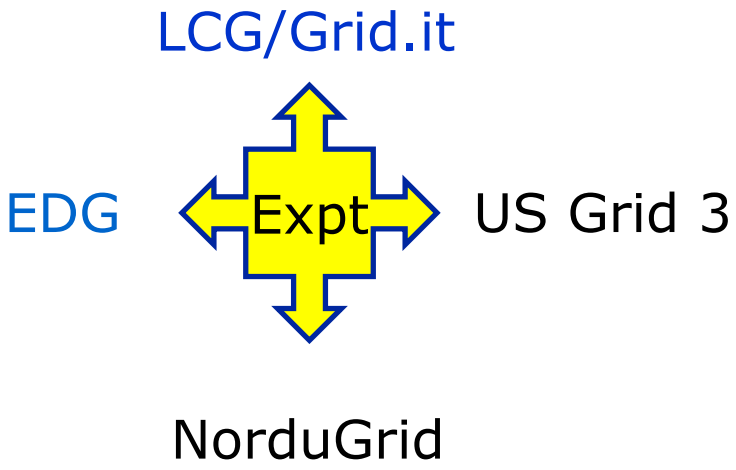
- Overview of objectives and achievements
- Key points in the achievements of the 6 WP8 experiments
- Lessons learned from the three years
- Summary of the exploitation of WP8 work, and of future HEP applications activity in LCG/EGEE
- Concluding comments
- Questions and discussion

Overview of objectives and achievements

OBJECTIVES	***ACHIEVEMENTS***
Continued work in Architectural Task Force (ATF)	◆ Walkthroughs of HEP use cases helped to clarify interfacing problems.
Reactivation of the Application Working Group (AWG)	◆ Extension of HEPCAL use cases covering key areas in Biomedicine and Earth Sciences. ◆ Basis of first proposal for common application work in EGEE
Work with LCG/GAG (Grid Applications group) in further refinement of HEP requirements	◆ HEPCAL-2 requirements document for the use of grid by thousands of individual users. ◆ In addition further refined the original HEPCAL document
Developments of tutorials and documentation for the user community	◆ WP8 has played a substantial role in course design, implementation and delivery

OBJECTIVES	***ACHIEVEMENTS***
Evaluate EDG Application Testbed, and integrate into experiment tests as appropriate.	◆ Further successful evaluation of 1.4.n throughout summer. ◆ Evaluation of 2.0 on the EDG Application Testbed since October
Liase with LCG regarding EDG/LCG integration and the development of the LCG service.	◆ EIPs (Loose Cannons) helped testing of EDG components on the LCG Cert TB prior to LCG-1 start in September. ◆ Performed stress tests on LCG-1.
Continue work with experiments on data challenges throughout the year.	◆All 6 experiments have conducted data challenges of different scales throughout 2003 on EDG App TB or LCG/Grid.it.

Comments on experiment work



Experiments are living in an international multi-grid world using other Grids

- DataTag project is very important for inter-operability (GLUE schema)

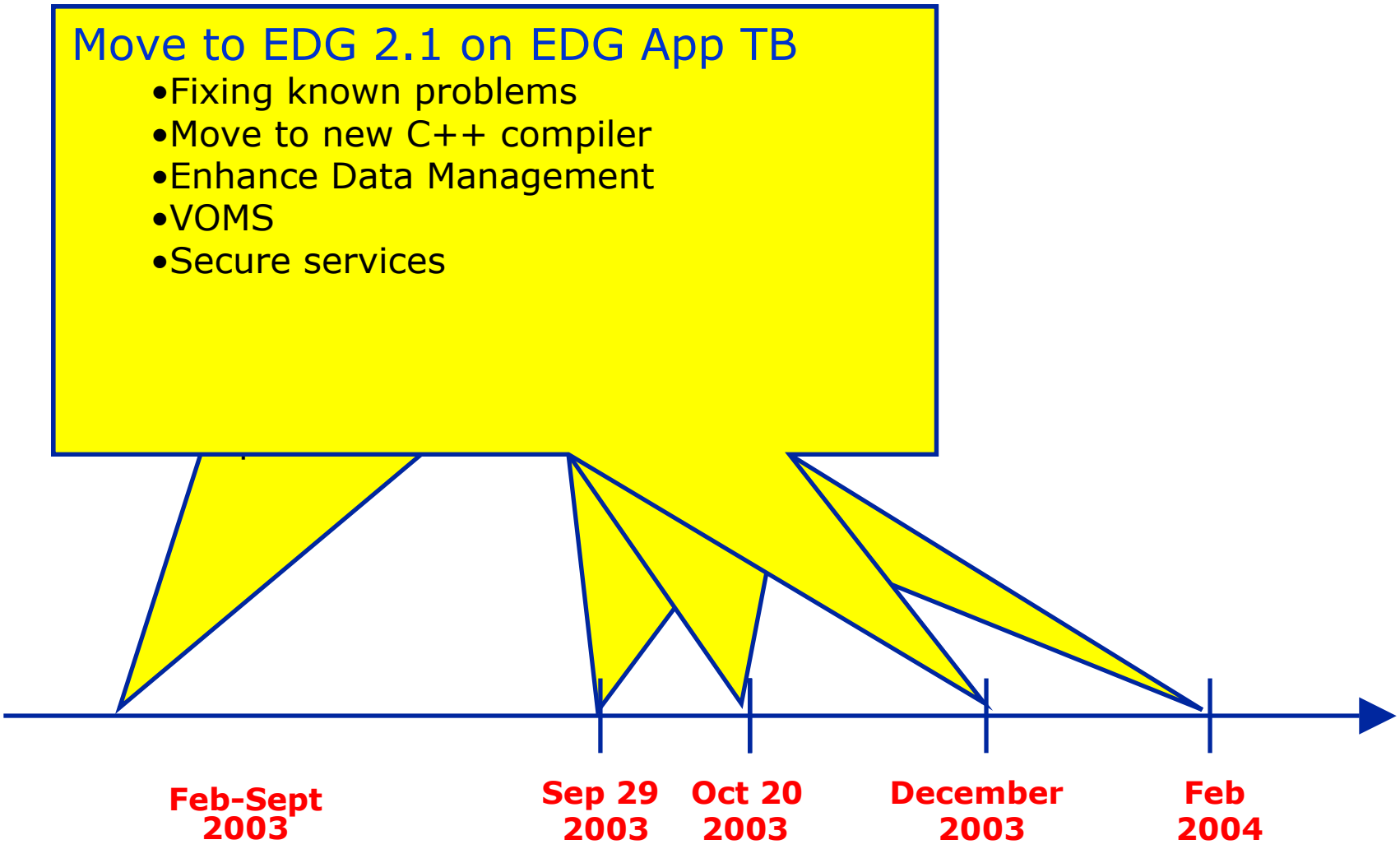
Used EDG software in a number of grids

- EDG Application Testbed
- LCG Service (LCG-1 evolving to LCG-2)
- Italian Grid.it (identical with LCG-1 release)

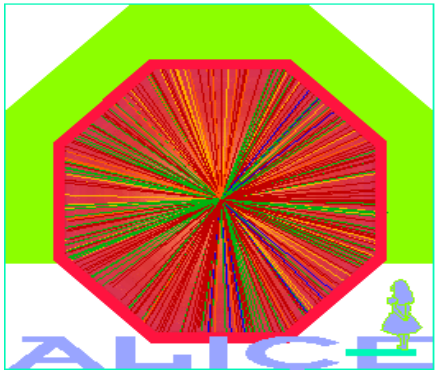
Having 2 running experiments (in addition to the 4 LHC experiments) involved in the evaluations has proved very useful

- BaBar and work on Grid.it
- D0 and work on EDG App TB

Evolution in the use of EDG App TB and the LCG service (and Grid.it)



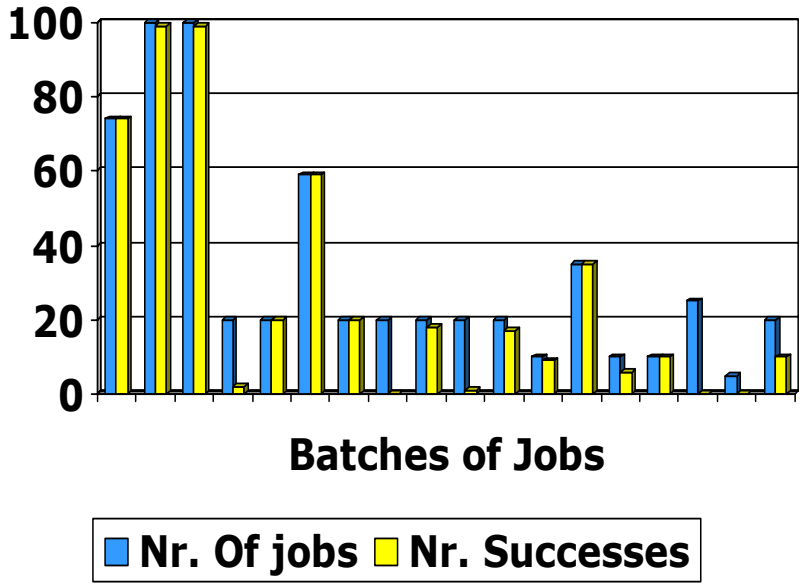
Key points in achievements of the 6 WP8 experiments



- evaluation on LCG-1 and Grid.it
- Significant improvement in terms of stability with respect to tests in Spring 2003
- Jobs were sensitive to space on worker nodes

Projected load on LCG1 during ALICE DC(start Feb 2004) when LCG-2 will be used

- 10^5 events (1 event/job)
- Generate ~30 TB output
- Test LCG Mass Storage
- Parallel data analysis (AliEN/PROOF) including LCG



PLOT of PERFORMANCE with JOB BATCHES IN SEP-NOV

Performance was generally a step function for batches (either close to 0 or close to 100). With long jobs and multi files very sensitive to long-term system stability

ATLAS

- ***Use of EDG 1.4.11 (mod for RH7.3)***

  In May reconstructed 500 K events in 250 jobs with 85% 1st pass efficiency

  With privately managed configuration of 7 sites in Italy, Lyon and Cambridge

- **LCG-1(+ Grid.it) production**

  Have simulated 30000 events in 150 jobs of 200 events each (the jobs required ~20 hours each with efficiency ~80%)

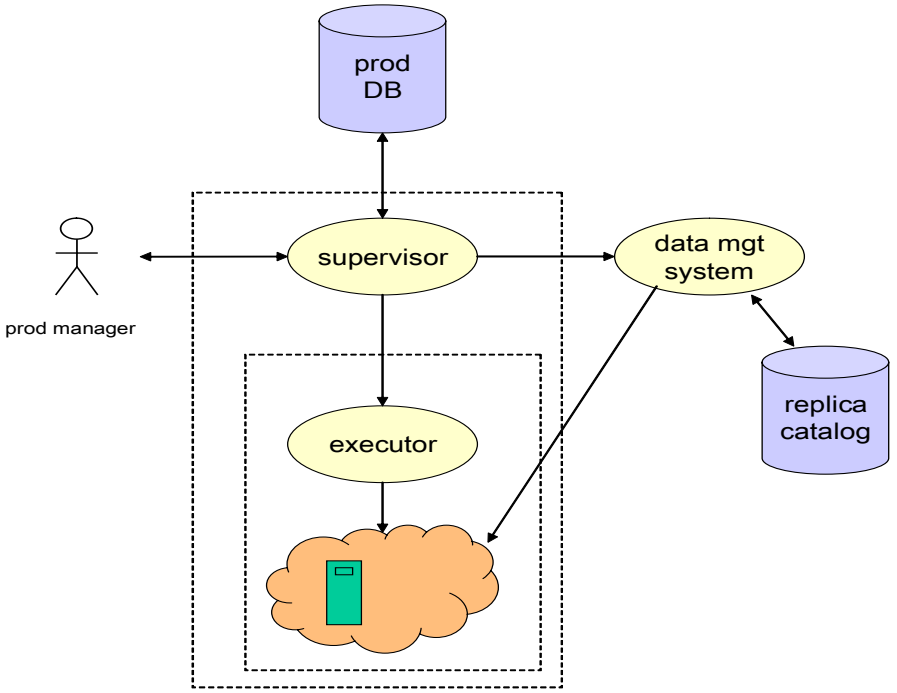
- **LCG-2 plans**

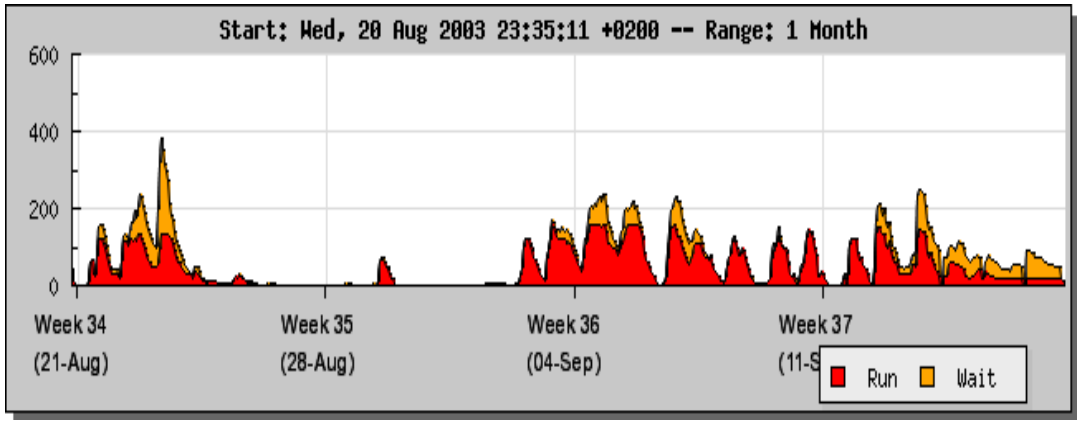
  Start around April

Main features of new DC2 system for multi-grid environment

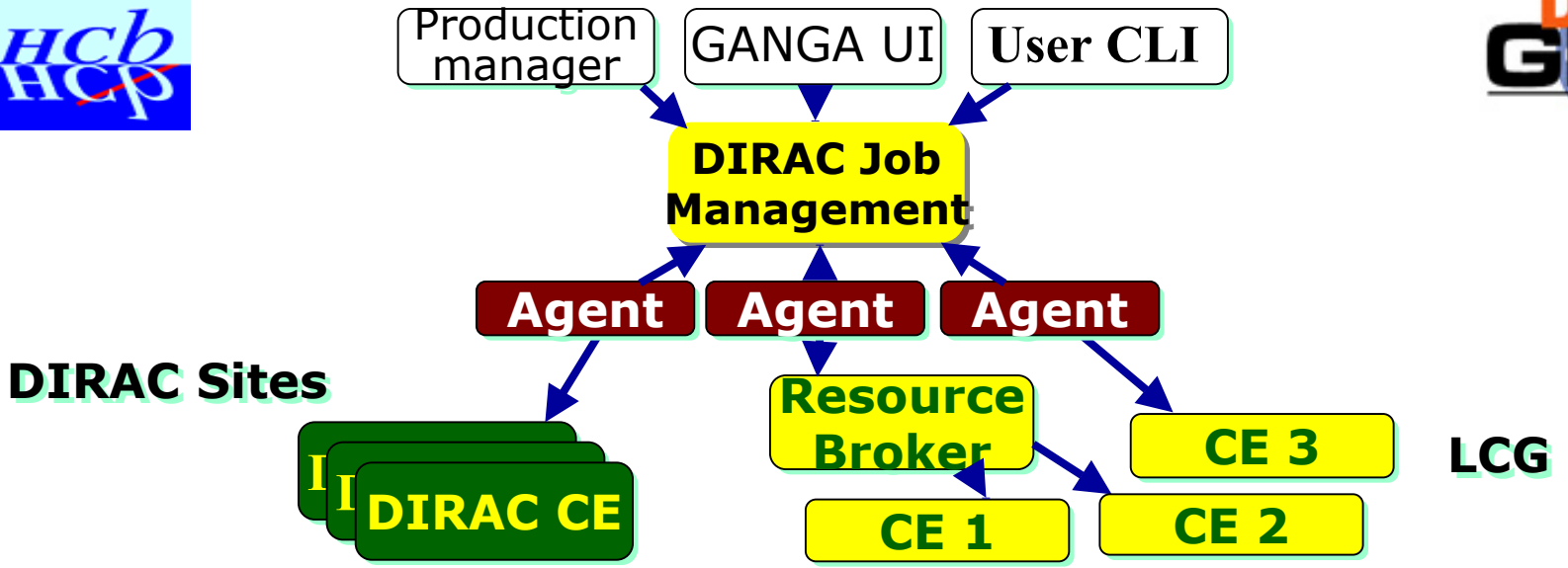
Common production database, supervisor and data management system for all of ATLAS

Executors developed by middleware experts (LCG, NorduGrid, US).





- LCG-0 (summer 2003)
  - Components from VDT 1.1.6 and EDG 1.4.11
  - DataTAG (GLUE)
  - VOMS + RLS + R-GMA
  - *14 sites configured and managed by CMS*
  - *Substantial improvements in efficiency compared to first EDG stress test (~80%)*
  - **76000 CPU hours on LCG-0**
    - *500K Pythia 2000 jobs 8h*
    - *1.5M CMSIM 6000 jobs 10h*
- LCG-1
  - Ran for 9 days on LCG-1 over Xmas
  - In total 600,000 events (30-40h jobs) were produced
  - Sites used mainly in Italy, Spain
  - *Efficiency around 75% over XMAS*
  - Used GENIUS portal
- LCG-2 -data challenge Mar 1



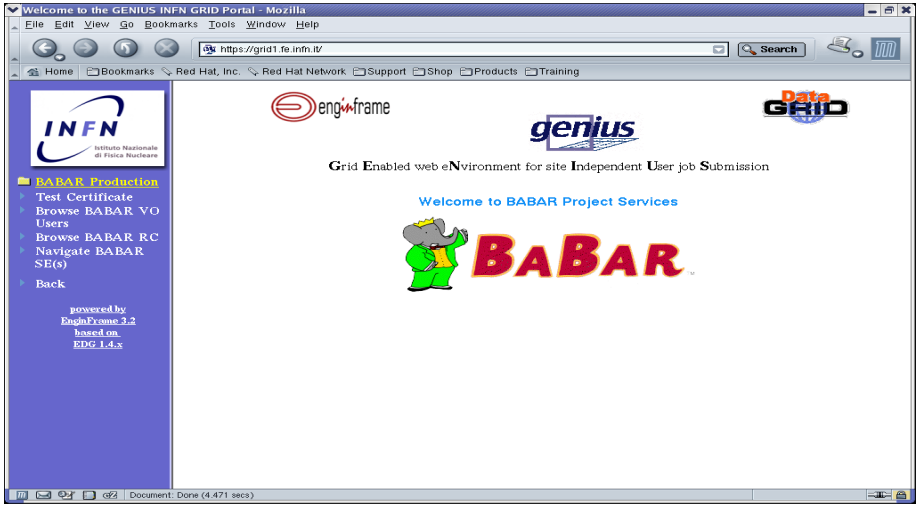
- **Tests on the EDG1.4 application testbed (Feb-Mar 2003):**
  - Standard LHCb production tasks, 300K events produced;
  - ~35% success rate. (TB support running down)
  - Software installation by the running job;

- **EDG2.0 tests (November 2003):**

Submission of the jobs:

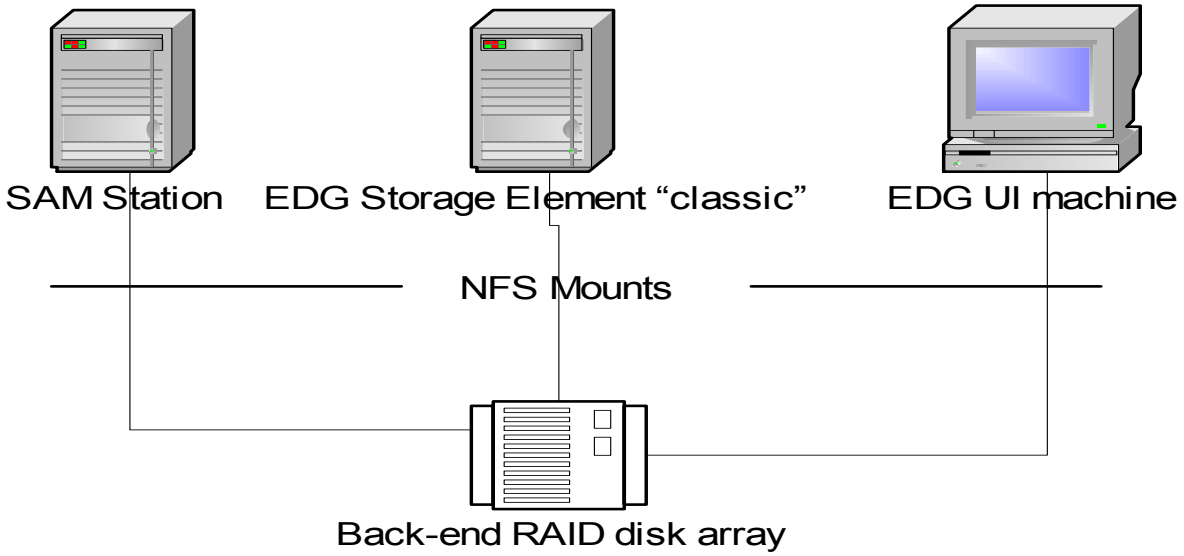
- To EDG RB;
- *Directly to a CE with the CE status information obtained from the CE GRIS server: 90% efficiency*

GETTING READY NOW FOR LCG-2 and DC in April (all tests are positive)



- **Strategy for first integration**
  - Created 'simulation' RPM to be installed at sites
  - Data output stored on closest SE
  - Data copied to Tier-1 or SLAC using edg-copy
  - Logfiles via sandbox
- **Scheme first tested with EDG 1.4.11 on 5 Italian sites**

- **Operation on Grid.it with LCG-1 release**
  - RB at CNAF - farms at 8 sites
  - 1 week test with ~ 500 jobs
  - *95% success at Ferrara(site with central DB)*
  - *60% success elsewhere*
    - *33% failures due to network saturation due to simultaneous requests to remote applications database (looking at distributed solutions)*
- **Positive experience with use of GENIUS portal**
  - https://genius.ct.infn.it
- **Analysis applications also have been successfully tested on EDG App TB**



- Interfaced EDG software and resources to D0 re-processing
  - *Frequent software updates so don't use RPMs*
    - *Registered compressed tar archives in RLS as grid files for installation by jobs*
  - *Use RGMA for monitoring*
    - *Allows users and programs to publish information for inspection by other users, and for archiving in production database*
- Found EDG s/w generally satisfactory for task (with caveats)
  - Used 'Classic' SE s/w while waiting for mass storage developments
  - Very sensitive to RGMA instability. Recently good progress with RGMA, and can run at ~90% efficiency when RGMA is up

Summary of middleware evaluations

- Workload management
  - Tests have shown that software is more robust and scalable
    - Stress tests were successful with up to 1600 jobs in multiple streams – efficiencies over 90%
    - Problems with new sites during tests– VOs not set up properly (though site accepted job)
- Data Management
  - Has worked well with respect to functionality and scalability (have registered ~100K files in ongoing tests)
    - Tests so far with only 1 LRC per VO implemented
  - Performance needs enhancement
    - Registrations and simple queries can take up to 10 seconds
  - We have lost (with GDMP) bulk transfer functions
  - Some functions needed inbound IP connectivity (Globus). D0 had to program round this (problem since fixed)

Summary of middleware evaluations(2)

- Information System
  - Partitioned MDS has worked well for LCG following on from work accomplished within EDG (BD II work), but limited to ~100 sites probably.
  - R-GMA work is very promising for 'life after MDS', but needs 'hardening'.
- Mass Storage support (mission critical for data challenges)
  - We await 'accepted' uniform interface to disk and tape systems
    - Solution coming with SRM/GFAL software
    - WP5 have made important contribution to the development of SRM interface
      - EDG 2.0 had mass storage access to CERN (Castor) and RAL(ADS)
    - The 'Classic-SE' has been a useful fallback (gridftp server) while waiting for commissioning of developments

Site Related Issues (major factors in overall efficiency)

- Site Certification
  - Official, standard procedure as part of release
  - Consistency checks of published information
- Site Configuration
  - Large parameter space with insufficient defaults so please can we have…
    - Automated configuration
    - Automated tests
    - Run-time checks of parameters
- Space management and publishing
  - Running out of space on SEs and WNs is still a problem. Jobs need to check availability before running

The Deliverables + 'extra' outputs from WP8

- The formal EU deliverables
  - D8.1 The original HEP requirements document
  - D8.2 'Evaluation by experiments after 1st year'
  - D8.3 'Evaluation by experiments after 2nd year'
  - D8.4 'Evaluation after 3rd year'
- Extra key documents (being used as input to EGEE)
  - HEPCAL Use cases May 2002 (revised Oct 2003)
  - AWG Recommendations for middleware (June 2003)
  - AWG Enhanced use cases (for Biomed,ESA) Sep 2003
  - HEPCAL2 Use cases for analysis (several WP8 people)
- Generic HEP test suite used by EDG/LCG
- Ongoing consultancy from 'loose cannons' to all applications
- *Interfacing of 6 experiment systems to middleware*

Main lessons learned

Architecture & Software Life-cycle

- Information system is nerve centre of grid. We look to R-GMA developments for long term solution to scaling problems
- Globally HEP applications feel it would have been 'better' to start with simpler prototype, and to have more frequent incremental releases
- Applications should have played larger role in architecture in defining interfaces (so we could all learn together!) .

Deployment & Operations of the Middleware

- Formation of Task Forces (applications+middleware) was a very important step midway in project
- Loose Cannons (team of 5) were crucial to all developments. Worked across experiments. This team comprised all the funded effort of WP8.

Main lessons learned (cont'd)

Site Related Lessons

- Site configuration must be automated.
- Site certification needs to be improved. Incompliant sites screw up the brokering.
- Space management on SEs and WNs is a still outstanding problem
- We look to SRM/GFAL as solution to uniform mass storage interfacing
- Must have flexible application s/w installation. Application needs and site policies vary.

Exploitation of the work of WP8, and future HEP applications work in LCG/EGEE

- All experiments have exploited the EDG middleware using WP8 effort, and this exploitation is being carried into the data challenges in 2004
- The HEPCAL and AWG documents are essential inputs to the future LCG/EGEE work
- Future developments will be in the context of LCG/EGEE infrastructure carrying over the experience from WP8
- The NA4 activity in EGEE will include dedicated people for interfacing middleware to experiments (8 people at CERN + others distributed in the community)
- Within the EGEE project middleware will be 'hardened' (including EDG components) and evaluated by the HEP applications

Concluding comments

- Over the past 3 years the HEP community has moved to the use of grid services in physics production systems using world-wide configurations
- Experiments are using several managed grids (LCG/EGEE,US Grids, Nordugrid) so inter-operability is crucial
- Existing software can be used in production services, with parallel 'hardening' of middleware taking advantage of lessons learned (the ARDA project in LCG/EGEE)
- THANKS
  - Personal thanks to the members of WP8 for all their efforts
  - And on behalf of WP8 to all the other WPs (middleware,testbed,networking,project office) for their full support and cooperation
  - To the EU and all our national funding agencies

Questions and discussion