

The experience of the 4 LHC experiments with LCG-1

F Harris (OXFORD/CERN)

Structure of talk (and sources of input)

- For each LHC experiment
 - Preparatory work accomplished prior to use of LCG-1
 - Description of tests (successes, problems, major issues)
 - Comments on user documentation and support
 - Brief statement of immediate future work and its relation to other work(e.g. DCs) and other grids + comments on manpower
- Summary
- *Inputs for this talk*
 - *4 experiment talks from internal review on Nov 17*
<http://agenda.cern.ch/fullAgenda.php?ida=a035728#s2>
 - *Extra information obtained since by mail and discussion*
 - *Overview talk on 'grid production by LHC experiments' of Nov 18*
(link as above)

ALICE and LCG-1

- ALICE users will access EDG/LCG Grid services via AliEn.
 - The interface with LCG-1 is completed; first tests have just started.
- Preparatory work commenced in August on LCG Certification TB to check working of Alice software in LCG environment.
 - Results of tests in early September on LCG Cert TB(simulation and reconstruction)
 - Aliroot 3.09.06 fully recontructed events
 - CPU-intensive, RAM-demanding (up to 600MB ,160MB average) ,long lasting jobs (average 14 hours)
 - Outcome:
 - > 95 % successful job submission, execution and output retrieval in a lightly loaded GRID environment
 - ~ 95 % success (first estimate) in a highly job-populated testbed with concurrent job submission and execution (2 streams of 50 AliRoot jobs and concurrent 5 streams of 200 middle-size jobs)
 - MyProxy renewal succesfully exploited

ALICE –details of latest LCG-1 test

- 200 Pb-Pb events
- 1 job/event -> 200 jobs
- 1.8 GB/job -> 360 GB
- 12-24 hours per job
-
- Started on November 14/11/2003
- 17/11 11:00 : 137 done; 31 cancelled; 32 Waiting
-> 82.2%

ALICE - Comments on first tests and use of LCG-1 environment

- Results: monitoring of efficiency and stability versus job duration and load
 - Efficiency (algorithm completion): if the system is stable eff ~90% , if any instability eff=0%. (looks like a step function!)
 - Efficiency(output registration to RC): 100%
 - Automatic Proxy-renewal: always OK
 - Comments on geographical job distribution by Broker:
 - A few sites accept event until they saturate and then RB looks for other sites
 - When submitting a bunch of jobs and no WN is available, all the jobs enter the Schedule state always on the same CE.
 - Disk space availability on WN has been a source of problems
 - .
- User documentation and support of good quality
 - But need more people
- Mass Storage support missing now
 - This is essential in LCG-2

ALICE –comments on past and future work

- EDG1.4(March) versus LCG1
 - Improvement in terms of stability
 - Efficiency 35% -> 82% (preliminary)...of course we want 90+% to be competitive with what we have with traditional batch production
- Projected load on LCG1 during ALICE DC(start Jan 2004) when LCG-2 will be used
 - 10^4 events
 - Submit 1 job/3' (20 jobs/h; 480 jobs/day)
 - Run 240 jobs in parallel
 - Generate 1 TB output/day
 - Test LCG MS
 - Parallel data analysis (AliEN/PROOF) including LCG

Atlas LCG-1 developments

- ATLAS-LCG task force was set up in September 2003
- October 13: allocated time slots on the LCG-1 Certification Testbed
 - Goal: validate ATLAS software functionality in the LCG environment and vice versa
 - 3 users authorized for the period of 1 week
 - Limitations: little disk space, slowish processors, short time slots (4 hours a day)
- ATLAS software (v6.0.4) deployed and validated
 - 10 smallest reconstruction input files replicated from CASTOR to the 5 SEs using the **edg-rm** tool
 - The tool is not suited for CASTOR timeouts
 - Standard reconstruction scripts modified to suit LCG
 - Script wrapping by users is unavoidable when managing input and output data (EDG middleware limitation)
 - Brokering tests of up to 40 jobs showed that the workload gets distributed correctly
 - Still, time was not enough to complete a single real production job

Atlas LCG-1 testing phase-2 (late Oct-early Nov)

- The LCG-1 Production Service became available for every registered user
 - A list of deployed User Interfaces was never advertised (though possible to dig out on the Web)
 - Inherited old ATLAS software release (v3.2.1) together with the EDG's LCFG installation system
- Simulation tests at LCG-1 were possible
 - A single simulation input file replicated across the service
 - 1/3 of replication attempts failed due to wrong remote site credentials
 - A full simulation of 25 events submitted to the available sites
 - 2 attempts failed due to remote site misconfiguration
 - This test is expected to be a part of the LCG test suite
 - At the moment, LCG sites do not undergo routine validation
- New ATLAS s/w could not be installed promptly because it is not released as RPM
 - Interactions with LCG : define experiment s/w installation mechanisms
 - Status of common s/w is unclear (ROOT, POOL, GEANT4 etc)

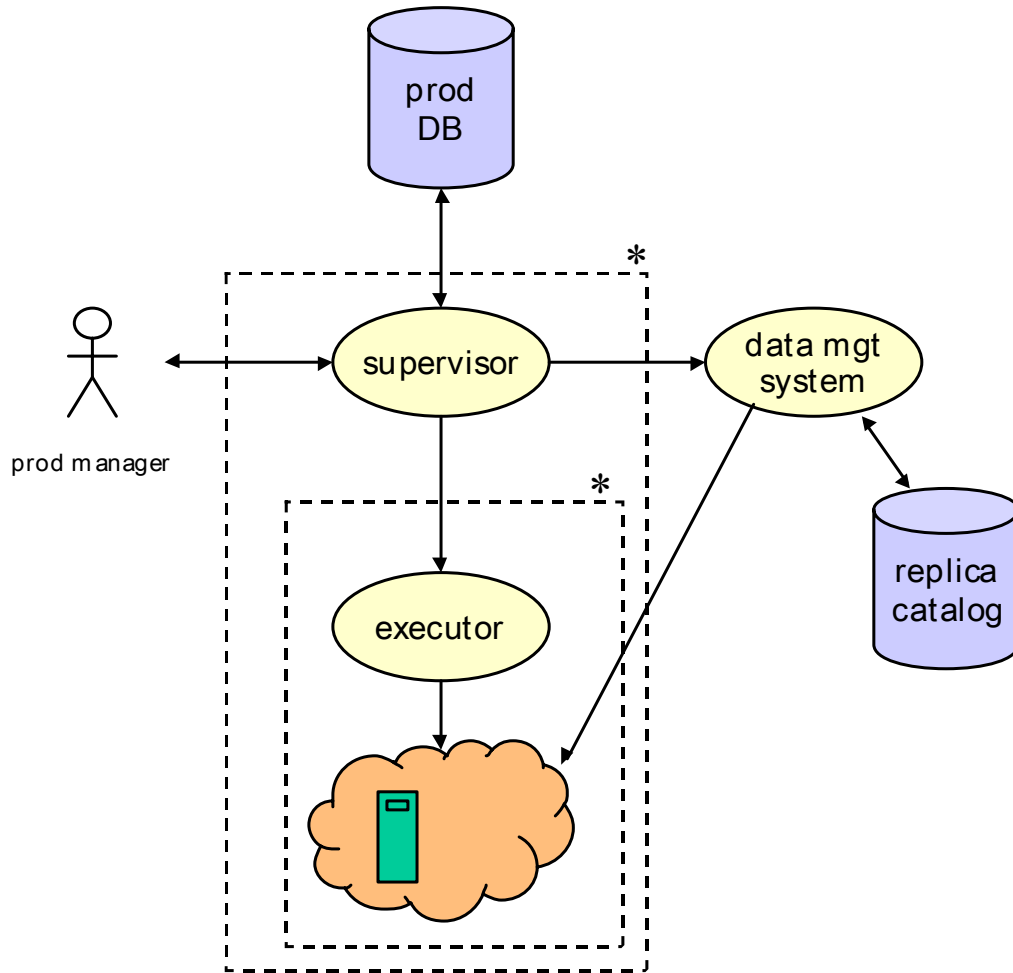
Atlas LCG-1 testing phase 3(Nov 10 to now...)

- By November 10, a newer (not *newest*) ATLAS s/w release (v6.0.4) was deployed at LCG-1 from tailored RPMs
 - PACMAN-mediated (non-RPM) software deployment is still in the testing state
 - Not all the sites authorize ATLAS users
 - 14 sites advertise ATLAS-6.0.4
 - Reconstruction tests are possible
- ATLAS s/w installation validated by a single-site simulation test
- File replication from CASTOR test repeated
 - 4 sites failed the test due to site misconfiguration
- Tests are ongoing

Atlas overview comments

- Site configuration
 - Sites are often mis-configured
 - Need a clear picture of VO mappings to sites
- Mass storage support is ESSENTIAL
- Application s/w deployment
 - System-wide experiment s/w deployment is a BIG issue, especially when it comes to 3d party s/w (e.g., that developed by the LCG's own Applications Area)
- The deployed middleware, as of **today**, does not provide the level of efficiency provided by existing production systems
 - Some services are not fully developed (data management system, VOMS), others are crash-prone (WMS, Infosystem – from EDG)
 - User interfaces are not user-friendly (wrapper scripts are unavoidable, non-intuitive naming and behavior) – **very steep** learning curve
- Manpower is a problem
 - Multi counting the same people for several functions (DCs +LCG testing + EDG evaluation+..)
- LCG are clearly committed to resource expansion, middleware stabilization and user satisfaction
 - ATLAS is confident it will provide reliable services by DC2
 - EDG-based m/w has improved dramatically, but still imposes limitations
 - There is good quality documentation and support - but need more people + tutorials and improved flow of information

Schematic of New ATLAS DC2 System - integrating use of LCG, Nordugrid and US production



– Main features

- Common production database for all of ATLAS
- Common ATLAS supervisor run by all facilities/managers
- Common data management system a la Magda
- Executors developed by middleware experts (LCG, NorduGrid, Chimera teams) -? Can Chimera drive US and LCG
- Final verification of data done by supervisor

preparatory work by CMS with 'LCG-0' – started in May

- CMS/LCG-0 is a CMS-wide testbed based on the LCG pilot distribution (LCG-0), owned by CMS (joint CMS/LCG/DataTag effort)
 - Red Hat 7.3
 - Components from VDT 1.1.6 and EDG 1.4.X
 - GLUE schemas and info providers (DataTAG)
 - VOMS
 - RLS
 - Monitoring: GridICE by DataTAG
 - R-GMA (as BOSS transport layer for specific tests)
- Currently configured as a CMS RC and producing data for PCP
- 14 sites configured
- Physics data produced
 - **500K Pythia** **2000 jobs** **8 hr**
 - **1.5M CMSIM** **6000 jobs** **10 hr.**
- **Comments on performance**
 - **Had substantial improvements in efficiency compared to first EDG stress test**
 - **Networking and site configuration were problems, as was 1st version of RLS**

CMS use of RLS and POOL

RLS used in place of the Replica Catalogue

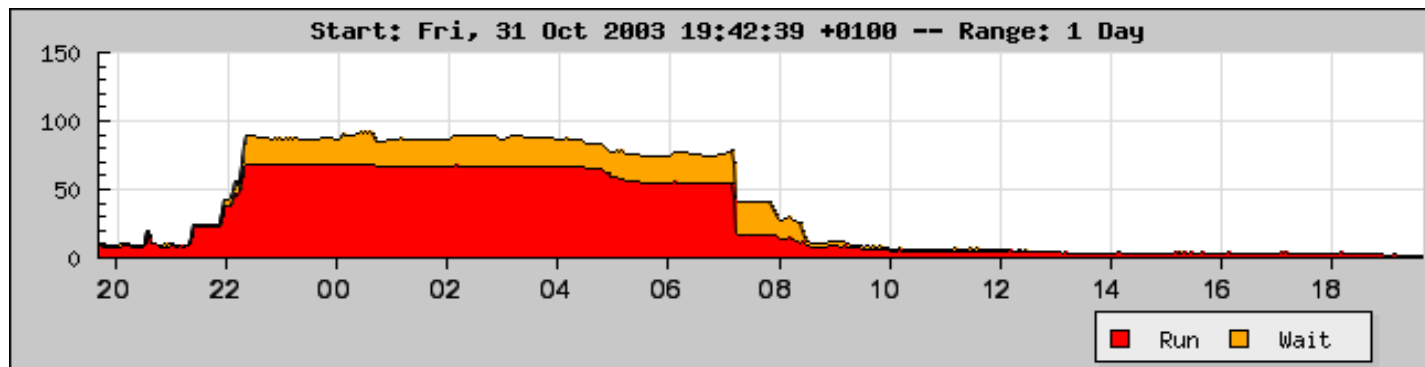
- Thanks to IT for the support

POOL based applications

- CMS framework (COBRA) uses POOL
- Tests of COBRA jobs started on CMS/LCG-0. Will move to LCG-1(2)
- Using SCRAM to re-create run-time environment on Worker Nodes
- Interaction with POOL catalogue. Two steps:
 - COBRA uses XML catalogues
 - OCTOPUS (job wrapper) handles XML catalogue and interacts with RLS
- definition of metadata to be stored in POOL catalogue in progress

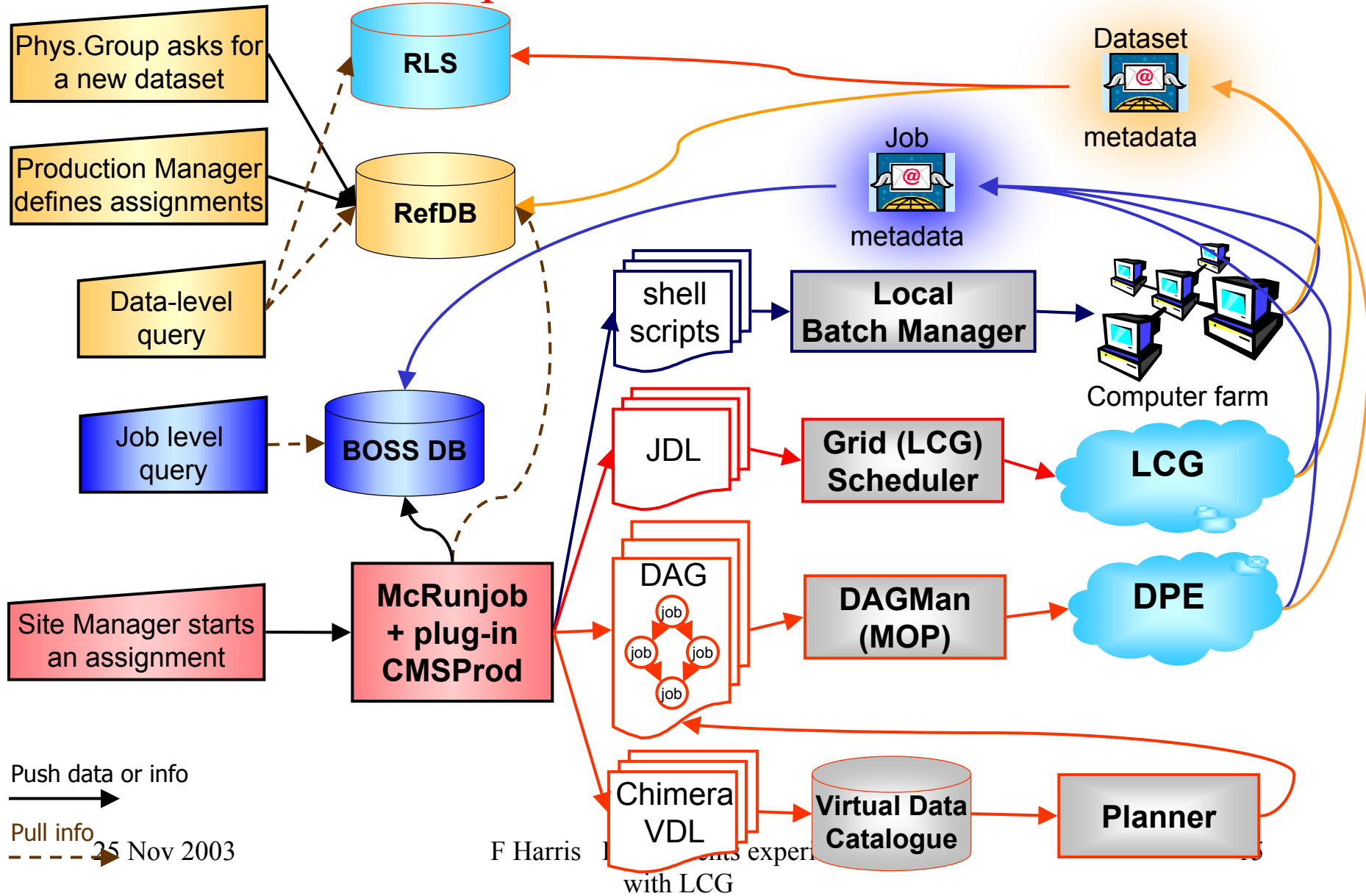
CMS Tests on LCG-1

- Porting of CMS s/w production software to LCG-1
 - on Italian (Grid.it) testbed and on LCG Certification & Testing testbed
 - improved interface to user simplifies job preparation
- Testing on official LCG-1 testbed
 - CMS software deployed everywhere on oct 28th 2003
 - CMKIN (few min's) & CMSIM (7 hours) submitted in bunches of ~50 jobs
 - Failure rate is 10-20% for short jobs and ~50% for long jobs
 - Mainly due to sites not correctly configured
 - excluded in the JDL (until ClassAd size exceeded maximum limit!)



- Will move all activities on LCG-1(2) official system as soon as CMS software to be deployed grid-wide will be more stable
 - Stress test before the end of the year

CMS OCTOPUS Production System – integrating all production modes

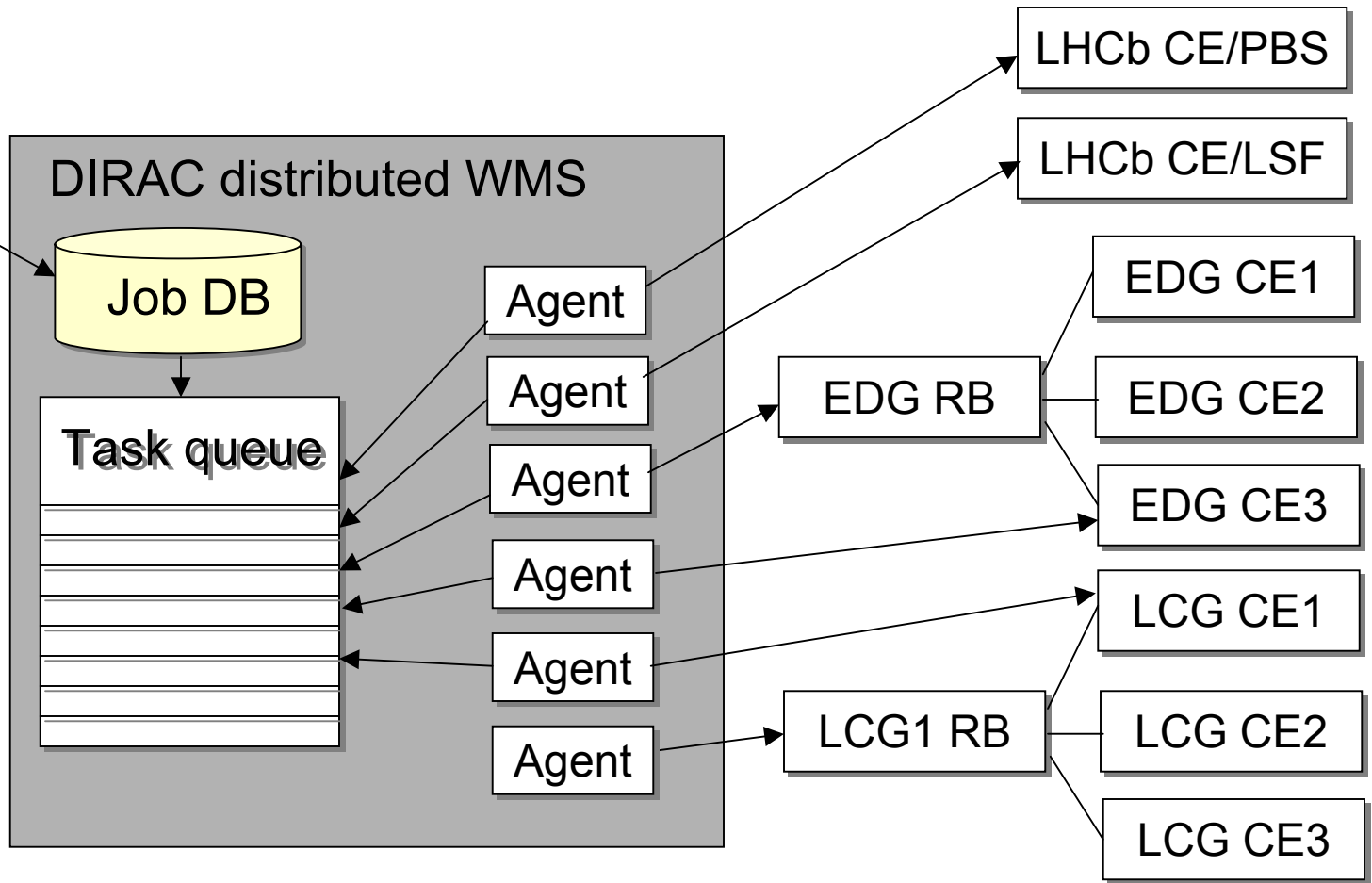


CMS Overview comments

- Good experience with CMS/LCG-0
 - LCG-1 components used in CMS/LCG-0 are working well
 - Close to production-quality
- First tests with LCG-1 promising
 - main reason of failure are mis-configured sites
- POOL/RLS tests under-way
 - CMS reconstruction framework (COBRA) is “naturally” interfaced to LCG grid catalogs
- Large scale tests still to be done on LCG-1(2)
 - LCG-2 preferred because it will likely have VOMS, SRM, GFAL
- Thanks to LCG for very good documentation and support
 - With more people now need more support

LHCb DIRAC WMS architecture

Production manager



LHCb LCG tests commenced mid October (following short period on Cert TB)

- New software packaging in rpms ;
 - Testing the new LCG proposed software installation tools;
- New generation software to run:
 - Gauss/Geant4+Boole+Brunel+...
- Using the LCG Resource Broker
 - Direct scheduling if necessary.

LHCb LCG tests (2)

- Tests of the basic functionality
 - LHCb software correctly installed from rpms ;
 - Tests with standard LHCb production jobs:
 - 4 steps – 3 simulation datasets, 1 reconstructed dataset;
 - Low statistics – 2 events per step;
 - Applications run OK ;
 - Produced datasets are properly uploaded to a SE and registered in the LCG catalog;
 - Produced datasets are properly found and retrieved for the subsequent use.

LHCb LCG tests – next steps

- Long jobs:
 - 500 events ;
 - 24-48 hours depending on CPU ;
- Large number of jobs to test the scalability:
 - Limited only by the resources available.
- LCG-2 should bring important improvements for LHCb which we will try as soon as they will be available:
 - Experiment driven software installation;
 - Testing now on the “installation” testbed.
 - Access to MSS (at least Castor/CERN)

LHCb LCG tests – next steps continued

- LCG-2 seen as an integral part of the LHCb production system for the DC 2004 (Feb 2004)
- Necessary conditions :
 - The availability of major non LHC dedicated centres both through usual and LCG workload management system;
 - E.g CC/IN2P3, Lyon.
 - The LCG Data Management tools accessing to major MSS (Castor/CERN, HPSS/IN2P3, FZK, CNAF, RAL);
 - The overall stability and efficiency (>90%) of the system providing basic functionality – develop incrementally but preserve the 90% please!
- Manpower is a problem
 - Same people running DCs, interfacing to LCG/EDG and doing software development – this is natural but there is a shortage of people
- Happy with quality of LCG support and documentation

Summary 1

- . Experiments have had access to LCG Cert TB from August, and to LCG-1 from early October (later than planned due to late delivery of EDG 2.0 software), so these are early days for the LCG service
- **Feedback from experiments on experiences so far**
 - Documentation and support
 - good quality – need more people now
 - Stability of service
 - has had good and bad days in start-up
 - ALICE and CMS have had some positive running on LCG-1
 - Experiments have appreciated careful approach of LCG in certifying releases
 - Site management, configuration and certification tools are essential. This area remains a major source of errors
 - Error detection, reporting and recovery are still very basic or non-existent (though applications have done good work e.g. GRAT,BOSS,CHIMERA...)
 - Application Software installation at sites is an issue (being worked on)
 - Support of mass storage devices is absolutely essential
 - Scalability of middleware as configurations and N users grow is a ?

Summary 2

- We all look to LCG-2 to improve the situation (mass storage,VOMS,gcc 3.2.2 release...)
- Experiments live in a run in a multi-grid world and must maintain their existing data processing systems
 - As well as LCG we have US grids, Nordugrid, Alien, Dirac.....
 - Manpower is a big issue to keep all this going
 - What is going to be influence of ARDA in ‘improving’ all this?
- Experiments start with LCG-2 for data challenges (ALICE in Jan)
- These are very early days – community is learning to live with GRIDs!
- **Thanks to experiments for full cooperation in providing information**
 - **ALICE** R Barbera(Catania), P Buncic(CERN), P Cerello(Turin)
 - **ATLAS** K De(Univ of Texas), R Gardner(Argonne), G Poulard(CERN), O.Smirnova(Lund)
 - **CMS** C.Grandi(Bol), G.Graham(FNAL), D.Bradley(Wisc), A.Fanfani(Bol)
 - **LHCb** N Brook(Bristol), J Closier(CERN), A Tsaregorodtsev(Marseille)