Testbed 2.1 Priorities

J. Templon, NIKHEF/WP8

Caveats

- EDG release 2.0 needs to work
- hard to discuss 2.1 priorities are until 2.0 is evaluated; it's still not really clear *exactly* what 2.0 will be
- input is preliminary (expect better around June)
- SE 'file' protocol needs to work!

Use Case Mapping

- Applied all of HEPCAL use cases to EDG 1.4.x testbed
- Evaluated in terms of "implemented", "mostly implemented", "partially implemented", or "not implemented"
- 26 of 43 use cases were at least partially implemented
- 17 were not implemented
  - Metadata use cases
  - Virtual Data use cases
  - Enhanced VO support (such as resource reservation)
- For more info, see chapter 5 of D8.3

Scales involved for TB2.x

recent results (LHCb) for distributed MC production with DIRAC system.

- 50 days running from 13 Feb 2003-04-15
- 17 centres in production configuration
- 36,600 jobs ran
  - 34,000 ran successfully 92% success
  - 2% LHCb s/w problems
  - 6% mixture data transfer + site + farm problems
- ~250,000 files produced
  - 100 MB /file
  - ~20 TB data
- Total CPU used 1.4 M hours 1GHz PC (~1000 CPUs for 50 days)

High Priority Items (I)

- **Robustness**
  - R3 Need less than 2% job failures for internal middleware reasons (not including network, quota etc.. problems)
- **Reliability and Scalability**
  - RS1 System must handle tens of thousands of concurrent jobs, and be able to cope with job submission rates of one/few secs
  - RS2 Must be able to register millions of files
  - RS3 Must be able to handle multi-GB file transfers with global reliability of ~99%
- **Security**
  - S4 Need a comprehensive grid-security implementation (multiple user groups with different levels of privilege)
  - S6 Security procedures should permit outbound IP connectivity from WNs

High Priority Items (II)

- **Accessibility**
  - A1 Need programmable APIs(C++ and Java) to middleware services
  - A2 Need 'recent' system support (Redhat 7.3,Redhat 8.0)
- **Data Management**
  - DM1 Need grid access to to files from running application (gridopen etc..)
  - DM2 Need support for application metadata(as user defined fields associated to the LFN)
  - DM5 Must have a system for SE,WN space management
- **Job Handling**
  - JB1 Scheduling of jobs should take into account job input/output files location and trigger replication if needed.
  - JB7 WMS handle job sequences (DAGs) … CMS request

High Priority Items (III)

- **Installing Application Software**
  - Currently being reviewed by Atlas (watch this space!)
  - Cal says "we need to see the requirements"

There are other issues that must be addressed by the project which are of a procedural nature. These are high priority for developing a production environment.

- **Site Certification**
  - We need procedures for certifying new sites
- **System Monitoring and Control**
  - We need a system for monitoring grid services and taking appropriate actions, e.g. taking sites in and out of configuration according to watchdog statistics over a period of time