Status of BaBar GRID and future plans

David Smith
University of Birmingham, UK.

16 September, 2002
GridPP Collaboration Meeting at Imperial College

- BaBar overview
- Existing infrastructure
- The development of a BaBar GRID
- Future plans

BaBar in brief

- BaBar is an HEP experiment which has been collecting data since May 1999, at SLAC.
- Primary storage of both data and MC is in federated Objectivity databases.
- Collaborators based at many sites around the world.
- Database now in excess of 500Tb, significant portions of which are replicated at sites other than SLAC.

Data distribution

- Data distributed outside SLAC in either root or Objectivity DB format.
- RAL is the main repository for the root format data, and many smaller sites copy from there.
- For the Objectivity distribution many dedicated tools have been written to handle the replication of databases and logical groups of databases. However these BaBar tools do not constitute a replica management system.

The skim metadata catalogue

- Each site hosts an SQL based metadata catalogue that contains entries for all of the data files produced. This is maintained at each site by mirroring databases at SLAC.
- The SQL database at each site also contains information regarding the availability of data at that location. Therefore the metadata catalogue also acts as a local replica catalogue.
- Data replication itself is initiated locally at each site with an explicit selection of data from the metadata catalogue.

Current Grid activity

Currently work is being done on:

- Distribution of any general data analysis task, via the Grid, to multiple execution sites. Each site should receive the task as multiple jobs, of reasonable size each.
- Submission of MC requests to production sites via Grid
- Security issues are also being considered.

Need a coherent replica catalogue:

- Need to construct a full replica catalogue, to resolve a logical $collection$ name to many $physical$ names.

Building a replica catalogue for job submission I

Two approaches, at the moment:

- An extension of the existing metadata catalogue at each site turns it into a replica catalogue.
- To build this replica catalogue every site is polled for availability of data and the result stored locally. This is repeated at intervals.
- Replica information can, of course, become slightly out of date.
- There is no replication facility with this approach.

...and...

Building a replica catalogue for job submission II

- Other approach is to use the WP2 Replica Location Service (RLS). The RLS is already being investigated for use in dataset replication using the Storage Resource Broker (SRB).
- The SRB is in use by BaBar at SLAC and in testing at CCIN2P3 (Lyon).

Building a replica catalogue for job submission III

- The intention is to setup SRB metadata catalogues, MCAT, using the RLS. This work is being done as one of the project activities within the Particle Physics Data Grid collaboration (PPDG).
- Tests using Giggle in the UK to construct the distributed RLS are foreseen to start soon.

Grid Analysis and data splitting

- An analysis job might want to access tens of thousands of data files (or objectivity $collections$), perhaps containing hundreds of millions of events.
- Overlapping subsections of the data may be available at tens of sites.
- Prior to submission to the grid the list of files is broken into jobs of a given number of events, and the jobs into groups. Each group contain data which are available at a unique combinations of sites.

Grid Analysis job submission

- Job submission being developed using European DataGrid (EDG) software.
- Allow Resource Broker (RB) to select between possible execution sites, given as resource specifiers.
- Actual job creation is involved due to the large number of dependent tcl files in the BaBar framework. An unrolling procedure will be used to make the jobs more self contained.
- Initially foresee to pass inputs and outputs via the sandboxes. In future, will need to decide under what conditions to register large files in the EDG Replica Catalogue (RC) and replicate the required files to a local Storage Element (SE).

Status of EDG installation and testing I

- A BaBar Virtual Organisation (VO) has been setup and is hosted in the UK. System to automatically update membership list from SLAC in place.
- EDG Compute Element (CE) and Worker Node (WN) installations in many BaBar sites. The UK BaBar dedicated farms are upgrading from EDG 1.0.x to 1.2.x now.

Status of EDG installation and testing II

- Extensive tests done at SLAC, first with 1.1.4 and now with 1.2.0.
- Problems with EDG at SLAC at first, from location of Globus Access to Secondary Storage (GASS) cache and WN firewall issues. Overcome with code patches, to be fed back into EDG.
- Much better performance now seen with 1.2.0. Stress tests showing about a 5% job failure rate.
- No SE being used yet.

BaBar GRID demonstrators

Two systems have been made to demonstrate the intended Grid functionality:

- One based on Globus only and demonstrates the possibility to split a BaBar job in several pieces, each one running on a different dataset located several sites. The dialog between the user and the grid application is done through a web interface.
- The second one is a standard BaBar analysis application run through EDG. The emphasis is on the remote job submission aspect.

BaBar GRID future plans

BaBar has a set list of milestones for GridPP. In particular:

- Prototype BaBar analysis job submission. Includes job file splitting, job submission and output retrieval via the RB. (December 2002)
- Production of a Replica Catalogue, also to be used by the SRB system. (March 2003)
- BaBar analysis job submission using the RLS. (June 2003)