Job management and bookkeeping in OSG productions

OSG Users meeting Fermilab 2007
Nikolay Kuropatkin

GGG (Generic Grid Gofer) - the tool to organize Grid productions.

Production – hundreds, thousands of jobs with the same set of executables but different parameters and data.

Objectives

- Flexible tool
- Easy to create new productions
- Reliable – should run weeks without interruption
- Bookkeeping – provide information about processed jobs, fail reasons, success, etc.

GGG

OSG Users meeting Fermilab 2007
Nikolay Kuropatkin

To create production means to create thousands of condor submit files and DAG files. This means that some automation of the process should be provided.

The easiest way to provide flexible and automatic creation of many files is to use templates.

GGG is using XML template format – human readable and editable.

GGG build in Java – easy to modify to meet requirements and data model of any production

GGG

Assumption – with any data model it should be possible to create a unique id for each job (jobID). This can be just a serial number.
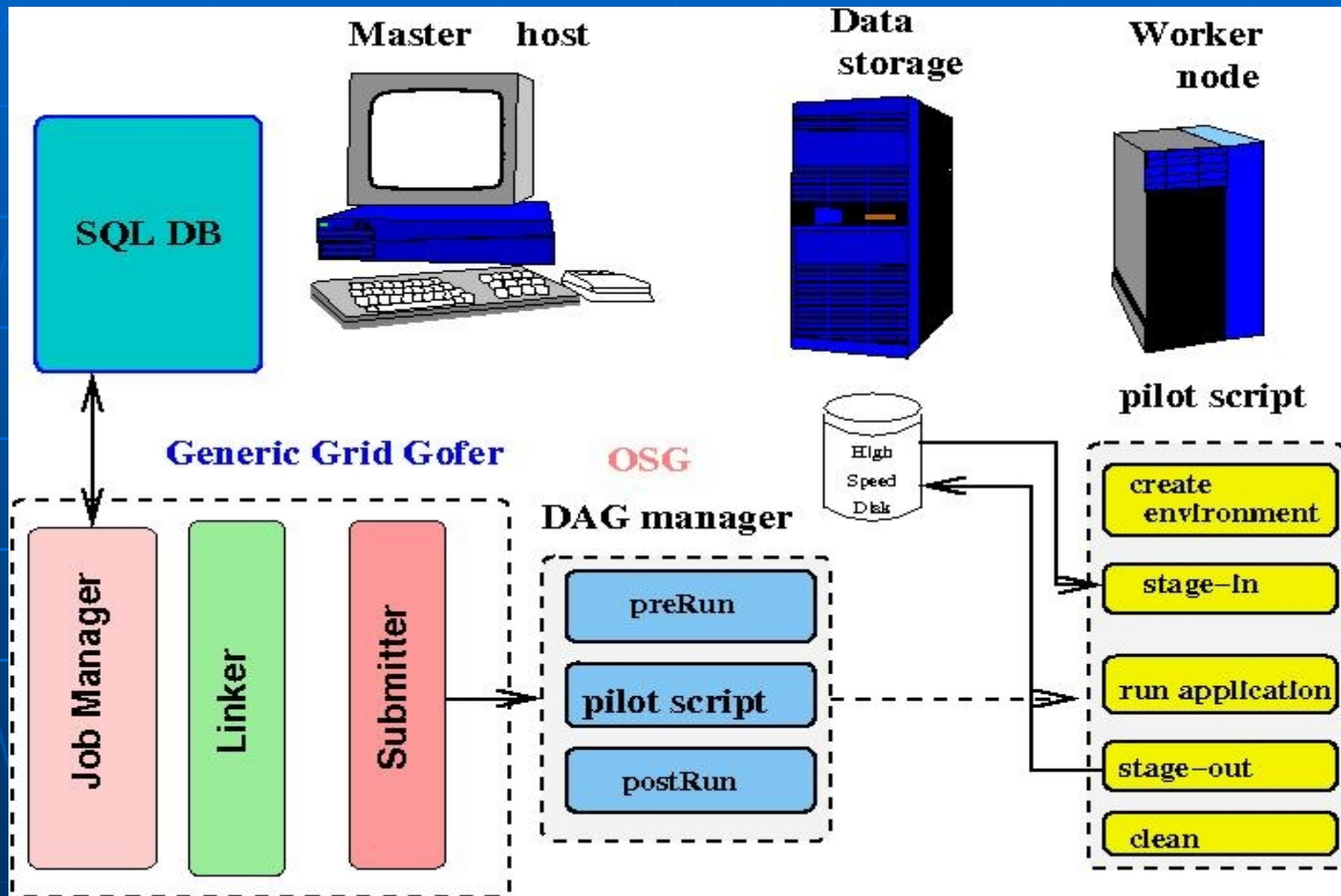
Requirement – all jobIDs are put in "jobs" table in the MySQL database. The package contains a simple Java program that can be used as an example

Requirement – All site names where production will be run should be put in "pool" table in the MySQL database.

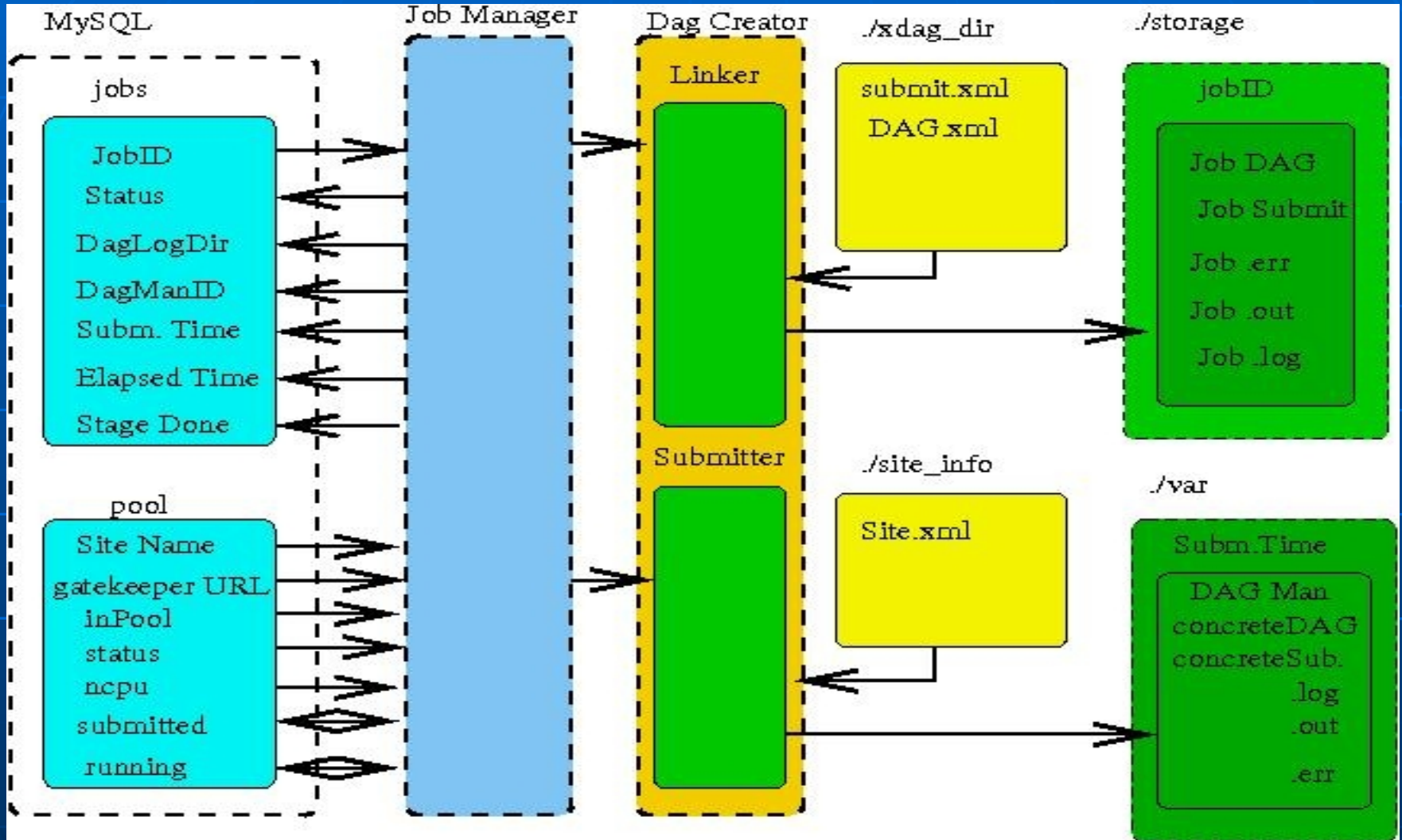The package also can use Postgress database.

GGG

Typical Job Structure

The package structure

Having all the information in the database it is easy to build graphs like this.

OSG Users meeting Fermilab 2007
Nikolay Kuropatkin

It is also easy to clean disk space after completed jobs, and to find logs for failed jobs.
A simple database operation permits to resubmit failed jobs, possibly to another site.
User can change production speed by switching on and off some sites or changing available worker nods on a site with simple database operation.

More details about the package, and a distribution file can be foud on following URLs
http://home.fnal.gov/~kuropat/sdss_grid/GGG.html
and

http://home.fnal.gov/~kuropat/sdss_grid/Overview.html

OSG Users meeting Fermilab 2007
Nikolay Kuropatkin

GGG