Transformation system

Federico Stagni

‣ What's this

‣ How it works

‣ Agents

Note to the audience: credits for the
development go to Andrew Smith

DIRAC user forum   Barcelona, May 12th – 13th 2011

What's this?

- A DIRAC system, as usually comprising:
  - MySQL tables
  - DIRAC Services
  - DIRAC Agents
  - Clients
  - Scripts
  - A set of **Plugins**
- Used for "repetitive" work
  - Can be used for different workflows
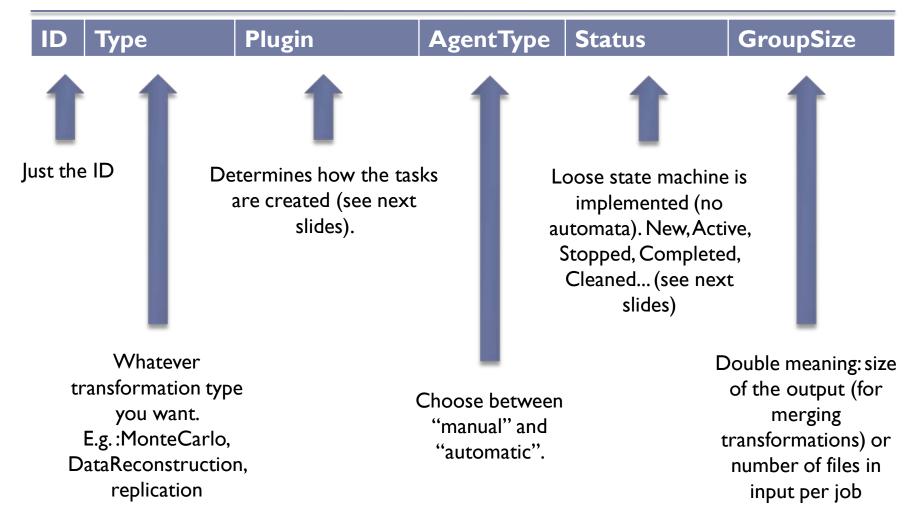 - A generic enough system

1. A new **transformation** is created

2. There are (or there aren't) **input** files

3. Each transformation has a **plugin**

4. **Tasks** are created based on the specified plugin

5. Tasks are treated by an **external system**

ID	Type	Plugin	AgentType	Status	GroupSize

Just the ID

Determines how the tasks are created (see next slides).

Loose state machine is implemented (no automata). New, Active, Stopped, Completed, Cleaned... (see next slides)

Whatever transformation type you want.
E.g. : MonteCarlo, DataReconstruction, replication

Choose between "manual" and "automatic".

Double meaning: size of the output (for merging transformations) or number of files in input per job

> Each Transformation can have a number of input files
>> E.g.: Data Reconstruction transformations

TransfID	FileID	Status	ErrorCount	TaskID

Unique FileID + TransfID. FileID associated to an LFN

Number of times the file failed to be processed

"Unused": ready to be "Assigned" to a task. "Processed": the task succeeded.

TaskID processing the File (no history)

▸ ## Standard():

   ▸ Group files by replicas (tasks create based on the file location)

▸ ## BySize():

   ▸ Group files until they reach a certain size (Input size in Gb)

▸ ## ByShare()

   ▸ Group files given the share (specified in the CS) and location

For replication:

▸ ## Broadcast()

   ▸ Takes files at the source SE and brodcast to a given number of locations

Tasks are treated by an external system:
- WMS (all but replication and removal transformations)
- DMS (replication and removal transformations)

TransfID	TaskID	ExternalID	ExternalStatus

Combination of Transformation ID and Task ID is unique

External system ID (e.g. WMS job ID)

External system status (e.g. "Done", "Failed", …)

Questions

DIRAC User Forum   Barcelona, May 12th – 13th 2011