An Introduction to Ganga

James Walder
Lancaster University

- Based on slides by Mark Slater

Overview

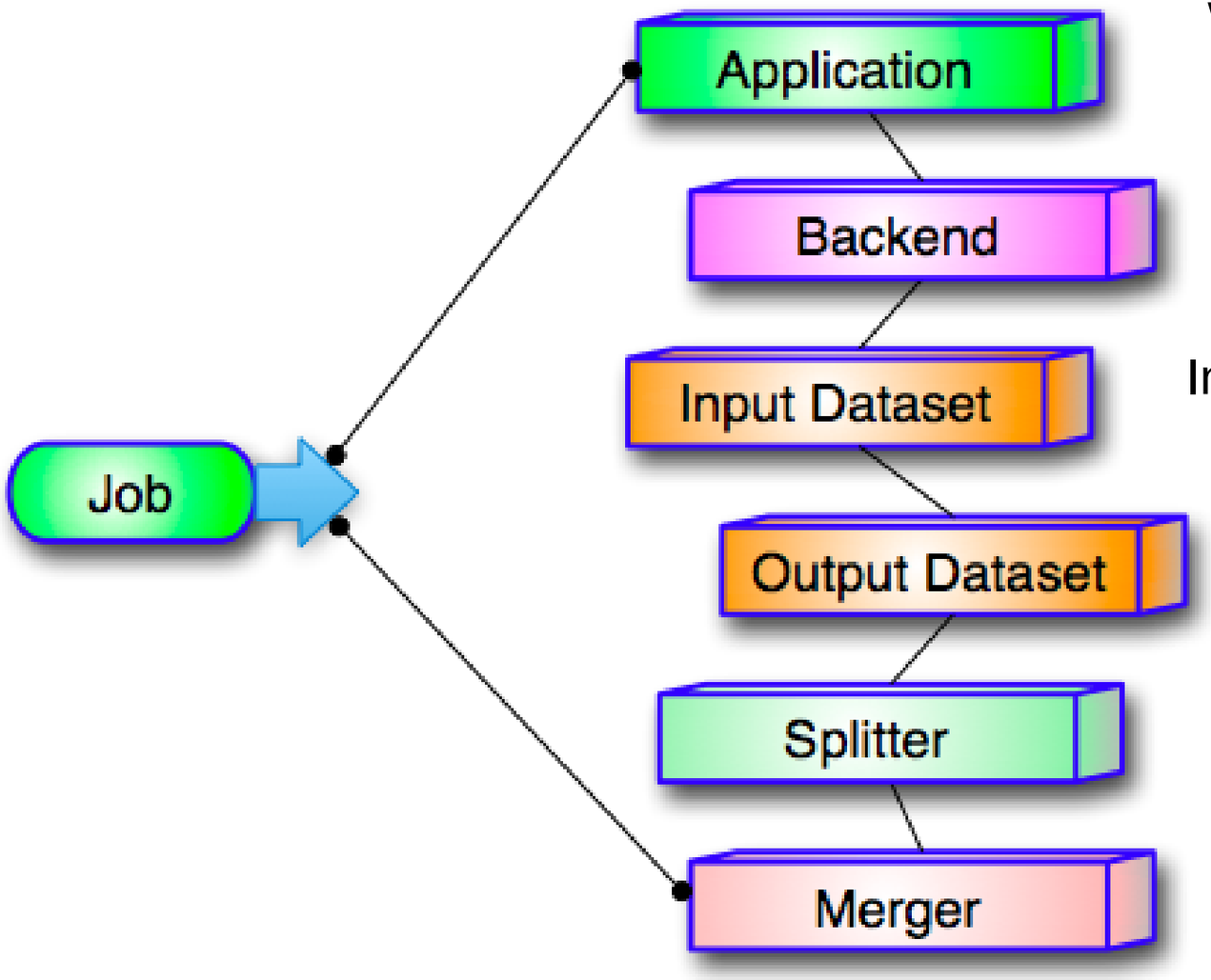
- Ganga is a general job management tool:
  - Designed to simplify the submission and book-keeping (or monitoring) of some form of “job”.
- Ganga is not exclusive to ATLAS
  - or even to High-Energy Physics

- Jobs can be submitted, monitored and retrieved by Ganga to:
  - Local computers
  - Local batch cluster
  - The GRID (LCG, EGEE, OSG, NorduGrid, Panda)

- You tell Ganga what you want to do,
  - It will prepare the job, and submit to the grid.
- When the job completes, Ganga will inform you.
  - Use the resubmission tools to retry any failed jobs.
  - Use the DQ2 tools to retrieve your output (from the Grid).

The Ganga Job

- This flexibility to run any kind of application on many different types of systems is provided by Ganga using the concept of:
  - independent modules



What to run - Athena

Where to run - Panda

Input data - Physics container

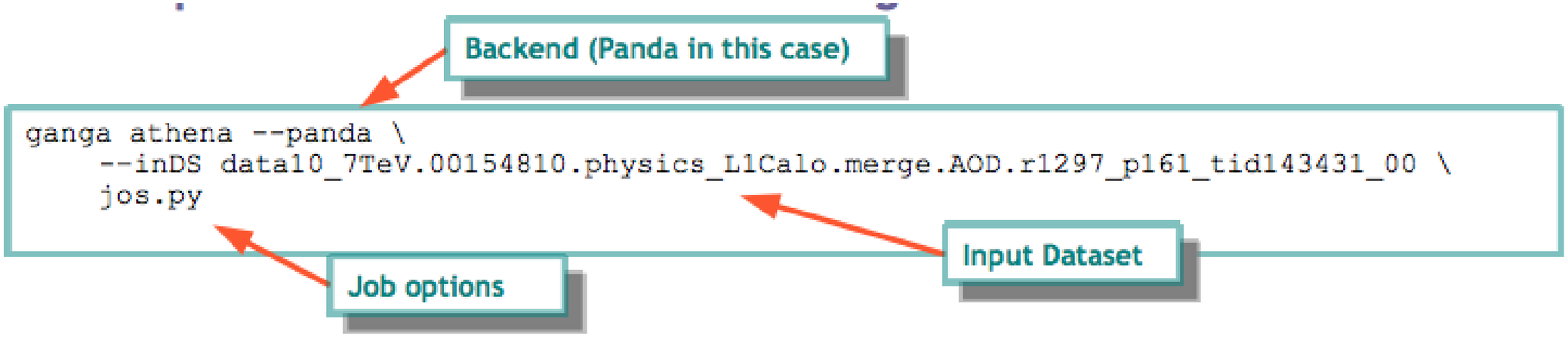
output data - user dataset, root files

Rule for dividing the workload

Rule for combining output

Ganga on the Command line

- Simplest method of job submission using Ganga is via the command line



- Additional arguments can be specified

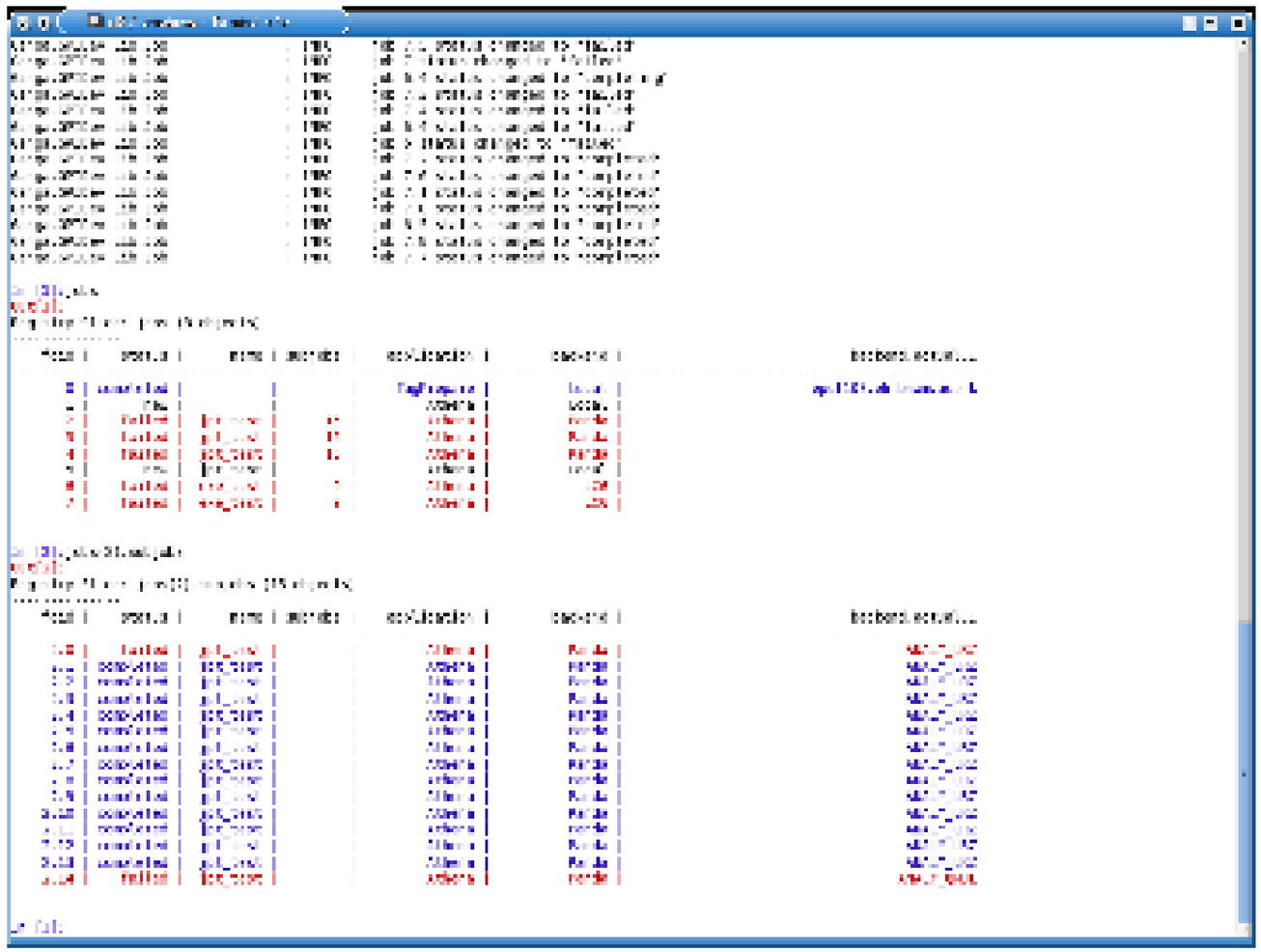
```
--outDS <ds_name>        # Specify an output datasetname
--cloud UK               # Specify a specific cloud to run on
--cloud <site>           # Specify a specific site to run on
--split 20               # Split over a number of subjobs

--athena_exe PYARA       # run as: python <jos.py>
--athena_exe ROOT        # run as: root <jos.C>
--athena_exe EXE         # run generic executables or TRFs (c.f. prun)
```

- Not restricted to Athena jobs,
  - Also run standard Python or ROOT scripts
  - Transforms can also be handled in this way.
- Ganga also provides a more advanced interface

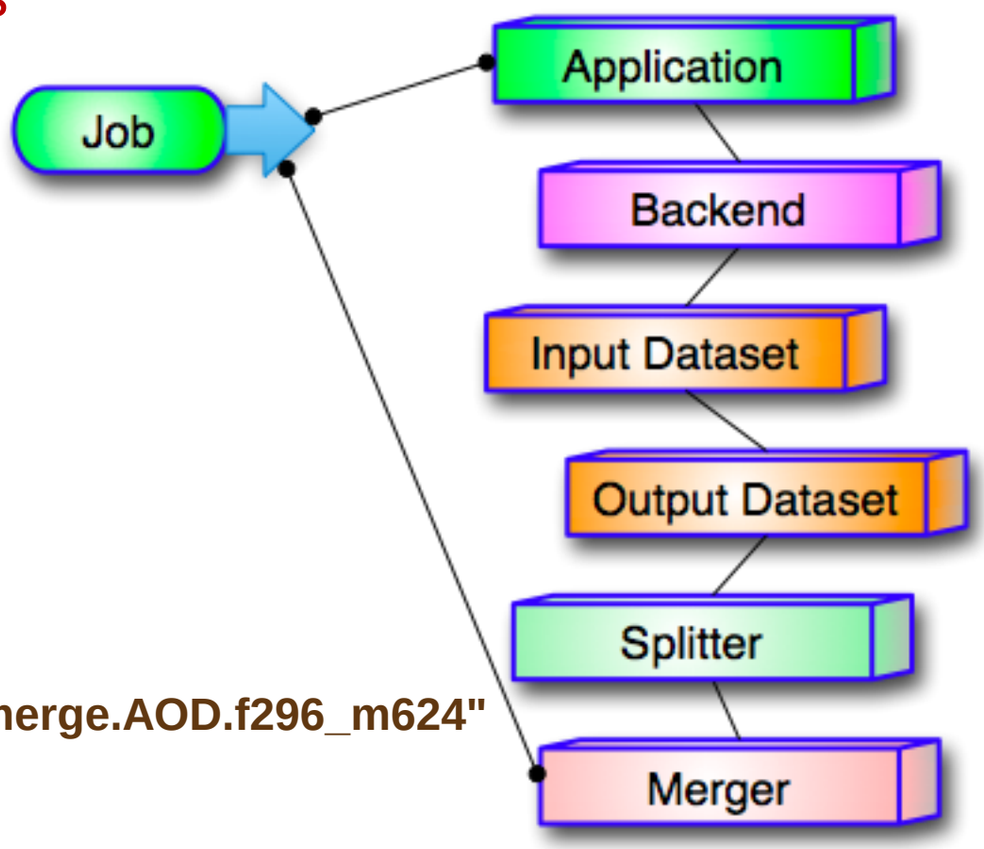
Ganga Interface

- Start Ganga with `ganga`
- Will start the Ganga interactive interface, using IPython.
  - Full python support - if you like scripting, full power is available.
- Simple commands to display monitoring information
  - `jobs` List all jobs, and their current status
- Additional details about a specific job using:
  - `jobs(22)`
    - List details about job 22, for example



Simple Athena Job

- Using the Ganga interface, we can construct a simple Job.
  - Run Athena using your favourite job options
  - Over your favourite datasets




```
j = Job()
j.application = Athena()

j.application.option_file = 'AnalysisSkeleton_topOptions.py'
j.application.prepare()

j.backend=Panda()

j.inputdata = DQ2Dataset()
j.inputdata.dataset = "data10_7TeV.00166786.physics_Muons.merge.AOD.f296_m624"

j.outputdata=DQ2OutputDataset()

j.splitter = DQ2JobSplitter()


j.submit()
```

- Update on progress by typing **jobs**
- Ganga will create an output dataset name (or you can choose your own).

Ganga Workflows

- Ganga supports all of the major workflows possible in Atlas, some of the major ones being:
  - Basic Athena usage (as shown)
  - AthenaROOTAccess via ROOT or python
  - ROOT Processing (e.g. D3PDs)
  - Generic Executables and Transforms
  - Submitting using an AMI query as input
  - Submitting using a Good Runs List as input
  - Submitting using TAG files as input (through production or ELSSI)

- Ganga Tasks:
  - For when you need to run over many datasets.
- Tasks will:
  - Resubmit failed jobs if it seems sensible to do so
  - Submit more jobs when others complete
  - Monitor and track running jobs, storing output data in appropriate containers
- https://twiki.cern.ch/twiki/bin/viewauth/Atlas/GangaTasks

Links

- The General Ganga Manual
  - http://ganga.web.cern.ch/ganga/
- The Full GangaAtlas Tutorial
  - https://twiki.cern.ch/twiki/bin/viewauth/Atlas/FullGangaAtlasTutorial
- A GangaAtlas Quick Start Guide
  - https://twiki.cern.ch/twiki/bin/viewauth/Atlas/GangaAtlasQuickReferenceGuide
- GangaAtlas FAQ
  - https://twiki.cern.ch/twiki/bin/viewauth/Atlas/DAGangaFAQ
- The Distributed Analysis List
  - hn-atlas-dist-analysis-help@cern.ch

This Session: Ganga

- A very detailed full tutorial to Ganga is given in the ATLAS TWiki:
  - https://twiki.cern.ch/twiki/bin/viewauth/Atlas/FullGangaAtlasTutorial
- Will will try to cover only a small section of the TWiki:
  - In your own time you can continue through the rest of the tutorial.
- Ask questions !