A lightweight task submission and management infrastructure

Xiaomei Zhang Bing SUO Xianghu Zhao TianYan

CHEP2016, San Fransico

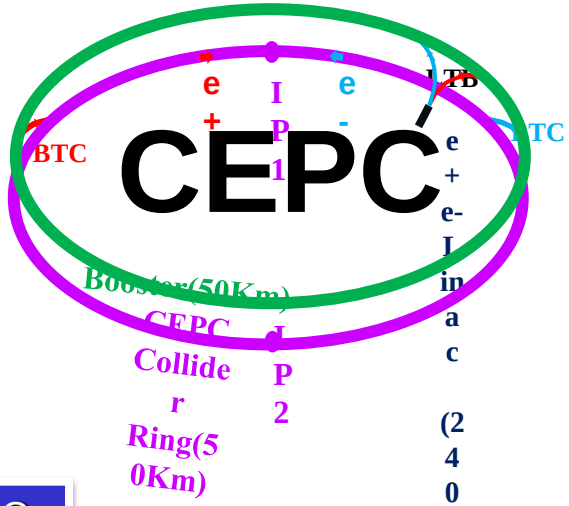
Oct 10, 2016

Experiments n IHEP

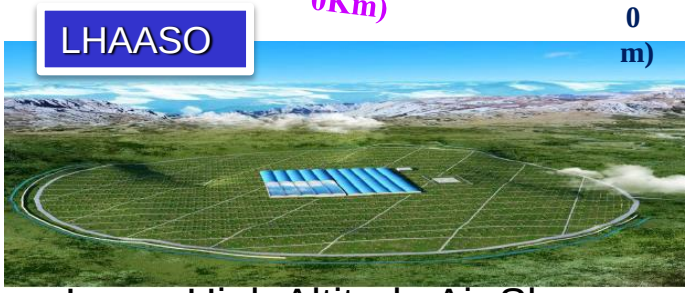
- Current experiments
  - BESIII
  - DYB
  - YBJ
- New experiments
  - JUNO
  - LHAASO
  - HXMT
  - CEPC

BESIII (Beijing Spectrometer III at BEPCII)

DYB (Daya Bay Reactor Neutrino Experiment)



JUNO (Jiangmen Underground Neutrino Observatory)

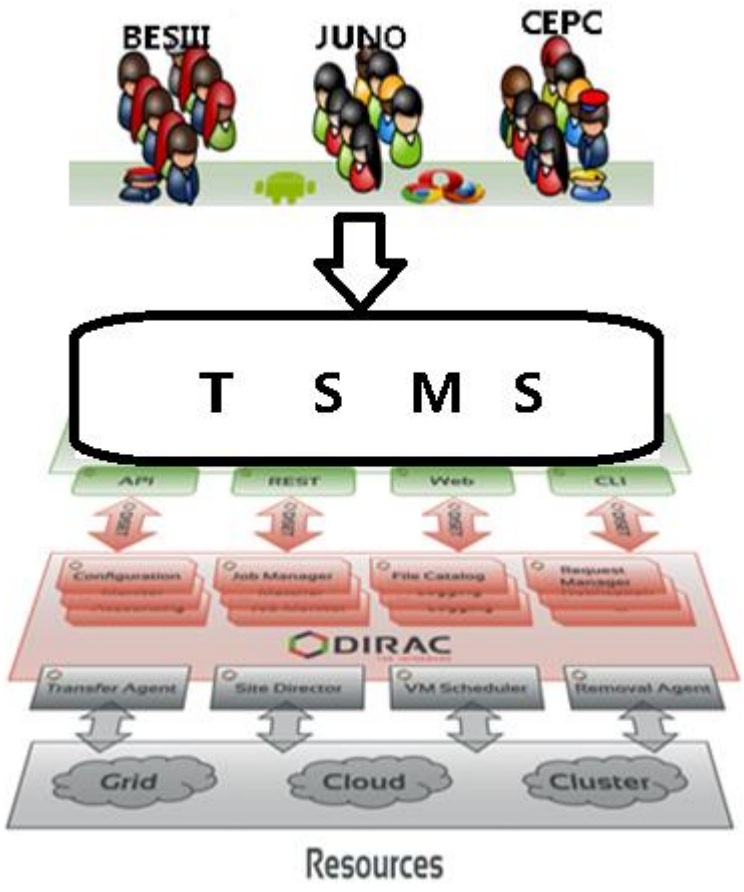


Large High Altitude Air Shower Observatory

Hard X-Ray Moderate Telescope

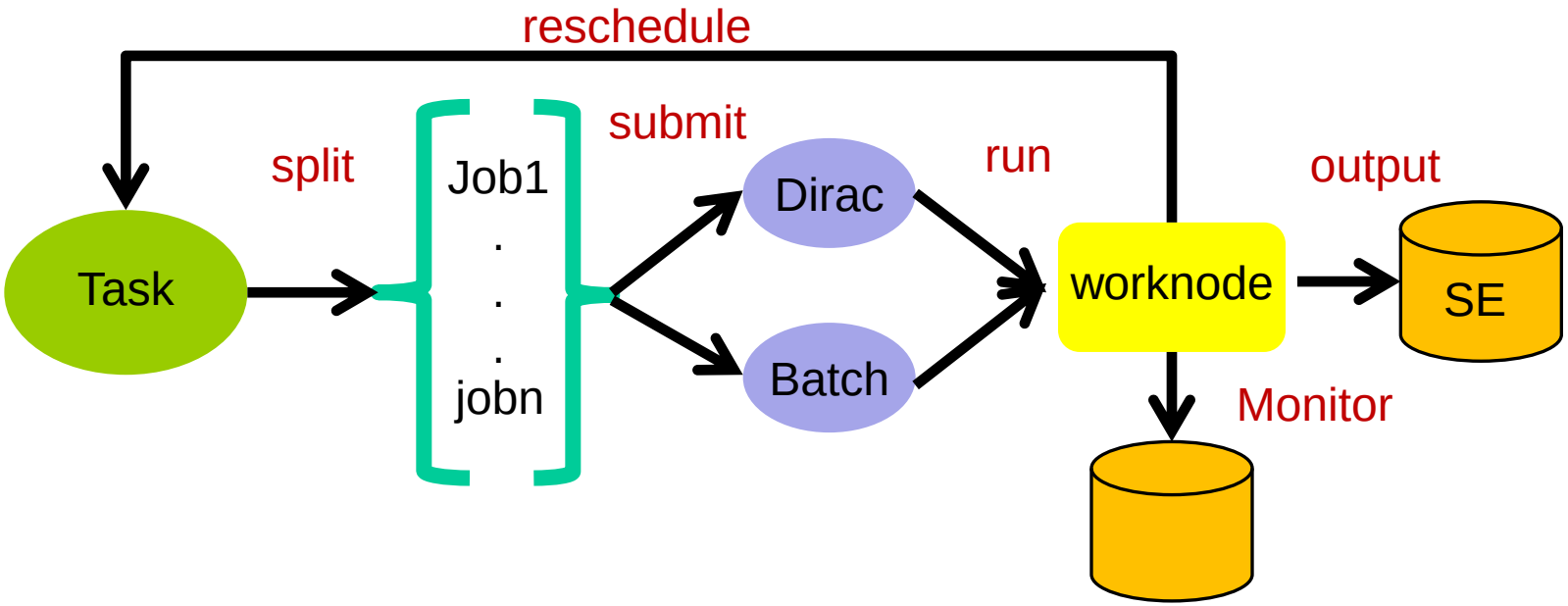
Multi-VO distributed computing

- More than one experiments express interests on using or evaluating distributed computing resources
- DIRAC was adopted to build distributed computing system for BESIII in 2012
- Multi-VO supports has been extended in one set-up to run JUNO and CEPC jobs
- Task Submission and Management System (TSMS) is one of the necessary steps for new experiments to have a try and use in their early stages

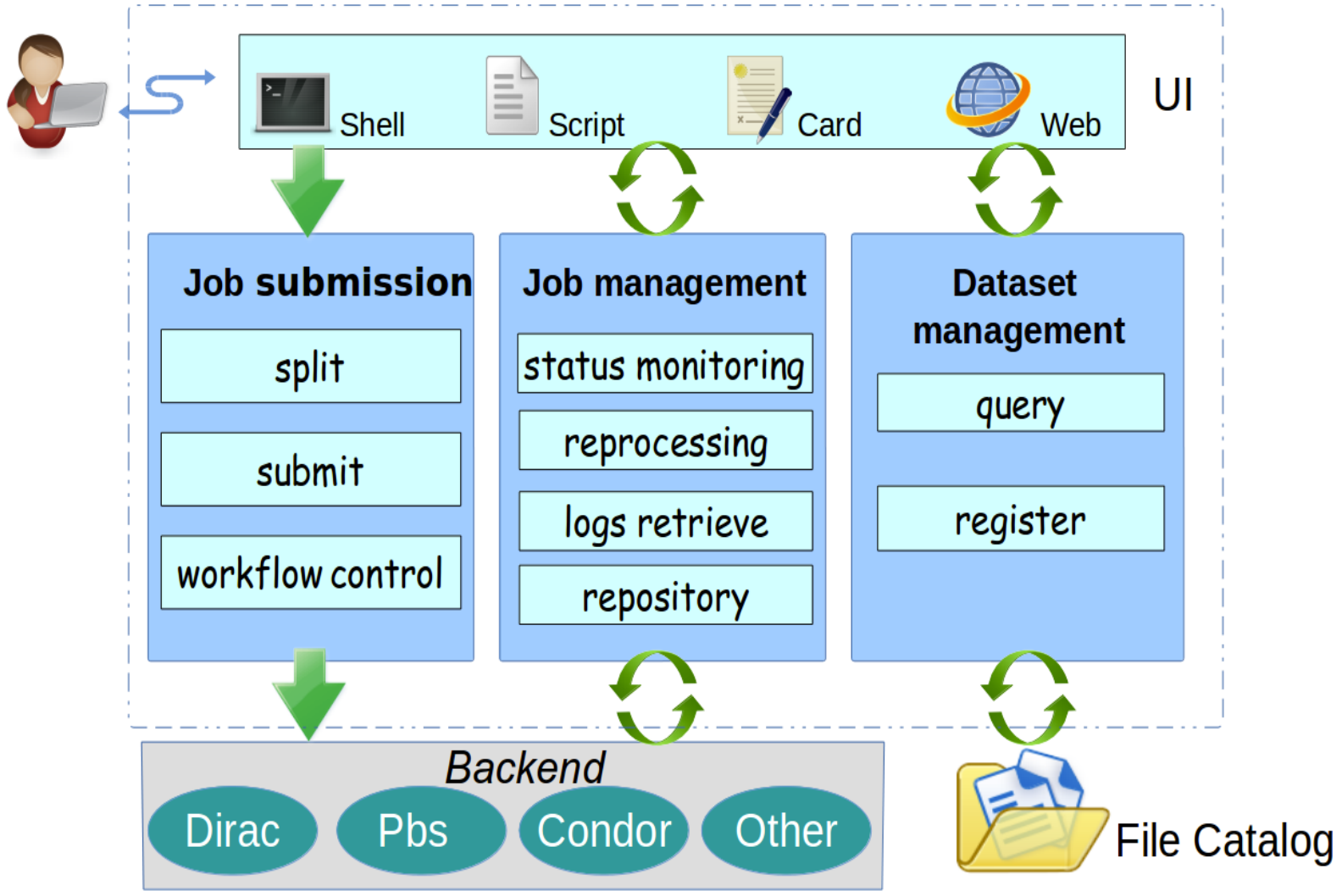


JSUB

- Goals:
  - Automatically manage massive jobs
  - Highly extensible for new experiments
  - Ease the procedure to use heterogeneous resources
- Life cycle of a task (a batch of jobs in one submission)
  - split->submit->run->status monitor->output retrieval -> reschedule



JSUB Architecture



User Interface

- YAML used to describe tasks
  - simple, clear, friendly with python
  - Data transformed into tree-like dictionary
- Shell commands are provided

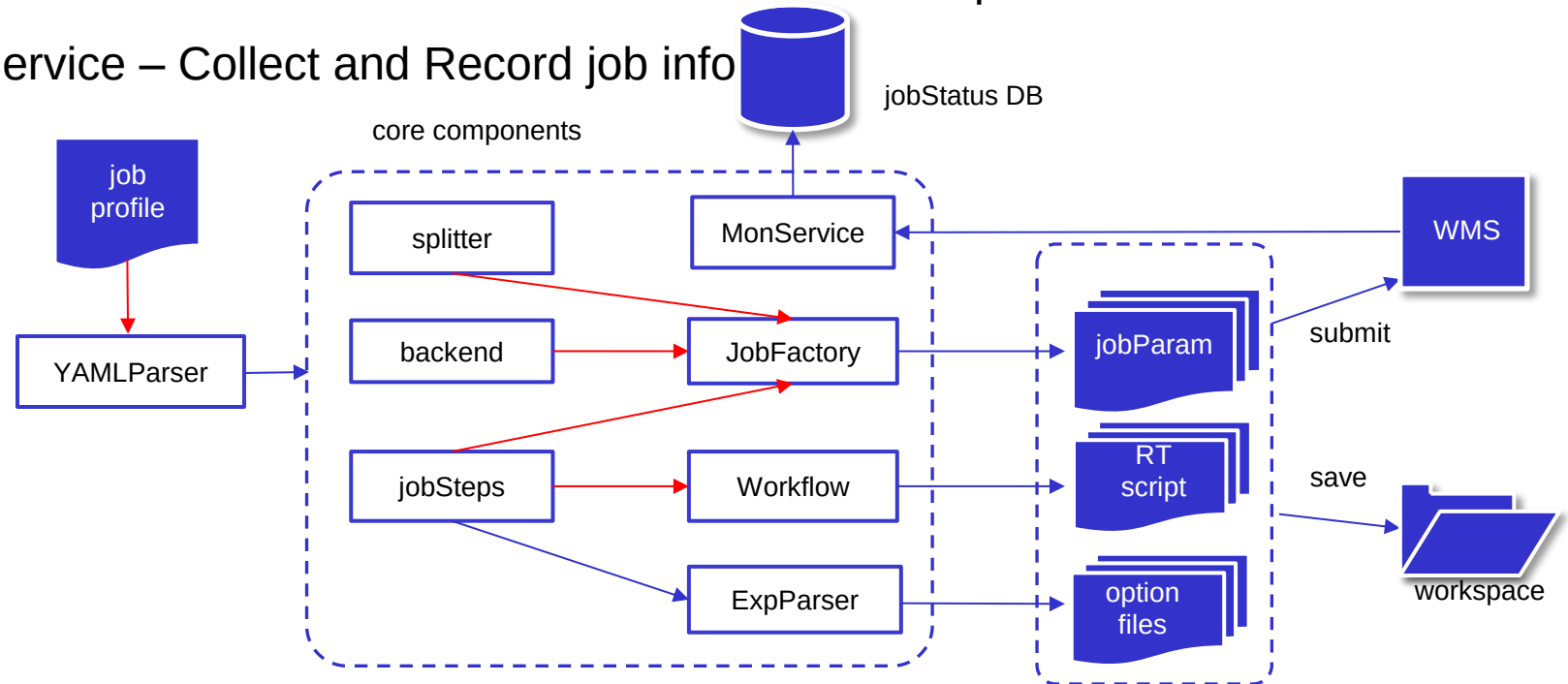
```
Country:
  Name: Liechtenstein
  Rank: 1
  Year: 2008
  Gdp: 141100
  Neighbor:
    - Name: Austria
      Direction: E
    - Name: Switzerland
      Direction: W
```

```
-bash-3.2$ ls -l
jsub_remove
jsub_resubmit
jsub_status
jsub_submit
-bash-3.2$ jsub_submit -h
usage: jsub_submit <options> <job profile>
options:
  -h    show the help information
  -v    show JSUB version
job profile: the file where the parameters of master job are defined

-bash-3.2$ jsub_status -h
usage: jsub_status <options> <target>
options:
  -h    show the help information
  -v    show JSUB version
  -all  show the status of all the jobs the user ever submitted
target: could be JobID, JobGroup
```

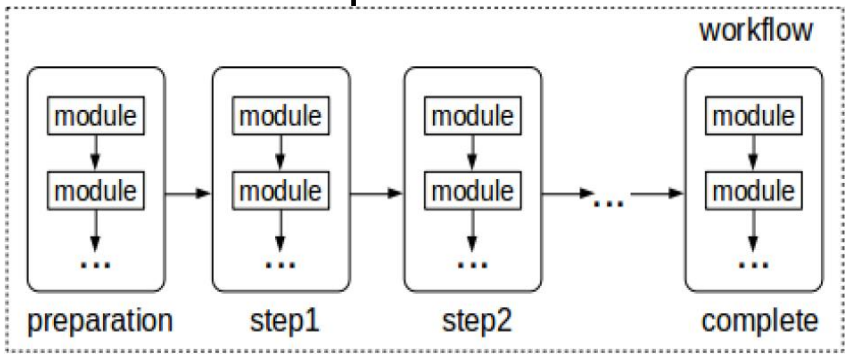
Job Split and Submit

- YAMLParser
- Splitter -- Split jobs with various way
- JobFactory -- Collect Job related parameters to generate Job submission scripts
- ExpParser -- Get runtime card or options for applications execution
- Workflow – Define runtime workflow to create RT scripts
- MonService – Collect and Record job info



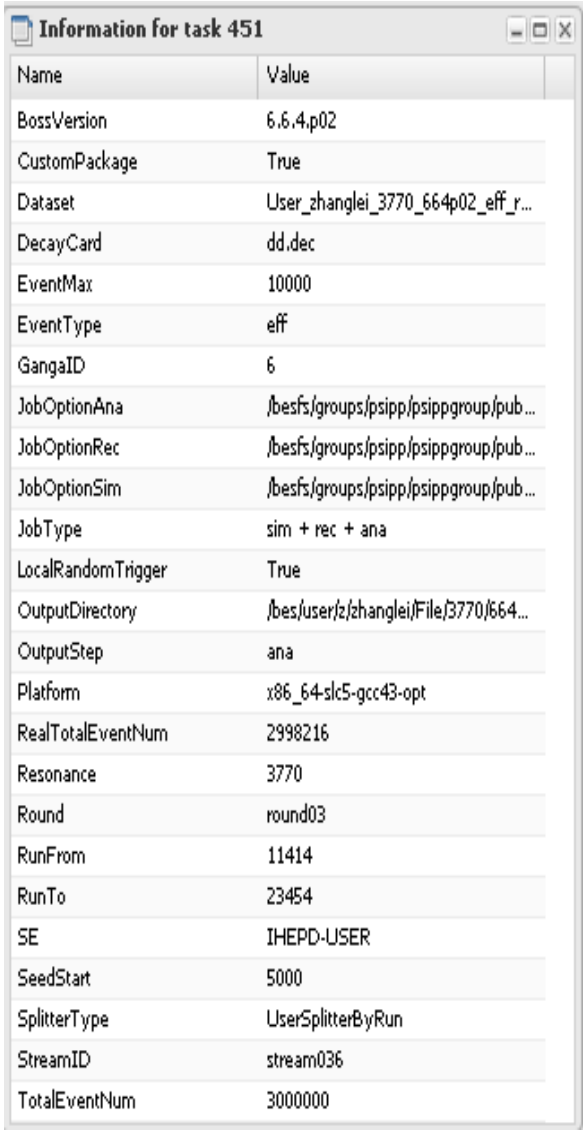
Workflow control

- Adopt ideas from “DIRAC workflow”
  - Highly extensible and reusable
- Classes for organizing run-time operations
  - Workflow
  - Step
    - Group of modules to complete a process, eg. preparation, simulation, reconstruction, complete
  - Module
    - Basic units for one operation or command, eg. checkenv



Task monitor and management(1)

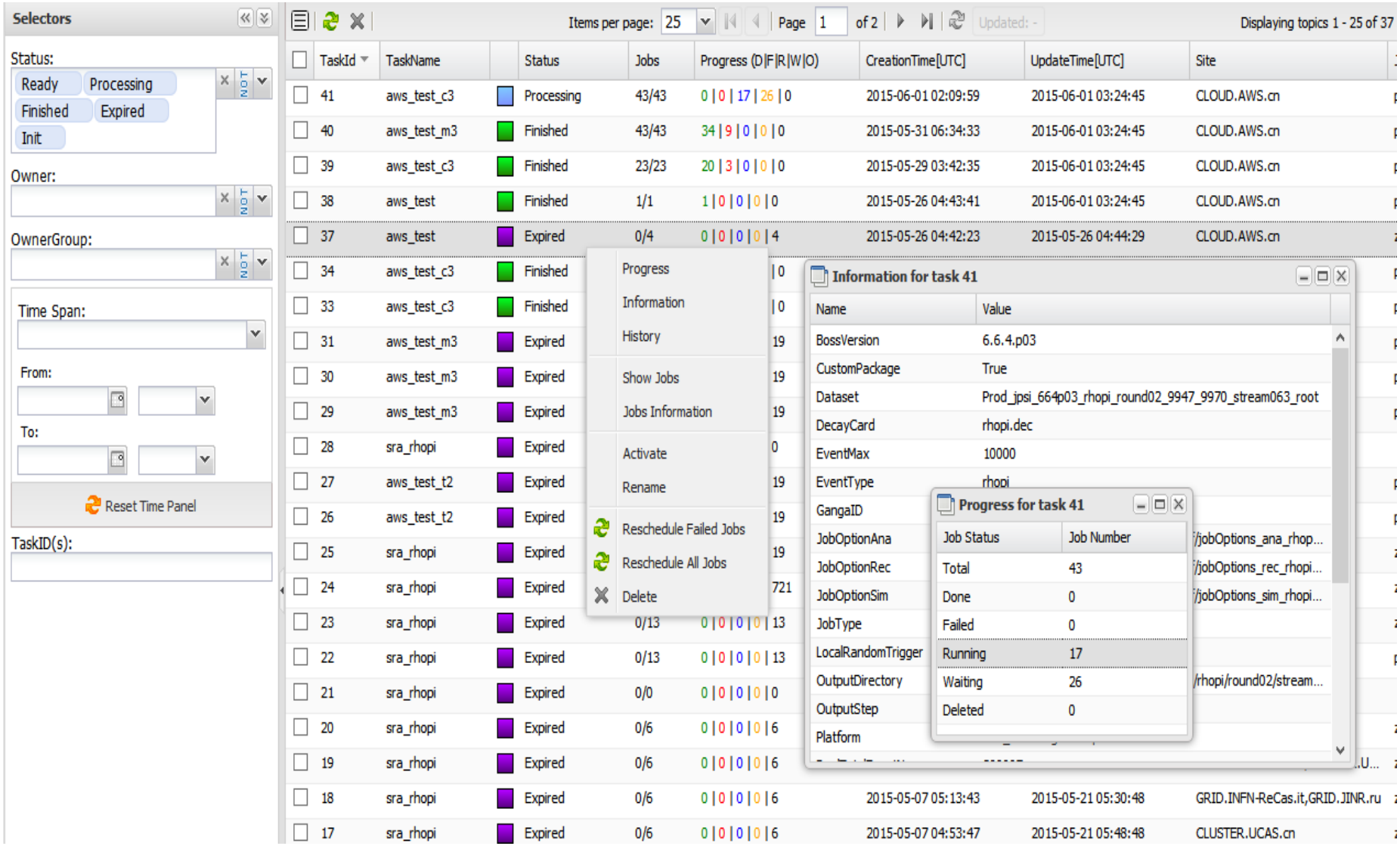
- TaskId uniquely identify the task, created after submitted
- Tasks can be monitored and controlled both trough commands and web portal (next slide)
  - Jsub_list, list all the tasks and the status
  - Jsub_status <taskId>, show task progress and information
  - Jsub_resubmit <taskId>, rescheule all or failed jobs
  - Jsub_remove <taskId>, delete the whole task
- Detailed task info collected from each stage of task processing
  - Physicists can easily track down specific task



Information for task 451

Name	Value
BossVersion	6.6.4.p02
CustomPackage	True
Dataset	User_zhanglei_3770_664p02_eff_r...
DecayCard	dd.dec
EventMax	10000
EventType	eff
GangaID	6
JobOptionAna	/besfs/groups/psipp/psippgroup/pub...
JobOptionRec	/besfs/groups/psipp/psippgroup/pub...
JobOptionSim	/besfs/groups/psipp/psippgroup/pub...
JobType	sim + rec + ana
LocalRandomTrigger	True
OutputDirectory	/bes/user/z/zhanglei/File/3770/664...
OutputStep	ana
Platform	x86_64-slc5-gcc43-opt
RealTotalEventNum	2998216
Resonance	3770
Round	round03
RunFrom	11414
RunTo	23454
SE	IHEPD-USER
SeedStart	5000
SplitterType	UserSplitterByRun
StreamID	stream036
TotalEventNum	3000000

Task monitor and management(2)

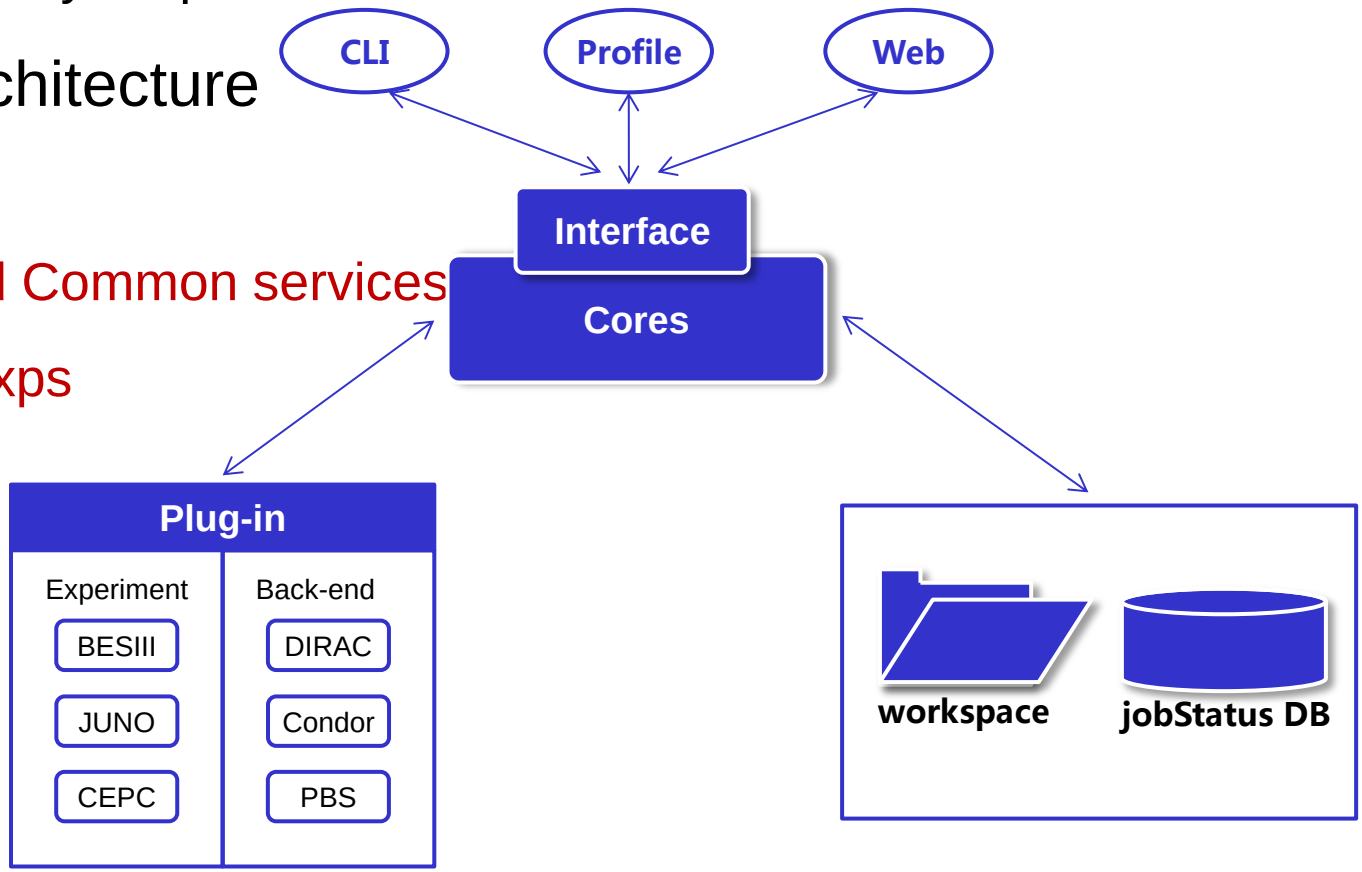


Dataset management

- Dirac File Catalog(DFC) used to build metadata and replica catalog
- Dataset class is built to contact with DFC
  - Query dataset with conditions or name
  - Register dataset for output of jobs
- Users can simply use Dataset as input data of task, move bulks of data by dataset

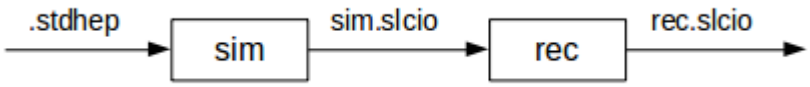
Implementation and Extensions

- Develop in python
- Object-oriented programing are used to make it reusable
  - Inheritance, polymorphism, reflection
- Main code architecture
  - Cores
  - Interface and Common services
  - Plug-in for Exps

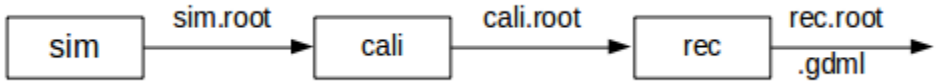


CEPC and JUNO case (1)

- CEPC



- JUNO



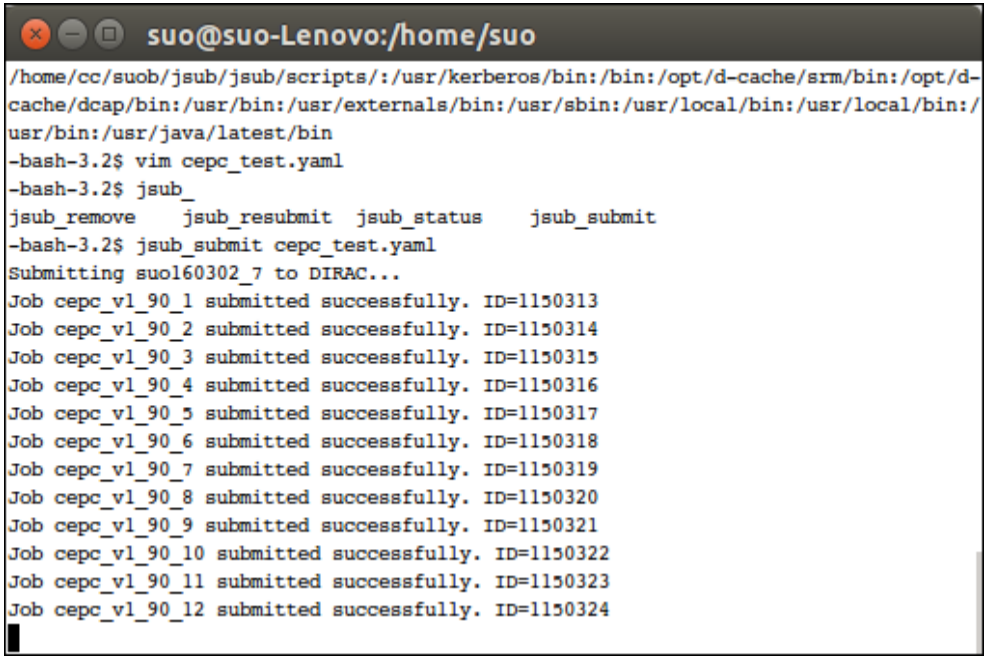
CEPC and JUNO case (2)

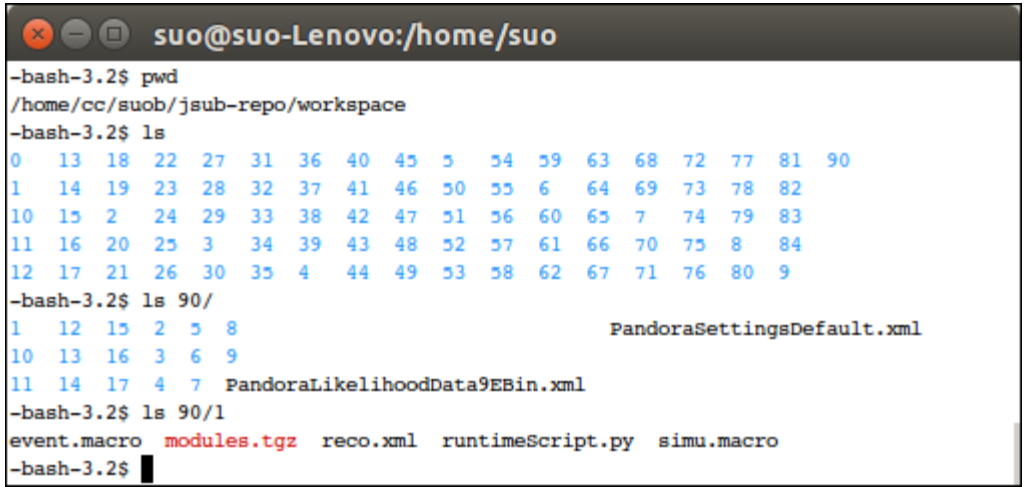
CEPC_test.yaml

```
Experiment:
  Name: cepc
  JobSteps:
    - Type: Sim
      Executable: Mokka
      JobOption: /home/cc/suob/jsub/workfiles/sim
      ReturnData: true
    - Type: Rec
      Executable: Marlin
      JobOption: /home/cc/suob/jsub/workfiles/rec
      ReturnData: true
```

```
Splitter:
  Type: ByFile
  EventMaxPerJob: 20
  SeedStart: auto
  InputData: ./stdhep.list
```

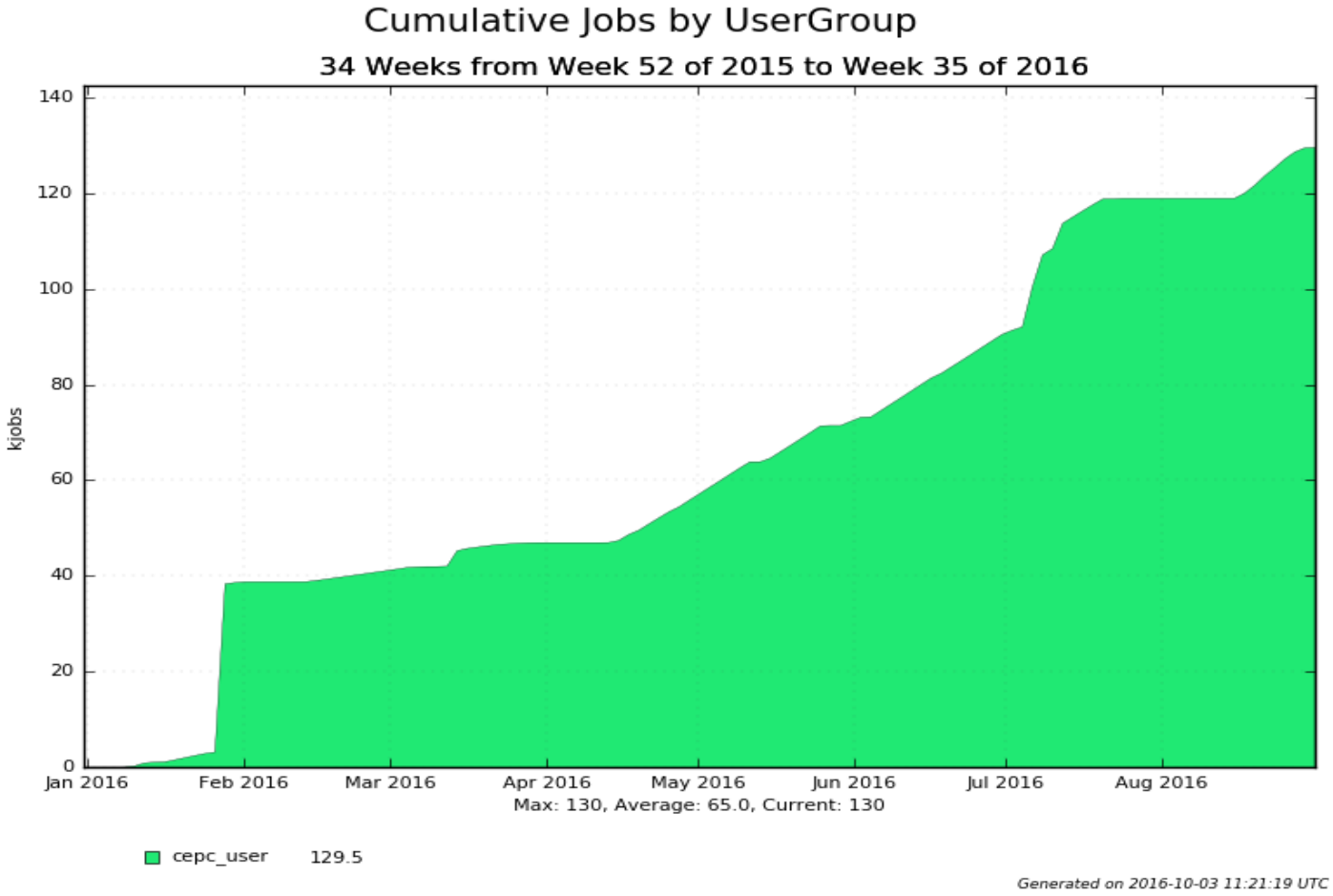
```
Backend:
  Name: dirac
  Site:
    - CLOUD.IHEP.cn
    - GRID.JINR.ru
  JobGroup: suob160302_7
  OutputSe: WHU-USER
  OutputDir: /test/cepc/0302_7
```





CEPC and JUNO case (3)

- In 2016, about 130K CEPC jobs have been successfully submitted through JSUB



Summary

- JSUB is a lightweight and extensible task submission and management tool
  - Small experiments or experiments in early stage can extend it for quickly access of distributed resources for massive production
- JSUB has been implemented in Python
  - Dirac and Condor supported
- Real cases of CEPC and JUNO experiments have shown its extensibility and simplicity
- CEPC usage in 2016 has proved its stability