JEDI-alpha Converter

Dmitry Golubkov

PanDA Workshop @ UTA

3 September 2013

Introduction

This is a schematic description of flow of logic JEDI-α and how it compares to Vanilla ProdSys I Sequence:

| Vanilla	[1] "Task Definition"	AKTR tables	[2] ProdSys tables	[3] ProdSys 1 sequence
α	-	-	[2'] DEfT tables	[3'] JEDI sequence

- New production system milestone required special converter
- Converter workflow
  - User authorization (CERN SSO)
  - Task submission
  - Generating Task parameters
  - Filling DEFT_TASK.TASK_PARAM

- Prototyped several protocols: XML, JSON, …
  - Adopted JSON Protocol
- Creating JSON dictionary
- Task parameters contain details required for JEDI:
  - JEDI generates jobs using those parameters and run execution

JEDI-alpha Converter

Production
Manager
Submits Tasks
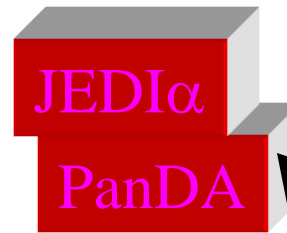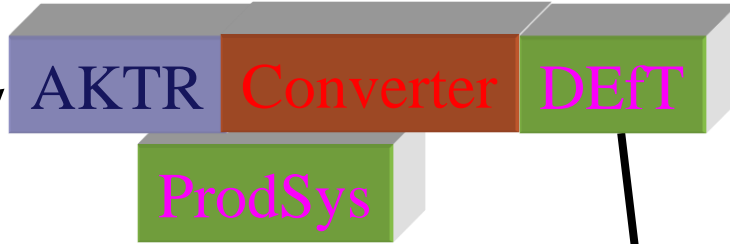
User

Database tables

- ATLAS_GRISLI.T_TASK_REQUEST @ ADCR_PANDA
- ATLAS_DEFT.DEFT_TASK @ ADCR
- Django models for all tables

T_TASK_REQUEST

	REQID	PROJECT	INPUTDATASET	TASKNAME	FORMATS	TRF
1	1246321	data11_cvalid	data12_8TeV.00212967.physics_EnhancedBias.recon.RAW.r4488	data11_cvalid.00212967.physics_EnhancedBias.recon.r4488_r4539	AOD	Reco_trf.py
2	1246319	data11_cvalid	data11_cvalid.00189421.physics_EnhancedBias.merge.RAW.	data11_cvalid.00189421.physics_EnhancedBias.recon.r4529	RAW	BatchHLTT...
3	1246317	data11_cvalid	data11_cvalid.00189421.physics_EnhancedBias.merge.RAW.	data11_cvalid.00189421.physics_EnhancedBias.recon.r4503	RAW	BatchHLTT...

DEFT_TASK

	TASK_ID	TASK_META	TASK_STATE	TASK_PARAM	TASK_TAG	TASK_COMMENT
1	1246321	(null)	pending	{ "architecture": "i686-slc5-gcc43-opt", "cloud": "CERN", "coreCoun...	r4539	(null)
2	1246319	(null)	pending	{ "architecture": "i686-slc5-gcc43-opt", "cloud": "CERN", "coreCoun...	r4529	(null)
3	1246317	(null)	pending	{ "architecture": "i686-slc5-gcc43-opt", "cloud": "CERN", "coreCoun...	r4503	(null)

Task submission via Converter

- Implemented non crones version of ProdSys
  - Calling as program, on demand
  - No writing into ProdSys tables
  - Multiple instances support
- Caching support in Google Drive API
- Support of non-direct URLs
- Filling TASK_REQUEST
- TASK_REQUEST.COMMENT_ includes link to file for edit (in Google Docs)
- Special flag TASK_REQUEST.GRID = "jedi@cern" ("ProdSys 2 processing")

Migration to SLC6 node

- Migration in progress together with *S. Baranov*
- Python 2.5 -> Python 2.6
- New service hostname

Conclusion

- JEDI can generate jobs using Task parameters generated by Converter

- Parallel task workflow was implemented

- CLI for Converter was implemented