



PyCon 2012

Grid Job management

Felix Lee, ASGC



About ASGC

Academia Sinica Grid & Cloud

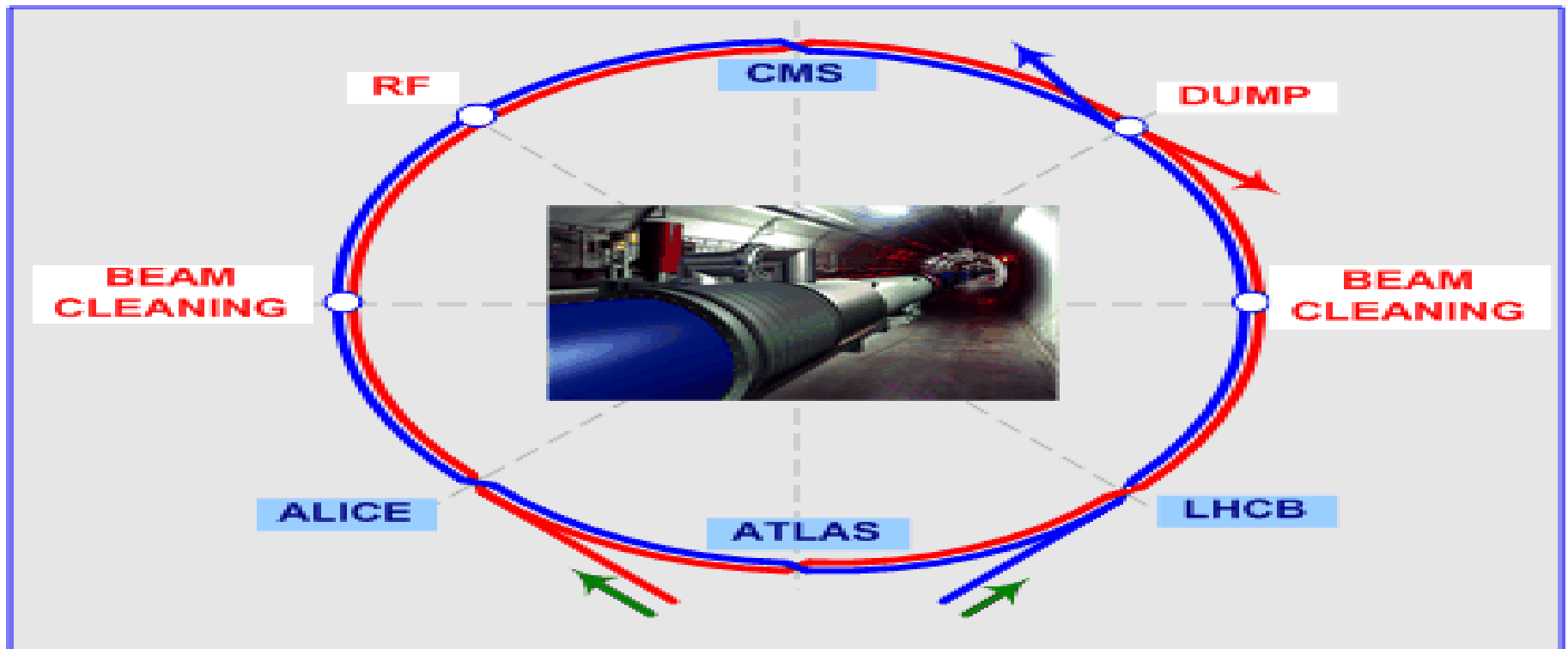


Something we might need to know..

- LHC
- WLCG
- Grid Computing

LHC experiment

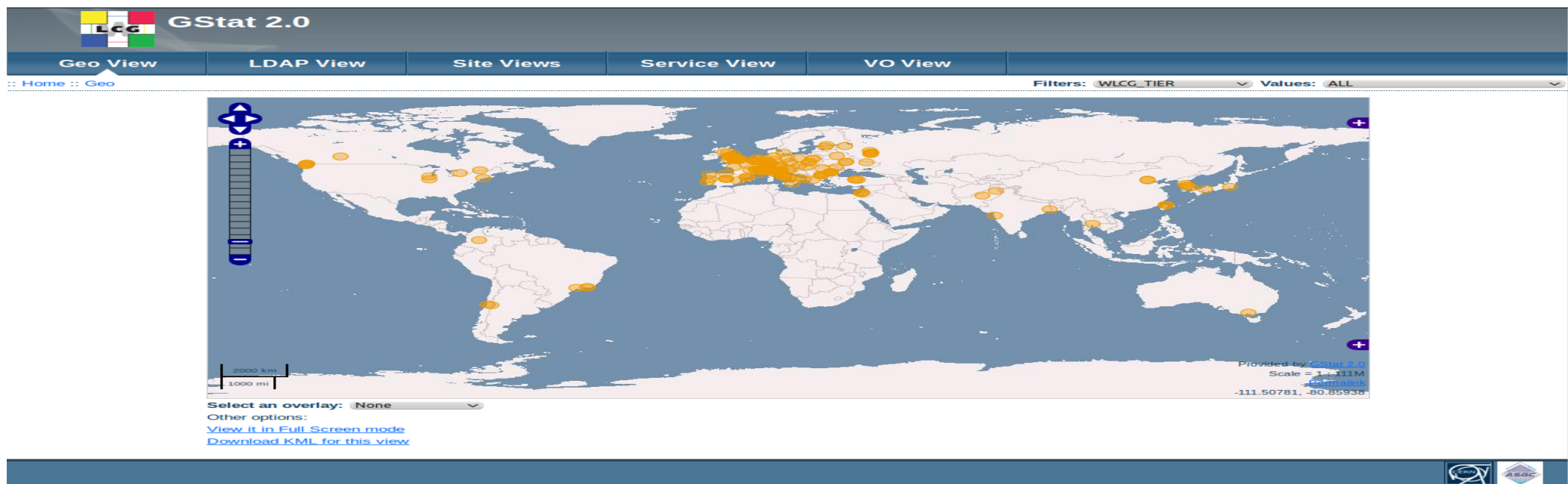
- LHC – The Large Hadron Collider.
 - It was built by European Organization for Nuclear Research (CERN)
 - 27KM tunnel in circumference, as deep as 175M





WLCG

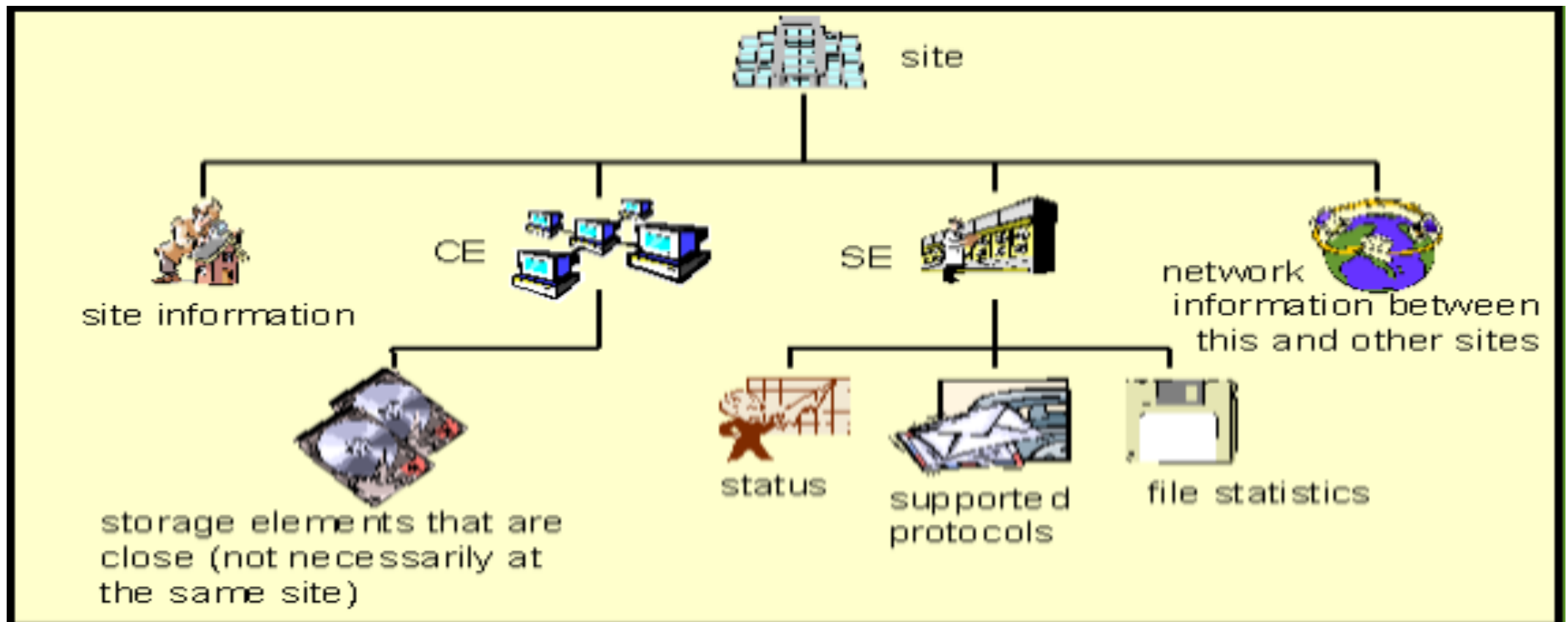
- World-wide LHC Computing Grid
- It's a distributed computing infrastructure to provide the production and analysis environment for LHC experiment.
- Currently, there are 11 tier1, 140 tier2 and several small tier3 in the world.
- There are 269299 CPU cores, 183PB storage capacity in the world.





Grid Computing

- It's one of distributed computing.
- Base on federal resources.
- It connects loosely-coupled computers by the Internet to be super virtual computer.





What we do

- ASGC is WLCG(World-wide LHC Computing Grid) Tier 1 operation center since 2005
- ASGC is also conducting Asia Pacific regional e-Science collaborations, development and infrastructure operation.
- Developing new generation distributed computing infrastructure and technologies.



Python for us



Python in WLCG & Grid

- It's widely used for high level integration.
- Clear code, clear syntax...
- Totally open source.
- Fast and flexible implementing.
 - It's script.
 - No need to be complied.
 - Plenty of mathematic and science modules.



Python in WLCG & Grid

- Work flow & Job Management.
- Data Management.
- Information system.
- Monitoring.
- HEP applications
 - Data processing.
 - Data analysis.



Computing system in WLCG/Grid

- They are all integrated/implemented by Python
 - WMAgent:
 - Workload Manager Agent.
 - GRAB:
 - CMS Remote Analysis Builder.
 - PanDA:
 - Production and Distributed Analysis system.
 - DIRAC:
 - Distributed Infrastructure with Remote Agent Control
 - AliEn:
 - Alice Environment
 - DIANE:
 - Distributed Analysis Environment



Python in ASGC

- Work flow & Job Management
 - GAP 1.0 (base on DIANE)
 - PanDA, collaborating with Atlas
- Monitoring and information
 - GSTAT 2.0, Nagios plugin.
- Integration of Grid & Cloud.
 - Virtual worker node on demand.
 - Virtual machine catalog service.
- Deployment and automation.



GStat 2.0



GStat 2.0

[GGUS](#) | [Contact](#) | [About](#)

[Geo View](#)

[LDAP View](#)

[Site Views](#)

[Service View](#)

[VO View](#)

[Home](#) :: [Site Summary](#)

Filters: [GRID](#) Values: [EGI](#)

Show 10 entries | Go to a site: [--SELECT A SITE NAME--](#)



Search:

Name	Status	CPUs			Online Storage Space (GB)		Nearline Storage Space (GB)		Grid Jobs		
		Physical	Logical	SI2000	TotalSize	UsedSize	TotalSize	UsedSize	Total	Running	Waiting
giNUIMie	OK	2	2	2,000	1,082	<div><div></div></div> 5%	0	<div><div></div></div> 0%	0	<div><div></div></div> 0%	<div><div></div></div> 0%
GoeGrid	OK	578	2,484	6,373,944	505,037	<div><div></div></div> 72%	0	<div><div></div></div> 0%	3,634	<div><div></div></div> 110%	<div><div></div></div> 24%
GR-01-AUTH	OK	48	272	572,560	7,874	<div><div></div></div> 43%	0	<div><div></div></div> 0%	131	<div><div></div></div> 12%	<div><div></div></div> 73%
GR-04-FORTH-ICS	CRITICAL	12	12	10,764	6,750	<div><div></div></div> 0%	0	<div><div></div></div> 0%	9	<div><div></div></div> 75%	<div><div></div></div> 0%
GR-06-IASA	OK	20	20	36,100	401	<div><div></div></div> 49%	0	<div><div></div></div> 0%	0	<div><div></div></div> 0%	<div><div></div></div> 0%
GR-07-UI-HEPLAB	OK	39	156	468,000	199,175	<div><div></div></div> 35%	0	<div><div></div></div> 0%	35	<div><div></div></div> 22%	<div><div></div></div> 0%
GR-09-UoA	OK	5	10	24,010	312	<div><div></div></div> 7%	0	<div><div></div></div> 0%	1	<div><div></div></div> 0%	<div><div></div></div> 100%
GR-10-UI	OK	120	120	231,720	939	<div><div></div></div> 65%	0	<div><div></div></div> 0%	904	<div><div></div></div> 32%	<div><div></div></div> 95%
GR-11-UPATRAS	OK	1	4	8,352	992	<div><div></div></div> 5%	0	<div><div></div></div> 0%	15	<div><div></div></div> 0%	<div><div></div></div> 100%
GRIF	CRITICAL	1,246	5,727	10,548,380	2,205,479	<div><div></div></div> 70%	0	<div><div></div></div> 0%	11,101	<div><div></div></div> 80%	<div><div></div></div> 58%
Total		53,017	233,842	551,770,799	118,487,281	73,966,375	85,191,066	38,201,119	478,881	272,304	349,546,447

Showing 61 to 70 of 248 entries



PanDA

The Integrated Grid

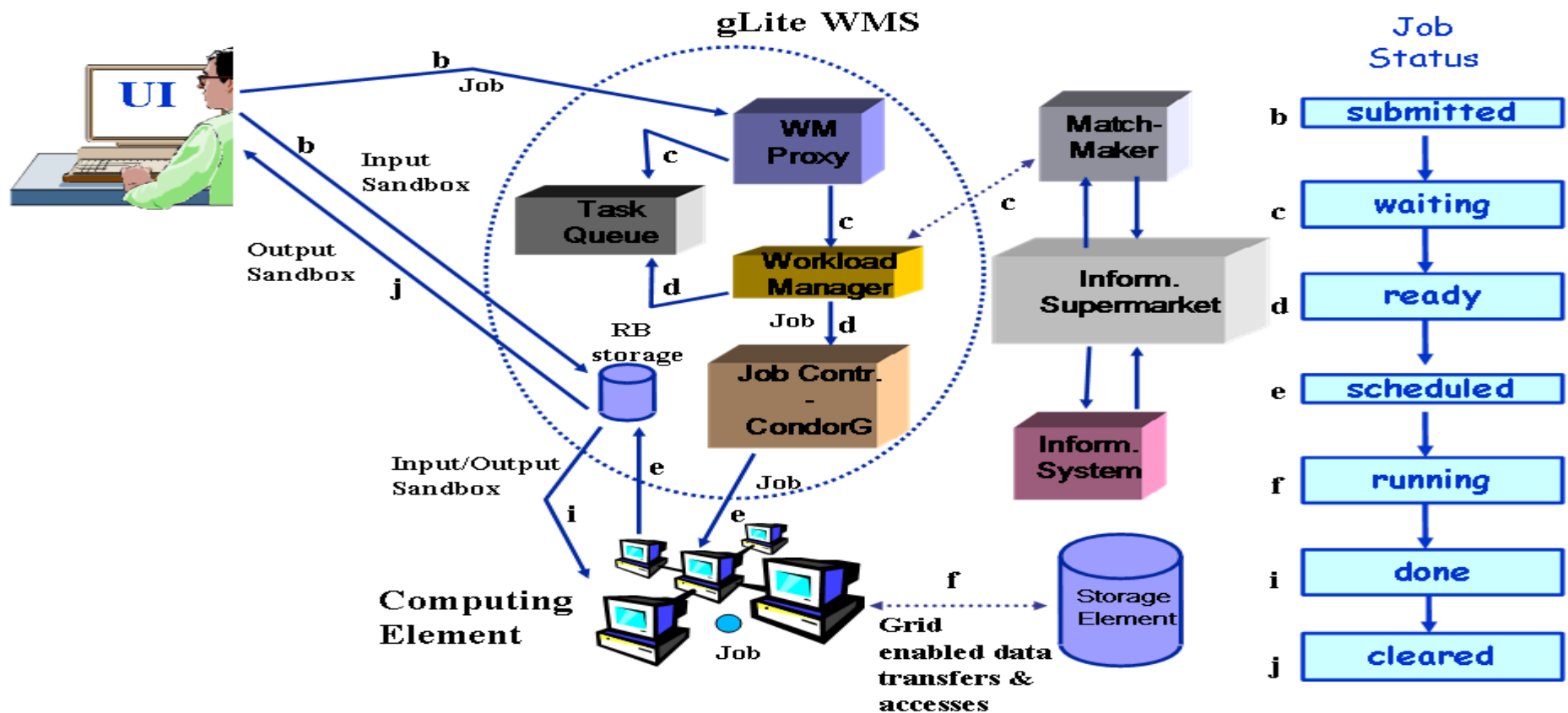
Computing System

with

Python

Work flow & Job management

- A typical Grid workflow



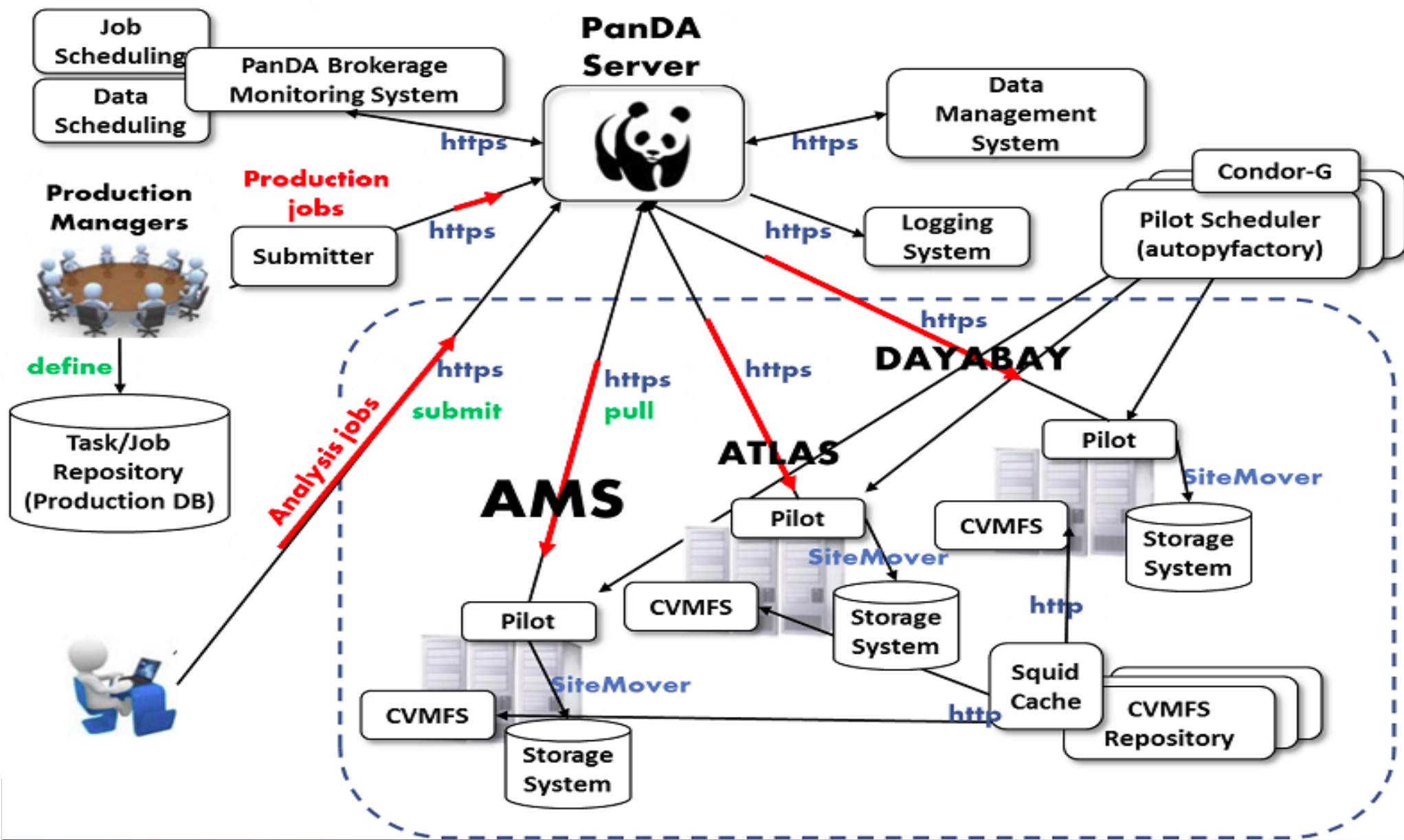


PanDA

- PanDA
 - Production and Distributed Analysis system.
 - Designed and developed by Atlas experiment.
 - It's data driven and pull model computing.
 - Including workflow, resource matchmaking and job management.
 - We are now working with Atlas to improve and deploy it for eScience users.

PanDA diagram

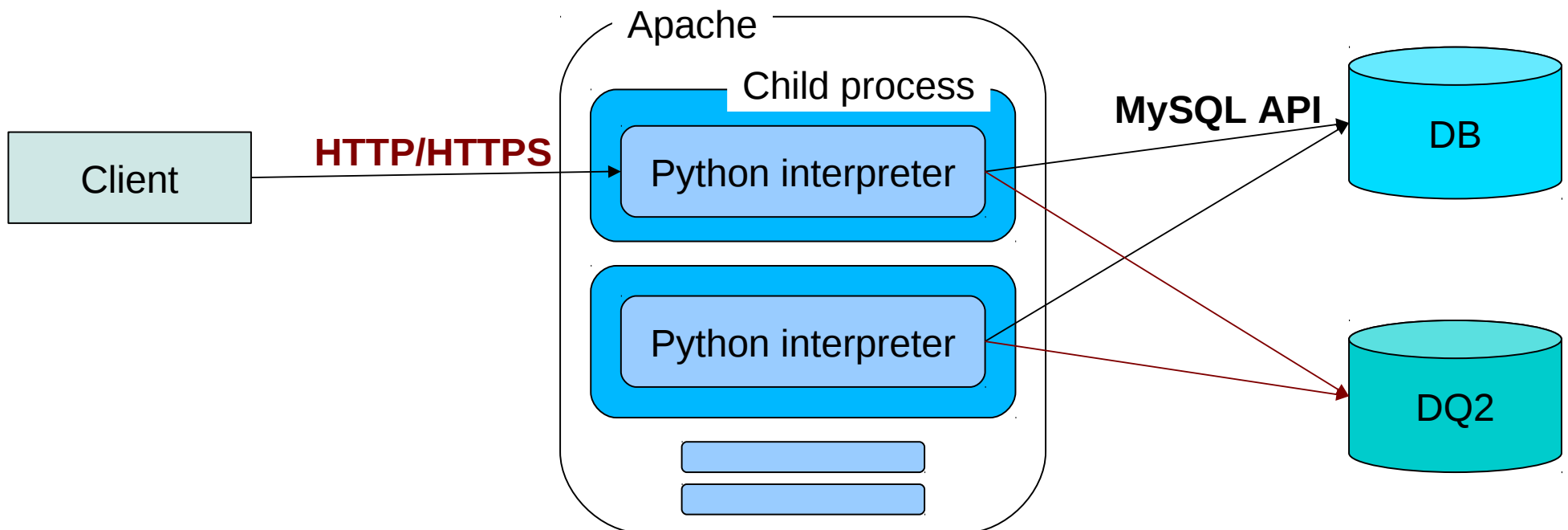
PanDA: Production and Distributed Analysis System





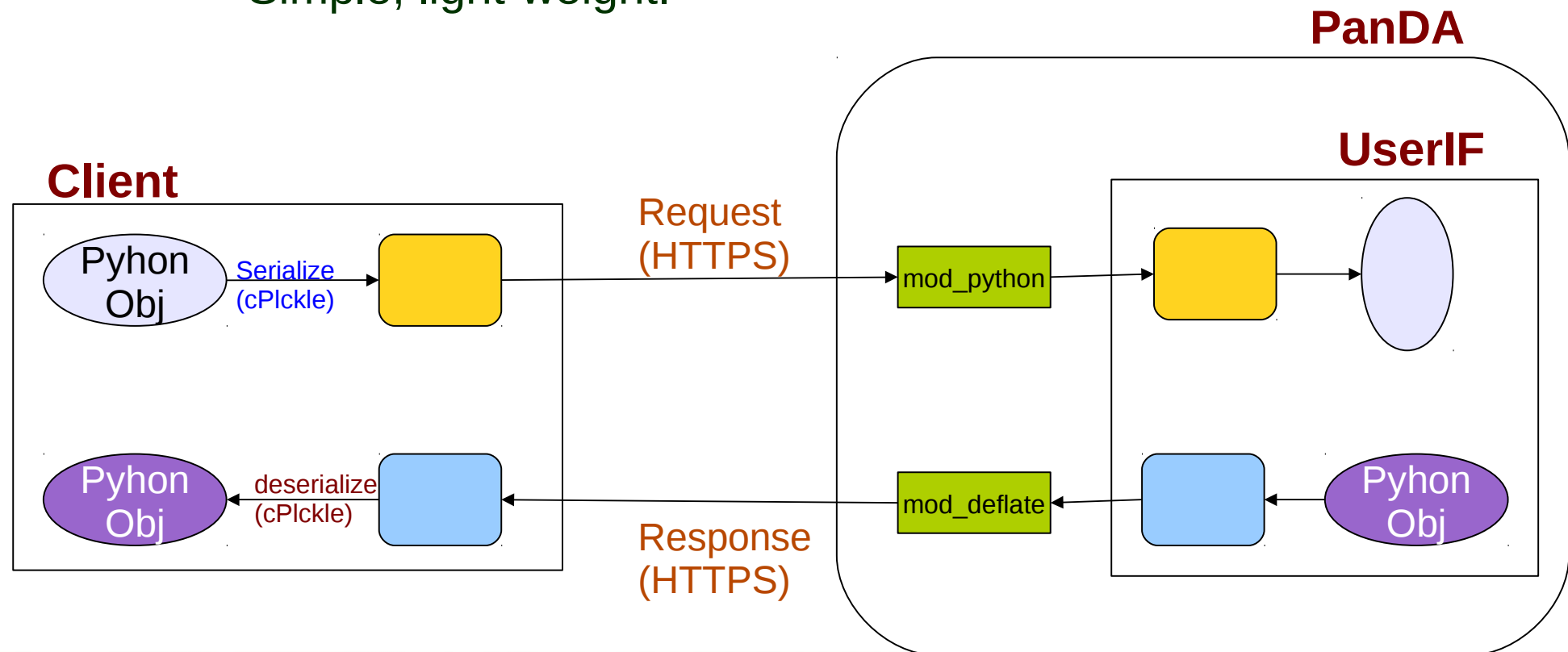
PanDA Server

- PanDA server design
 - Apache-based
 - Communication via HTTP/HTTPS
 - Multi-process
 - Global info in the memory resident database



PanDA Client

- PanDA client
 - Pickle module of python and native curl.
 - Client require python 2.3 or higher, curl and grid-proxy
 - Simple, light-weight.





PanDA screen shot

File Edit View History Bookmarks Tools Help

WLCGDB... Clouds ... Production... Indico [Ac... Ganglia: C... W Mean time... Google's ... DT DailyTech ... M Correlatio... Massive G... Research ... Anal > + v

panda.cern.ch/server/pandamon/query?dash=clouds#TW

Google

TW Tasks Jobs: assigned:1714 activated:12069 running:4628 holding:172 transferring:11949						
ANALY_TW_AMS	ANALY_TW_AMS	ANALY_TW_AMS	200			brokeroff
TW-FTT	ANALY_TAIWAN	ANALY_TAIWAN <i>set_queue_online</i>	5	6		online
	ANALY_TAIWAN_PNFS	ANALY_TAIWAN_PNFS	1			brokeroff
	ANALY_TAIWAN_TEST	ANALY_TAIWAN_TEST	20			brokeroff
	TW-FTT-f-cream01-atlas-pbs	TW-FTT	10			online
	TW-FTT-f-cream03-atlas-pbs	TW-FTT	10			online
	TW-FTT-AMS	TW-FTT-AMS	20			brokeroff
TW_AMS	TW_AMS	TW_AMS	20			brokeroff
TW_AMS_GO	TW_AMS_GO	TW_AMS_GO	5			brokeroff
Taiwan-LCG2	Taiwan-LCG2-cream01-atlas-pbs	Taiwan-LCG2	10			online
	Taiwan-LCG2-cream02-atlas-pbs	Taiwan-LCG2	10			online
	Taiwan-LCG2-cream03-atlas-pbs	Taiwan-LCG2	5			online
UK Tasks Jobs: assigned:38 waiting:10 activated:24780 running:17122 holding:272 transferring:7533						
ANALY_GLASGOW_GLEXEC	ANALY_GLASGOW_GLEXEC	ANALY_GLASGOW_GLEXEC	20			brokeroff
ANALY_OXFORD_GLEXEC	ANALY_OXFORD_GLEXEC	ANALY_OXFORD_GLEXEC	20			brokeroff
GLASGOWTESTGLEXEC	GLASGOWTESTGLEXEC	GLASGOWTESTGLEXEC <i>misconfig</i>	20			brokeroff
	ANALY_RAL	ANALY_RAL <i>HC.Blacklist.set.online</i>	5			online
	ANALY_RAL_XROOTD	ANALY_RAL_XROOTD <i>HC.Blacklist.set.online</i>	20			online