Dr. Paul Nilsson
University of Texas at Arlington

THE PANDA SYSTEM IN THE ATLAS EXPERIMENT

OUTLINE

- Introduction
- Design and implementation
- PanDA components
- Security
- Performance and system operations
- Distributed analysis
- Outlook and summary

INTRODUCTION

INTRODUCTION

- ***PanDA*** is the ***P*roduction** *an*d ***D*istributed *A*nalysis** system for the ATLAS Experiment
  - Designed to meet ATLAS requirements for a data-driven workload management system capable of operating at LHC data processing scale
- Single task queue and pilots
  - Apache based central server
  - Pilots retrieve jobs from server as soon as CPU is available (low latency)
- Highly automated
  - low operation manpower
  - integrated monitoring system
- Integrated with ATLAS Distributed Data Management system

A BRIEF HISTORY OF PANDA

PanDA development started in summer of 2005

Summer of 2005, **PanDA project** started

End of 2005, **US ATLAS** production and analysis

September 2006, workload management component of **US Open Science Grid**

October 2007, adopted by **ATLAS Collaboration** as sole system for distributed production

A BRIEF HISTORY OF PANDA

PanDA was initially developed for US based ATLAS production and analysis, and assumed that role in late 2005

Summer of 2005, **PanDA project** started

End of 2005, **US ATLAS** production and analysis

September 2006, workload management component of US Open Science Grid

October 2007, adopted by **ATLAS Collaboration** as sole system for distributed production

A BRIEF HISTORY OF PANDA

Since September 2006 PanDA has also been a principal component of the **US Open Science Grid** program in just-in-time (pilot-based) workload management

Summer of 2005, **PanDA project** started

End of 2005, **US ATLAS** production and analysis

September 2006, workload management component of **US Open Science Grid**

October 2007, adopted by **ATLAS Collaboration** as sole system for distributed production

A BRIEF HISTORY OF PANDA

In October 2007 PanDA was adopted by the ATLAS Collaboration as the sole system for distributed production across the Collaboration

Summer of 2005, **PanDA project** started

End of 2005, **US ATLAS** production and analysis

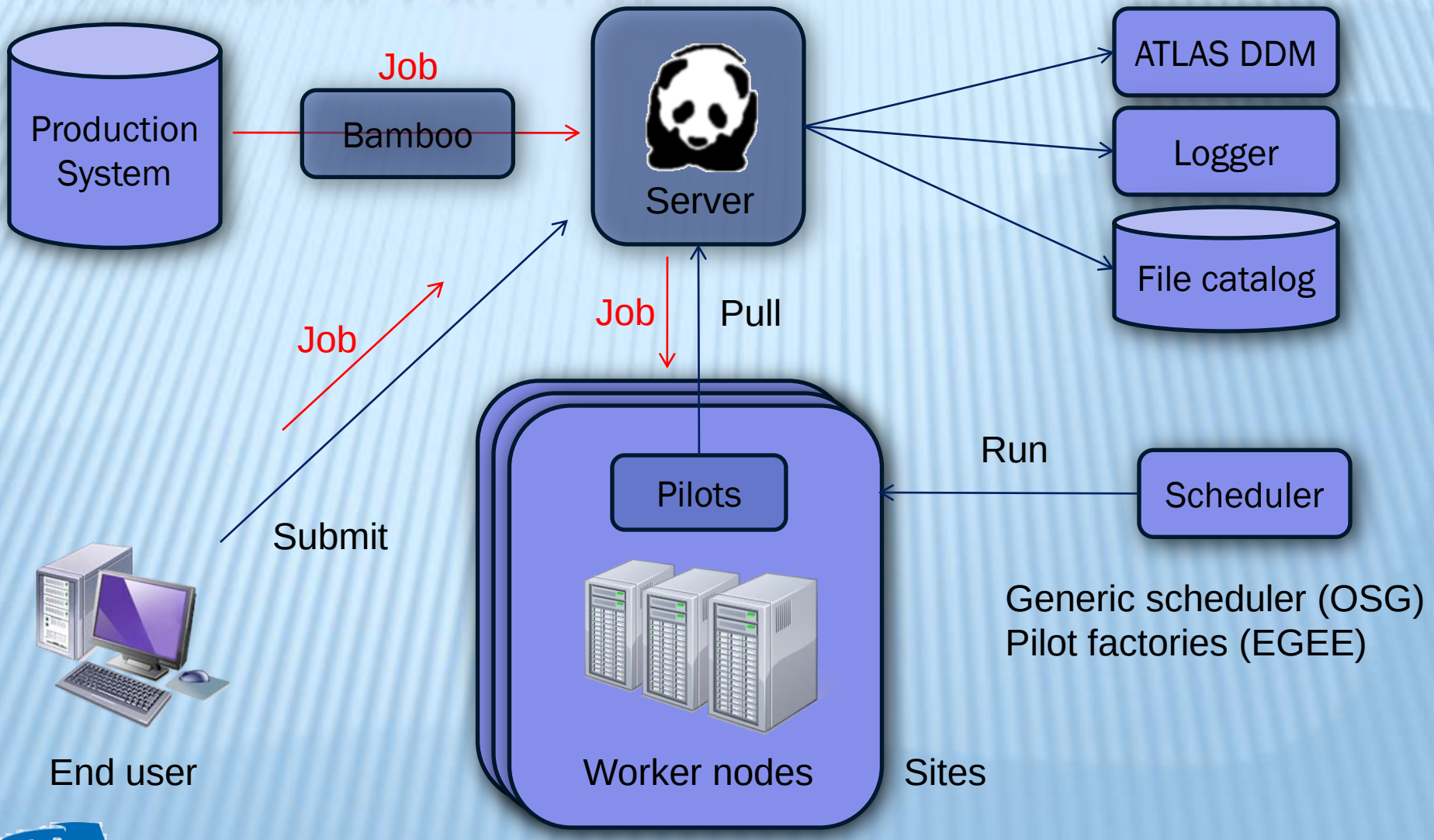
September 2006, workload management component of **US Open Science Grid**

October 2007, adopted by **ATLAS Collaboration** as sole system for distributed production

DESIGN AND IMPLEMENTATION

Design and Implementation

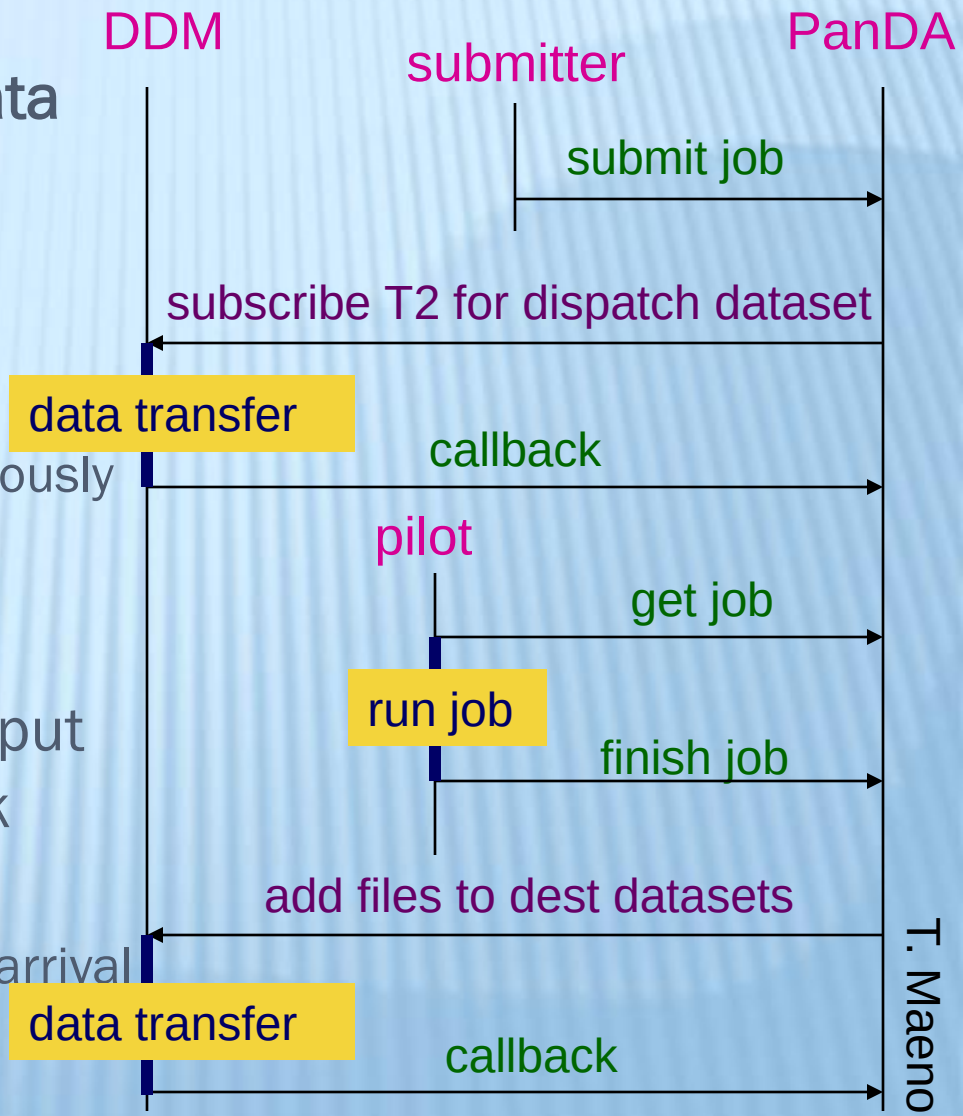
THE PANDA SYSTEM



Design and Implementation

DATA TRANSFER

- Relies on **ATLAS Distributed Data Management** (DDM)
  - PanDA sends request to DDM
  - DDM moves files and sends notifications to PanDA
  - PanDA and DDM work asynchronously
- Dispatch input files to T2 and aggregate output files to T1
- Jobs get 'activated' when all input files are copied, and pilots pick them up
  - Pilots don't have to wait for data arrival on WNs



PANDA COMPONENTS

PanDA components

CORE COMPONENTS

- *PanDA Server*: central hub composed of several components that make up the core of PanDA (task buffer, job dispatcher, etc)
- *PanDA DB*: MySQL database for PanDA (migration to Oracle in progress)
- *PanDA Client*: PanDA job submission and interaction client
- *Pilot*: execution environment for PanDA jobs. Pilots request and receive job payloads from the dispatcher, perform setup and cleanup work surrounding the job, and run the jobs themselves, regularly reporting status to PanDA during execution
- *AutoPilot*: Pilot submission, management and monitoring system
- *SchedConfig*: Database table used to configure resources
- *Monitor*: web based monitoring and browsing system that provides an interface to PanDA for operators and users
- *Logger*: logging system allowing PanDA components to log incidents in a database via the standard Python logging module
- *Bamboo*: interface between PanDA and the ATLAS production database

PanDA components

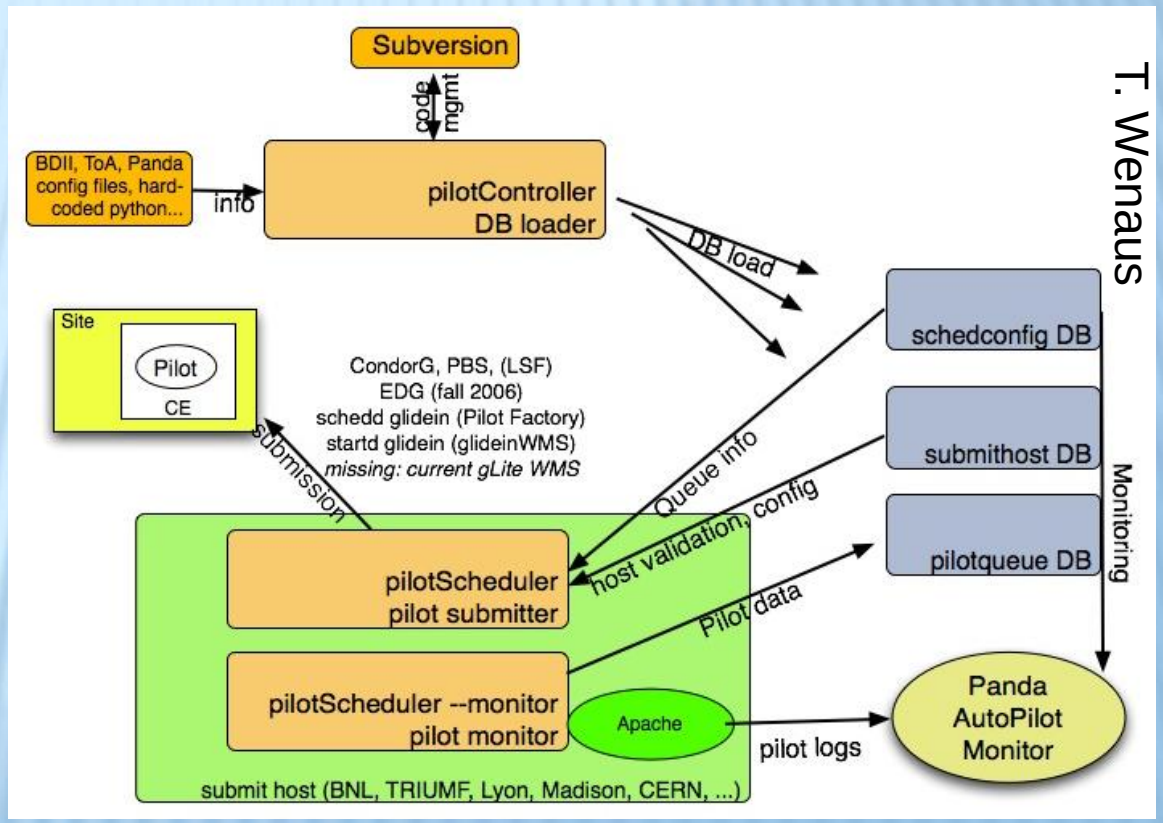
PANDA SERVER

- Implemented as a stateless multi-process REST web service with Apache mod_python and with a MySQL back-end
- Interaction with clients is via http (passive read operations) and https (active operations like job submission, pilot interaction). Secure https is authenticated using grid certificates, with mod_gridsite
- Composed of several components
  - *Task buffer* - PanDA job queue manager, keeps track of all active jobs in the system
  - *Brokerage* - matches job attributes with site and pilot attributes, manages the dispatch of input data to processing sites, and implements PanDA's data pre-placement requirement
  - *Job Dispatcher* - receives requests for jobs from pilots and dispatches job payloads; Jobs are assigned depending on how they match the capabilities of the site and worker node (data availability, disk space, memory etc)
  - *Data Service* – responsible for data dispatch to and retrieval from sites

PanDA components

AUTOPILOT

- "*AutoPilot*" is a simple and generic implementation of PanDA pilot and pilot-scheduler for use in more varied environments
- Governs submission of pilot jobs to target sites (using CondorG; another scheduler under development uses gLite job submission tools)
- Provides a pilot implementation that contains no ATLAS specific content
  - Used by *CHARMM*, a protein structure application that uses OSG VO



PanDA components

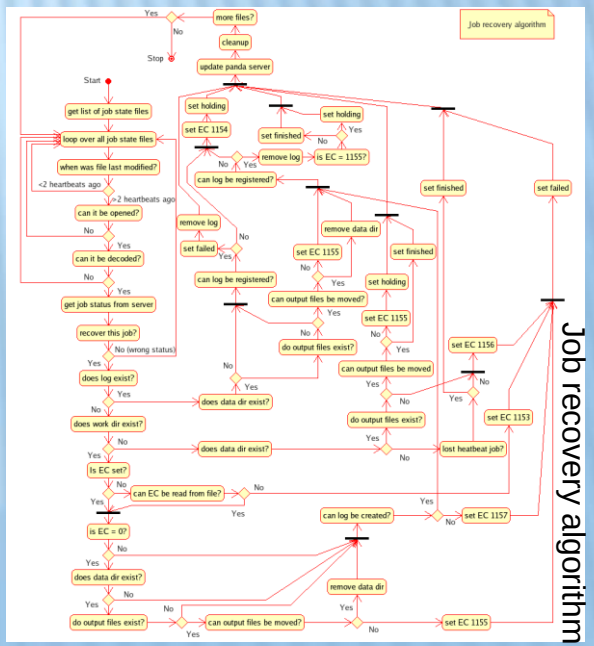
PILOT – FEATURES (1/2)

- **Pilot execution**
  - Pilot asks job dispatcher for a job, exits immediately if no real job is available
  - Pilot forks job wrapper for the real payload
  - Job wrapper prepares runtime environment, performs stage-in, executes payload, performs stage-out before wrapping up!
  - Pilot cleans up work dir after job is done (finished or failed) and saves the tarball of work dir to SE
- **Data transfer** of input/output and log files to/from local SE (Data already copied to site by DDM system)
  - File transfers using various copy tools; *cp, mv, dccp, rfcp, uberftp, gridftp, lcg-cp, xrdcp, dq2-get/put, globus-url-copy, and other tailor-made tools (necp, xcp, ..)*
    - *Different copy tools can be used for stage-in and stage-out*
    - *New copy tools can be added easily as plugins*
  - Direct reading support for dCache, Castor and Xrootd (i.e. no need for stage-in)
  - Time-outs for [hanging] transfers (supports multiple attempts)
  - Pilots can be centrally configured according to each site policy (so the pilot knows e.g. which copy tool to use, etc)

PanDA components

PILOT – FEATURES (2/2)

- Monitoring
  - Job monitor checking if job is still alive - kills sub processes if no output is produced within limit
  - Sends heartbeats to PanDA server every 30 minutes
    - Server will fail job if no heartbeats in 6 hours
    - Can kill a job if desired by user (message sent through heartbeat communication channel)
  - Checks size of user work dir and remaining local disk space
- Job recovery
  - Depending on site policy, a job that fails to transfer its output can leave its work dir (or move to external disk) which will be found by a later pilot that will re-attempt the transfer, file registration, cleanup, or whatever is necessary to finish the job



Job recovery algorithm

SECURITY

Security

PANDA SECURITY (1/2)

- PanDA services use the standard GSI grid security model of authentication and authorization based on X509 grid certificates, implemented via OpenSSL and https
- Interactions with PanDA require secure https (for anything other than passive read actions)
- Proxy's VOMS attributes are checked to ensure user is a member of a VO authorized for PanDA use
- Authenticated users' DN is part of the metadata of a PanDA job, so user id is known and tracked throughout PanDA operations
- Production job execution and file management relies on production certificates, with the pilot carrying a specific 'pilot' VOMS role to control its rights

PANDA SECURITY (2/2)

- Analysis jobs also run under a production proxy unless gLExec is employed in identity switching mode
  - Optional gLExec based identity change on WN to submitter identity for user jobs under testing (proxy management done by MyProxy)
  - ATLAS has performed all necessary preparations (including discussing security issues of its model of job execution within the wLCG working group).
    - gLExec is considered mature
- Proxy can be concealed by pilot during payload execution
- We have also defined a means of securing the job workload specification (transformation) from tampering in the PanDA DB (encryption of the transformation using RSA key pair, with decoding/validation in the pilot prior to execution)

PERFORMANCE AND SYSTEM OPERATIONS

PERFORMANCE

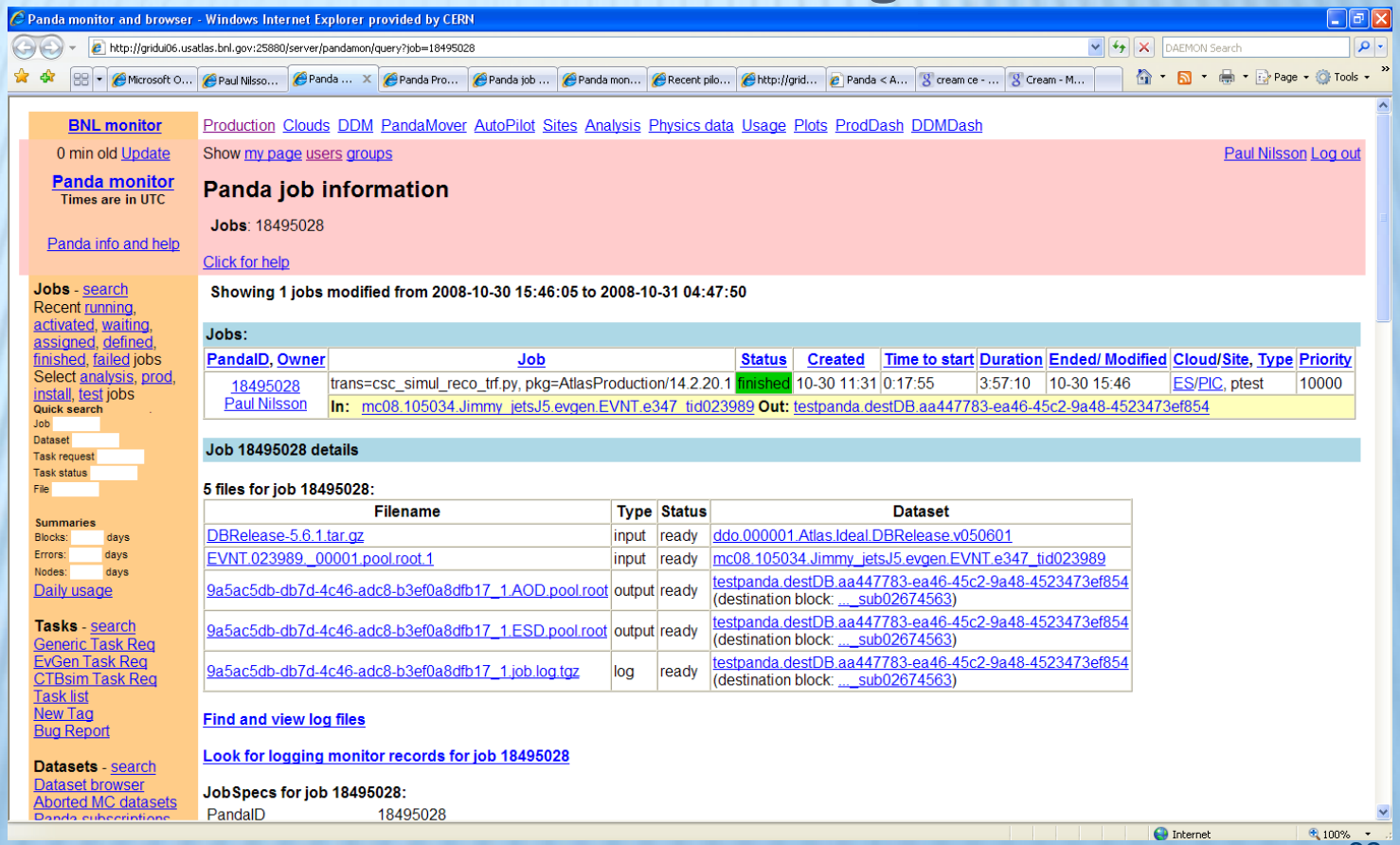
- 18+ million jobs as of Nov 2008, now at a rate of about 500k jobs/week for production at 100 sites around the world
- ~10k jobs/week for analysis, 3M analysis jobs by almost 400 users in last 6 months (673 users in total)
- As ATLAS data taking ramps up over the next few years, job counts are estimated to reach on the order of 500k jobs/day, with the greatest increase coming from analysis.

Performance and system operations

MONITORING

- Web based monitor, e.g. providing fast access to user jobs, output files, log files
- Multiple service instances for load-balancing

Example of a user job



Performance and system operations

PANDA DASHBOARD

Panda monitor

Shift log Wiki

Panda Production Operations Dashboard

Panda shift guide calendar mailing list

Click for help

Jobs - search
Recent running, activated, waiting, assigned, defined, finished, failed jobs
Select analysis, prod, install, test jobs

Quick search
Job
Dataset
Task req
Task status
File

Summaries
Blocks: days
Errors: days
Nodes: days
Daily usage

Tasks - search
Generic Task Req
EvGen Task Req
CTBsim Task Req
Task list
Task browser
Bug Report

Datasets - search
Dataset browser
Aborted MC datasets
Panda subscriptions

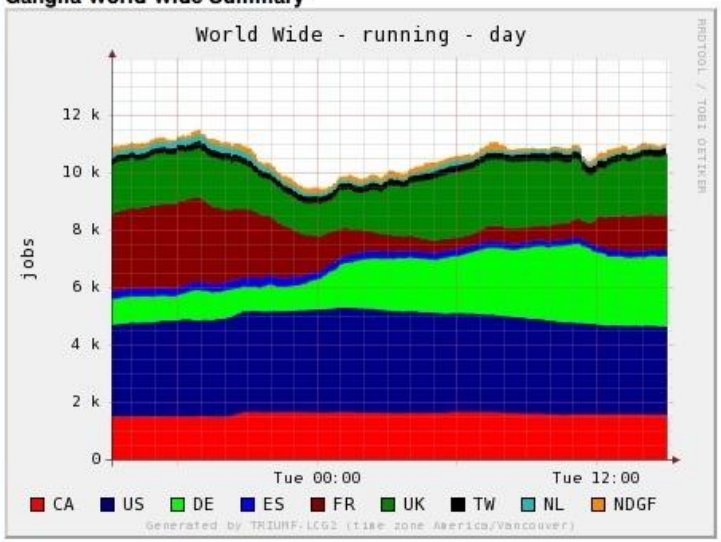
Datasets Distribution
DDM Req
Req list
AODs
EVNTs
RDOs
Conditions DS
DB Releases
Validation Samples

Servers: BNL:OK BNLdev:OK CERN:OK Logger:OK

Active tasks: CA:14 DE:10 ES:3 FR:13 NL:4 TW:3 UK:15 US:31
Bamboo submissions, status over last 12 hours

Jobs updated >12 hrs ago: activated:6528 running:none
Jobs updated >36 hrs ago: transferring:366

Ganglia World Wide Summary



Space available at sites:

Site	GB	As of (UTC)
US		
ANALY_MWT2-condor	53743	04-08 22:01
ANALY_SLAC-lsf	25241	04-08 21:10
BU_ATLAS_Tier2	23777	04-08 21:28
BU_ATLAS_Tier2o	23774	04-08 21:56
IU_OSG	58622	04-08 22:01
MWT2_IU	58622	04-08 22:02
MWT2_UC	53743	04-08 22:02
OU_OCHEP_SWT2-condor	5839	04-08 21:47
SLACXRD-lsf	25240	04-08 21:29
UC_ATLAS_MWT2	53744	04-08 21:58
UTA_SWT2	2830	04-08 22:01
Other SEs reporting in last 3 days		

Production job summary, last 12 hours (Details: errors, nodes)

Cloud Information	Nodes	Jobs	Latest	Pilots (3hrs)	defined	assigned	waiting	activated	running	holding	transferring	finished	failed	tot	trf	other
Overall Production	6541	67253	04-08 18:03	4864	0 / 0	1829 / 0	1 / 0	22964 / 0	10960 / 0	6047 / 0	8347 / 366	15489 / 0	1617 / 0	9%	0%	9%

Performance and system operations

PRODUCTION JOB SUMMARY

Production job summary, last 12 hours (Details: errors, nodes)

Cloud Information	Nodes	Jobs	Latest	Pilots (3hrs)	defined	assigned	waiting	activated	running	holding	transferring	finished	failed tot trf other
Overall Production	6541	67253	04-08 18:03	4864	0 / 0	1829 / 0	1 / 0	22964 / 0	10960 / 0	6047 / 0	8347 / 366	15489 / 0	1617 / 0	9% 0% 9%
CA	779	8021	04-08 18:03	399	0	402	0	3227	1559	147	800 / 0	1603	283	15% 0% 15%
DE	1103	13400	04-08 18:03	634	0	31	0	5350	2464	64	2324 / 0	2531	636	20% 0% 20%
ES	225	1561	04-08 18:02	198	0	10	0	681	204	81	216 / 0	359	10	3% 0% 3%
FR	1201	6082	04-08 18:03	1994	0	806	0	1573	1299	157	396 / 1	1797	54	3% 0% 3%
IT	0	0	0	0	0	0	0	0	0	0	0 / 0	0	0	
NL	231	2161	04-08 17:55	223	0	241	0	519	39	0	663 / 365	329	370	53% 0% 53%
UK	1649	17404	04-08 18:03	862	0	318	1	3784	2079	4887	1856 / 0	4327	153	3% 0% 3%
US	1170	17055	04-08 18:03	516	0	12	0	6951	3064	709	2036 / 0	4172	111	3% 1% 2%
TW	183	1569	04-08 18:03	38	0	9	0	879	252	2	56 / 0	371	0	0% 0% 0%

ERROR REPORTING

Panda job error summary for last 12 hours (0.5 days)

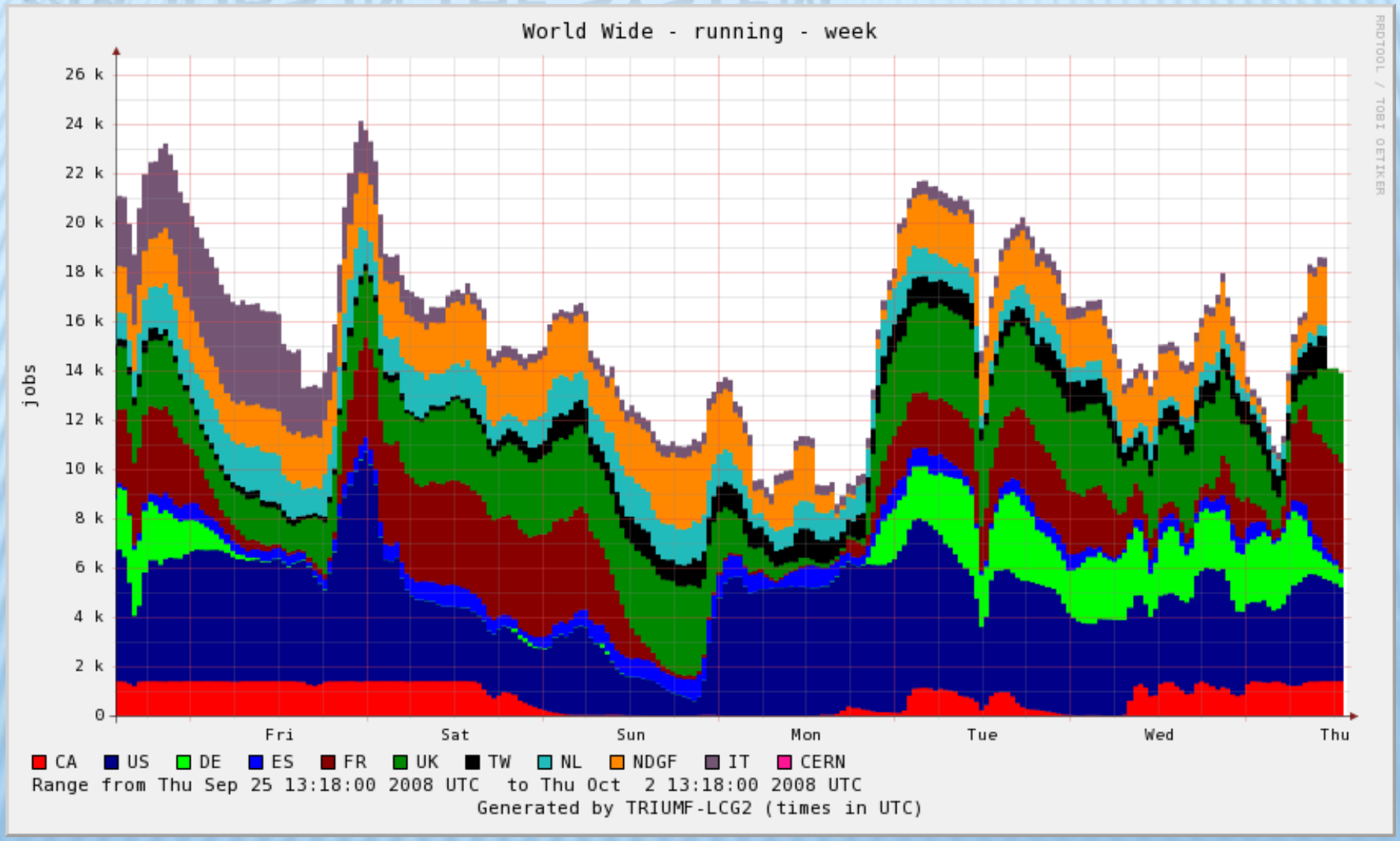
Managed production jobs only. Show **production**, analysis, test, all jobs/CEs

Job wall time: 37303 hrs Error losses: trans: 732 (2.0%) panda: 1586 (4.3%) ddm: 6552 (17.6%) other: 279 (0.7%)

Error type (type count)	Count	CPU-hrs	Latest	Code:	Description
All	defined:300 assigned:1213 waiting:5 activated:3541 sent:1 running:2617 holding:3091 transferring:6880 finished:15608 failed:8035 (34.0%)				
ddmErrorCode (1)	1	4.4	10-31 00:28	200:	Could not add output files to dataset
exeErrorCode (57)	2	8.9	10-30 09:43	1132:	LRC registration error (consult log file)
exeErrorCode (57)	2	10.8	10-29 23:20	1137:	Put error: Error in copying the file from job workdir to localSE
exeErrorCode (57)	3	9.0	10-29 04:12	1155:	Failed to move output files for lost job
exeErrorCode (57)	29	7.1	10-31 09:28	60010:	segmentation fault
exeErrorCode (57)	10	17.0	10-31 09:44	61200:	ServiceManager Unable to initialize Service
exeErrorCode (57)	4	1.6	10-31 01:13	64100:	Transform output file errors
exeErrorCode (57)	7	82.3	10-31 08:20	69999:	Unknown Transform error
jobDispatcherErrorCode (570)	548	748.6	10-31 03:15	100:	Lost heartbeat
jobDispatcherErrorCode (570)	4	12.5	10-29 05:17	101:	Job recovery failed for three days
jobDispatcherErrorCode (570)	18	0.0	10-31 08:04	102:	No reply to sent job
pilotErrorCode (7367)	4840	712.6	10-31 09:49	1099:	Get error: Staging input file failed
pilotErrorCode (7367)	1	0.1	10-31 03:50	1103:	Get error: No such file or directory
pilotErrorCode (7367)	5	5.7	10-31 04:09	1111:	Exception caught by runJob
pilotErrorCode (7367)	1	0.9	10-31 01:13	1112:	Exception caught by pilot
pilotErrorCode (7367)	16	0.5	10-31 00:12	1113:	Get error: Failed to import LFC python module
pilotErrorCode (7367)	1	6.9	10-31 06:58	1114:	Put error: Failed to import LFC python module
pilotErrorCode (7367)	1221	1704.1	10-31 09:06	1137:	Put error: Error in copying the file from job workdir to localSE
pilotErrorCode (7367)	45	2.2	10-31 09:31	1145:	Get error: md5sum mismatch on input file
pilotErrorCode (7367)	5	181.5	10-31 08:59	1150:	Looping job killed by pilot
pilotErrorCode (7367)	733	3501.8	10-31 09:03	1151:	Get error: Input file staging timed out
pilotErrorCode (7367)	45	72.3	10-31 05:46	1152:	Put error: File copy timed out

Performance and system operations

24K JOBS IN THE SYSTEM



DISTRIBUTED ANALYSIS

Distributed analysis

PRODUCTION VS ANALYSIS

- Run on same infrastructure
  - Same software, monitoring system and facilities
  - No duplicated manpower for maintenance
- Separate computing resources
  - Different queues → different CPU clusters
  - Production and analysis don't have to compete with each other
- Different policies for data transfer
  - Analysis jobs don't trigger data-transfer
    - Jobs go to sites which hold the input files
  - For production, input files are dispatched to T2s and output files are aggregated to T1 via DDM asynchronously

Distributed analysis

PANDA MONITOR USER INFO

Show my page **users** groups

Paul Nilsson Log out

Users

673 users
3.1M PanDA jobs in last 6 months
220 users with > 1k jobs
68 users with > 10k jobs

Users: 673 in last 3 days: 62 7: 72 30:131 90:271 180:386
Usage in last 6 months: Job count: 3125468 Users with >1000 jobs: 220 >10k jobs: 68
Listed by most recent usage

User	Jobs	Latest	Sites used	Job types run	Groups
Vladimir Savinov	3664	2008-10-27 07:07	ANALY_BNL_ATLAS_1 (3044) ANALY_LONG_BNL_ATLAS (620)	user (3578) panda (86)	all atlas usatlas
Christian Ohm	2992	2008-10-27 07:00	ANALY_BNL_ATLAS_1 (2568) ANALY_LONG_BNL_ATLAS (424)	user (2970) panda (22)	all atlas
Hannah DeBerg	6932	2008-10-27 06:59	ANALY_BNL_ATLAS_1 (6360) ANALY_LONG_BNL_ATLAS (568) ANALY_MWT2 (4)	user (6344) panda (588)	all atlas usatlas
Pavel Jez	1710	2008-10-27 06:40	ANALY_BNL_ATLAS_1 (852) ANALY_LONG_BNL_ATLAS (858)	user (1702) panda (8)	all atlas
Johannes Elmsheuser	5592	2008-10-27 06:25	ANALY_LONG_LYON (16) ANALY_BNL_ATLAS_1 (754) ANALY_TRIUMF (128) ANALY_OU_OCHEP_SWT2 (592) ANALY_SFU (124) ANALY_LYON (102) ANALY_SWT2_CPB (524) ANALY_TOKYO (120) ANALY_ALBERTA (120) ANALY_GRIF-IRFU (120) ANALY_GRIF-LAL (114) ANALY_LPC (116) ANALY_LAP	panda (2804) user (2788)	all atlas
Ana Damjanovic	11862	2008-10-27 06:14	CHARMM (11570) TESTCHARMM (292)	test (11862)	all atlas usatlas

Distributed analysis

USER TOOLS

- *pAthena* – client tool for PanDA
  - Submit user-defined jobs from the command line
  - A consistent user-interface to Athena users
  - Works on ATHENA runtime environment
- *GANGA* (equipped with a PanDA backend)
  - Graphical frontend for job definition and management, developed for ATLAS and LHCb
  - Built-in support for configuring and running applications based on the Gaudi / Athena framework

OUTLOOK AND SUMMARY

NEAR TERM PLANS

- Migration of PanDA DBs from MySQL to Oracle (in progress)
- Migration of PanDA central services from BNL to CERN (after Oracle migration)
- Deployment of gLExec-enabled pilot for analysis under user proxy rather than production proxy
  - Once gLExec is available on EGEE CEs
- Extension of analysis support to ROOT (in addition to ATLAS offline framework)
- Support for pilot submission to CREAM CEs using CondorG
  - To access all resources (even those exposing only CREAM I/F)
  - Depends on CondorG support in CREAM, in progress
- Ongoing extensions of production workflow automation
  - Intelligent production task brokering among clouds with associated data movement (via ATLAS DDM)
- Pilot scheduler security extensions providing pilot authentication via secure tokens

SUMMARY

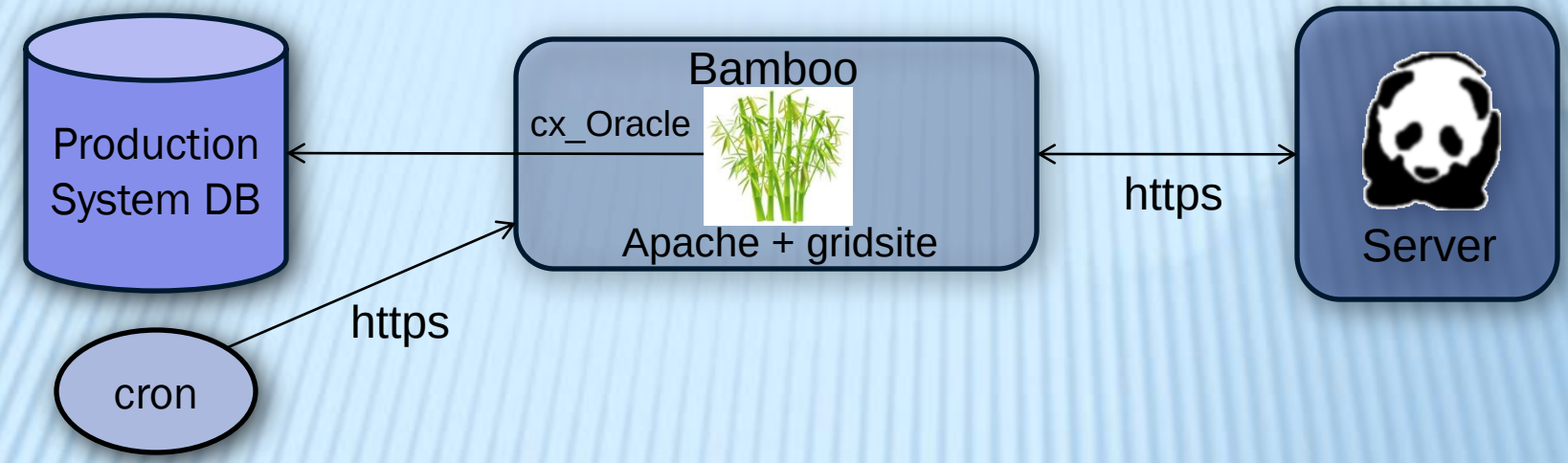
- PanDA production across ATLAS is going very smoothly
  - No performance issues, plenty of scaling headroom
  - High volume MC production
  - Huge computing resources available for individual analysis
- Ready to provide stable and robust service for data-taking

MORE INFORMATION

- PanDA homepage
  - https://twiki.cern.ch/twiki/bin/view/Atlas/Panda
- PanDA monitor
  - http://services.atlascomp.org/?redirect=pandamon
- Open Science Grid
  - http://www.opensciencegrid.org
- ATLAS Experiment
  - http://atlas.web.cern.ch/Atlas/index.html

PanDA components

BAMBOO



T. Maeno

- Get jobs from production DB and submit them to PanDA
- Update job status in production DB
- Assign tasks to clouds dynamically
- Kill jobs set to be aborted
- A cron job triggers the procedures every ten minutes

U.S. PRODUCTION FOR 2008