COMPASS
ProdSys Overview

Artem Petrosyan
28.11.2018, LIT, JINR

COMPASS collaboration

Common **M**uon and **P**roton **A**pparatus for **S**tructure and **S**pectroscopy

24 institutions from 13 countries
– nearly 250 physicists

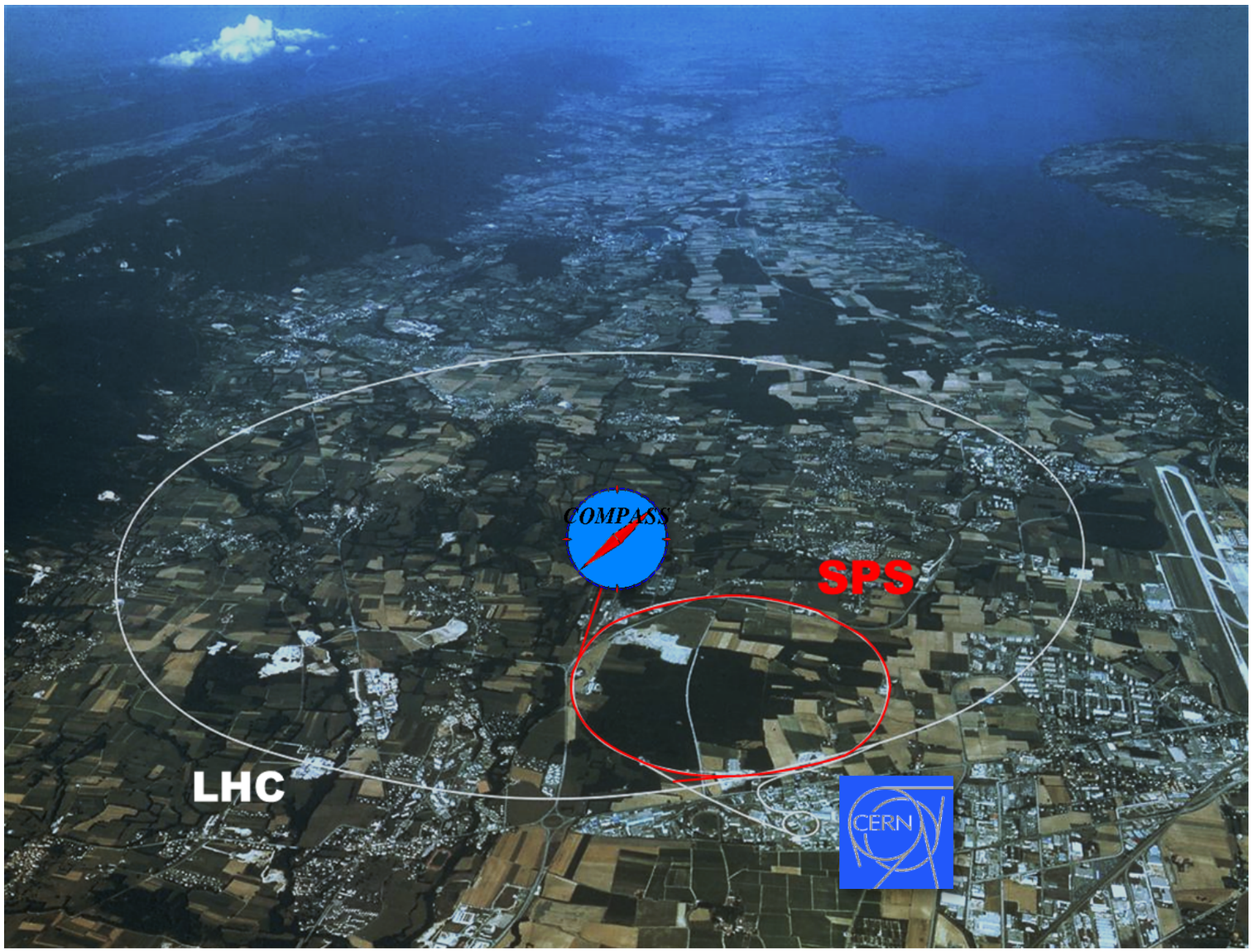
- **CERN SPS north area**
- **Fixed target experiment**
- **Approved in 1997 (20 years)**
- **Taking data since 2002**

Wide physics program

COMPASS-I
- **Data taking 2002-2011**
- **Muon and hadron beams**
- **Nucleon spin structure**
- **Spectroscopy**

COMPASS-II
- **Data taking 2012-2018 (2021?)**
- **Primakoff**
- **DVCS (GPD+SIDIS)**
- **Polarized Drell-Yan**
- **Transverse deuteron SIDIS**

Many "beyond 2021" ideas

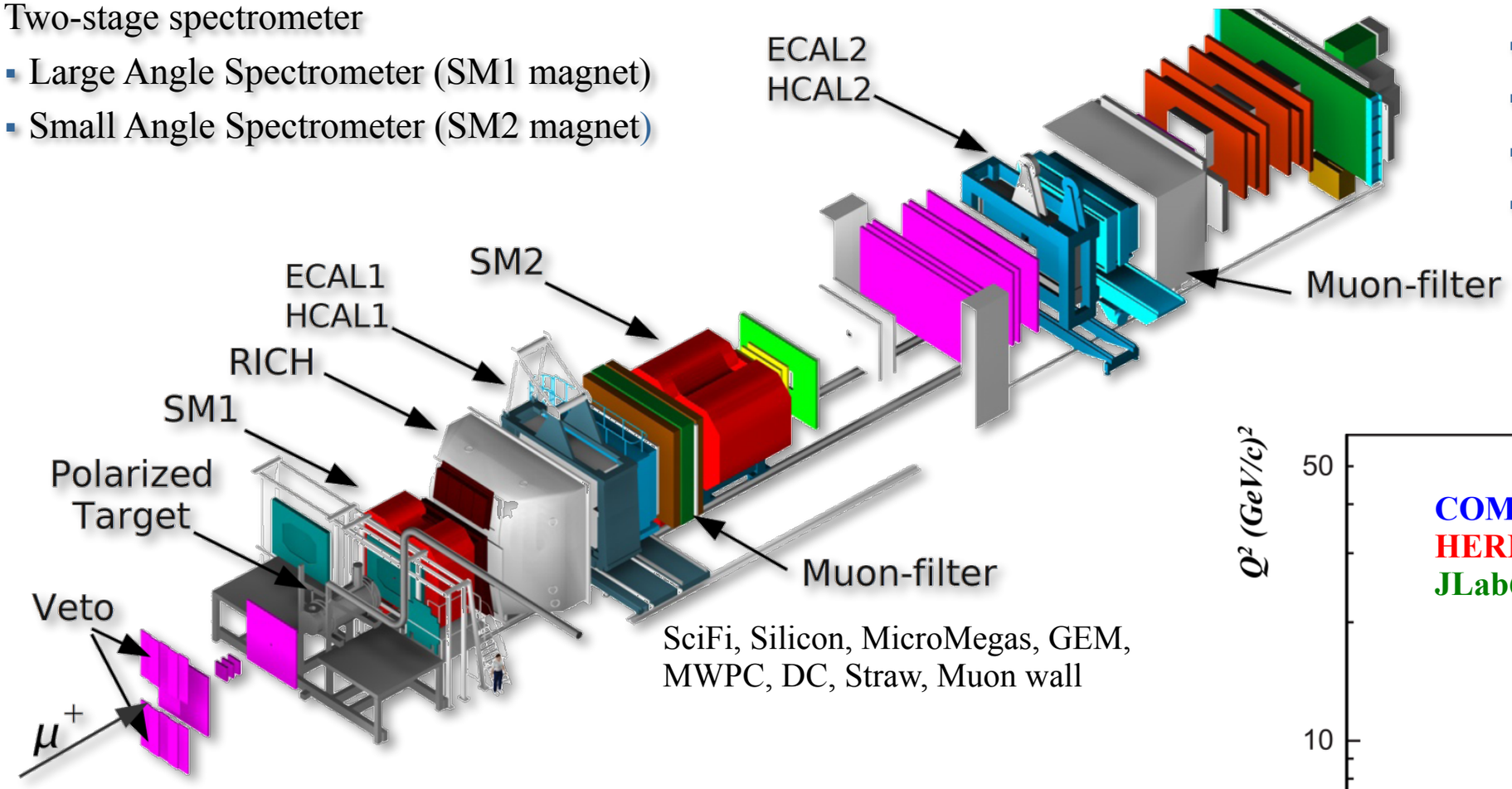


COMPASS web page: http://wwwcompass.cern.ch

COMPASS experimental setup: Phase I (muon program)

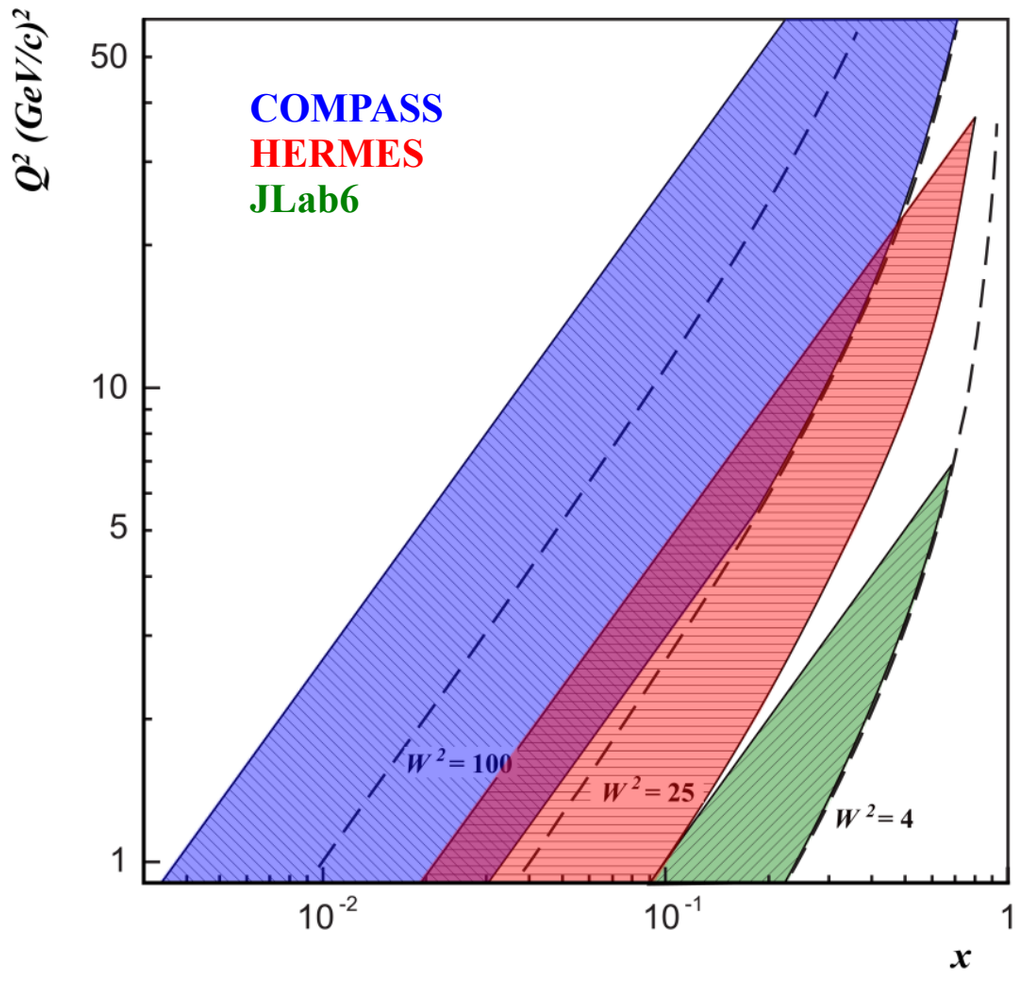
Two-stage spectrometer

- Large Angle Spectrometer (SM1 magnet)
- Small Angle Spectrometer (SM2 magnet)



SciFi, Silicon, MicroMegas, GEM, MWPC, DC, Straw, Muon wall

- High energy beam
- Large angular acceptance
- Broad kinematical range
- Momentum, tracking and calorimetric measurements, PID



Data-taking years: 2002-2011

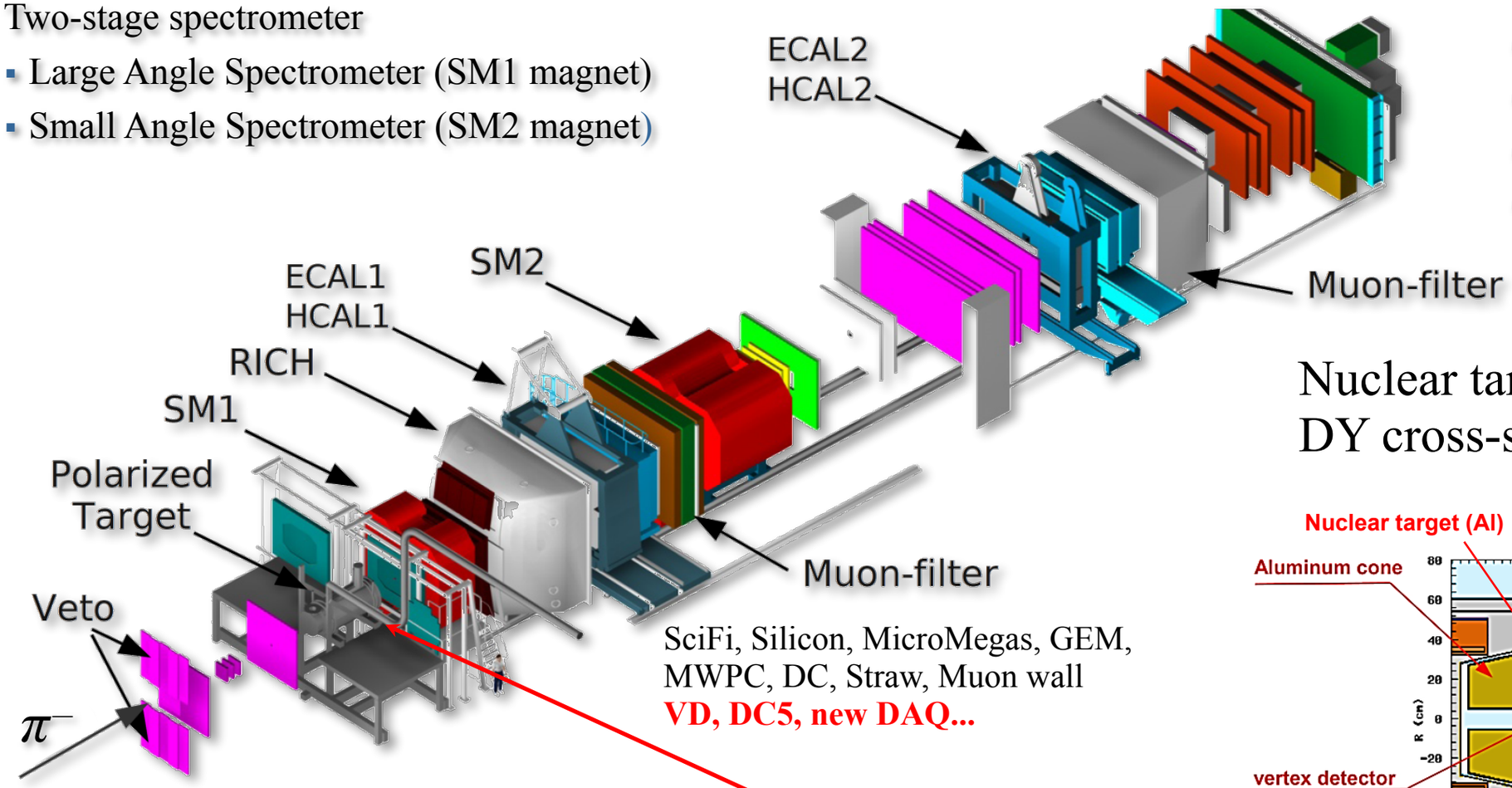
Longitudinally polarized (80%) μ^+ beam:
Energy: 160/200 GeV/c, Intensity: $2 \cdot 10^8$ μ^+/spill (4.8s).

Target: Solid state *(^{6}LiD or NH_3)*

- ^{6}LiD 2-cell configuration. Polarization (L & T) ~ 50%, f ~ 0.38
- NH_3 3-cell configuration. Polarization (L & T) ~ 80%, f ~ 0.14

COMPASS experimental setup: Phase II (DY program)

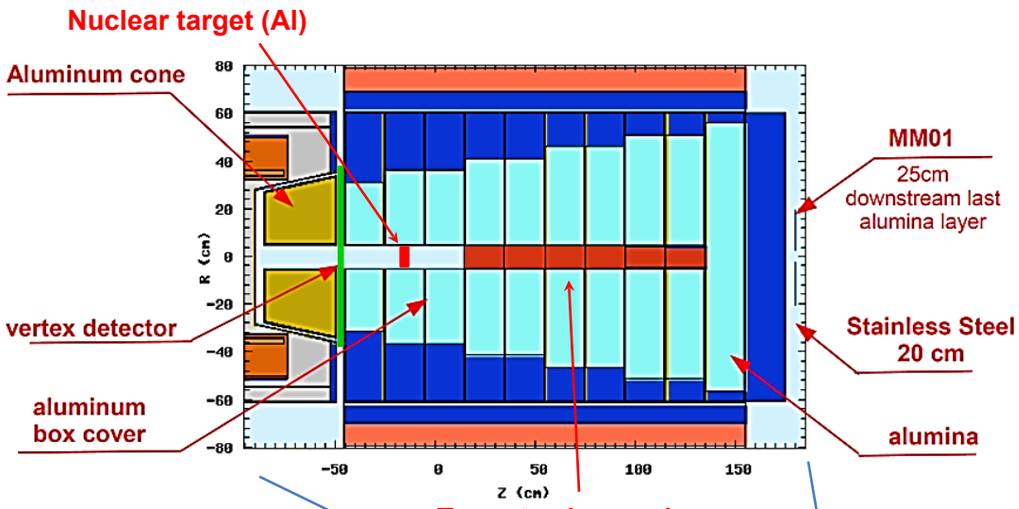
Two-stage spectrometer
- Large Angle Spectrometer (SM1 magnet)
- Small Angle Spectrometer (SM2 magnet)

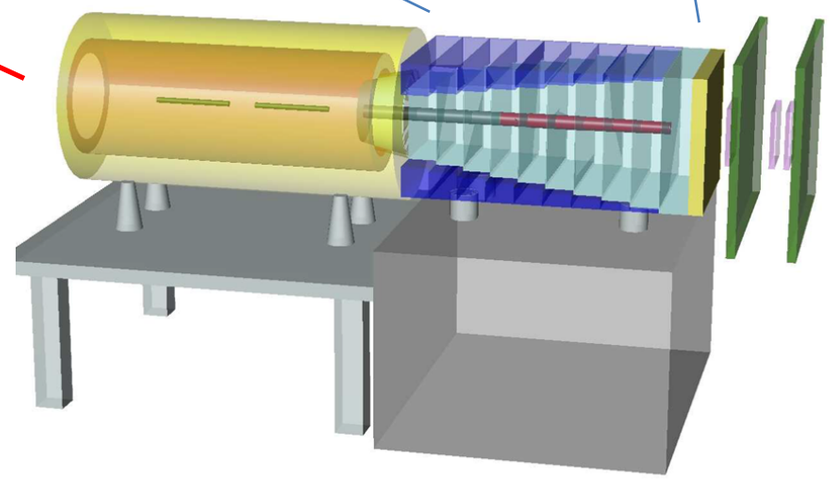


SciFi, Silicon, MicroMegas, GEM, MWPC, DC, Straw, Muon wall
VD, DC5, new DAQ...

- High energy beam
- Large angular acceptance
- Broad kinematical range
- Momentum, tracking and calorimetric measurements, PID

Nuclear targets → unpolarized DY, DY cross-sections, EMC effect





Data-taking years: 2014 (test) 2015 and 2018

High energy π^- beam:

Energy: 190 GeV/c, Intensity: 10^8 π/s

Target: Solid state

- NH_3 2-cell configuration. Polarization T ~ 73%, f ~ 0.18
- Data is collected simultaneously with both target spin orientations
  Periodic polarization reversal to minimize systematic effects

Raw data

2001 - 13 TB
2002 - 196
2003 - 230
2004 - 496
2006 - 390
2007 - 912
2008 - 523
2009 - 1223
2010 - 1740
2011 - 518
2012 - 878
2015 - 801
2016 - 571
2017 - 1391
2018 - 1450

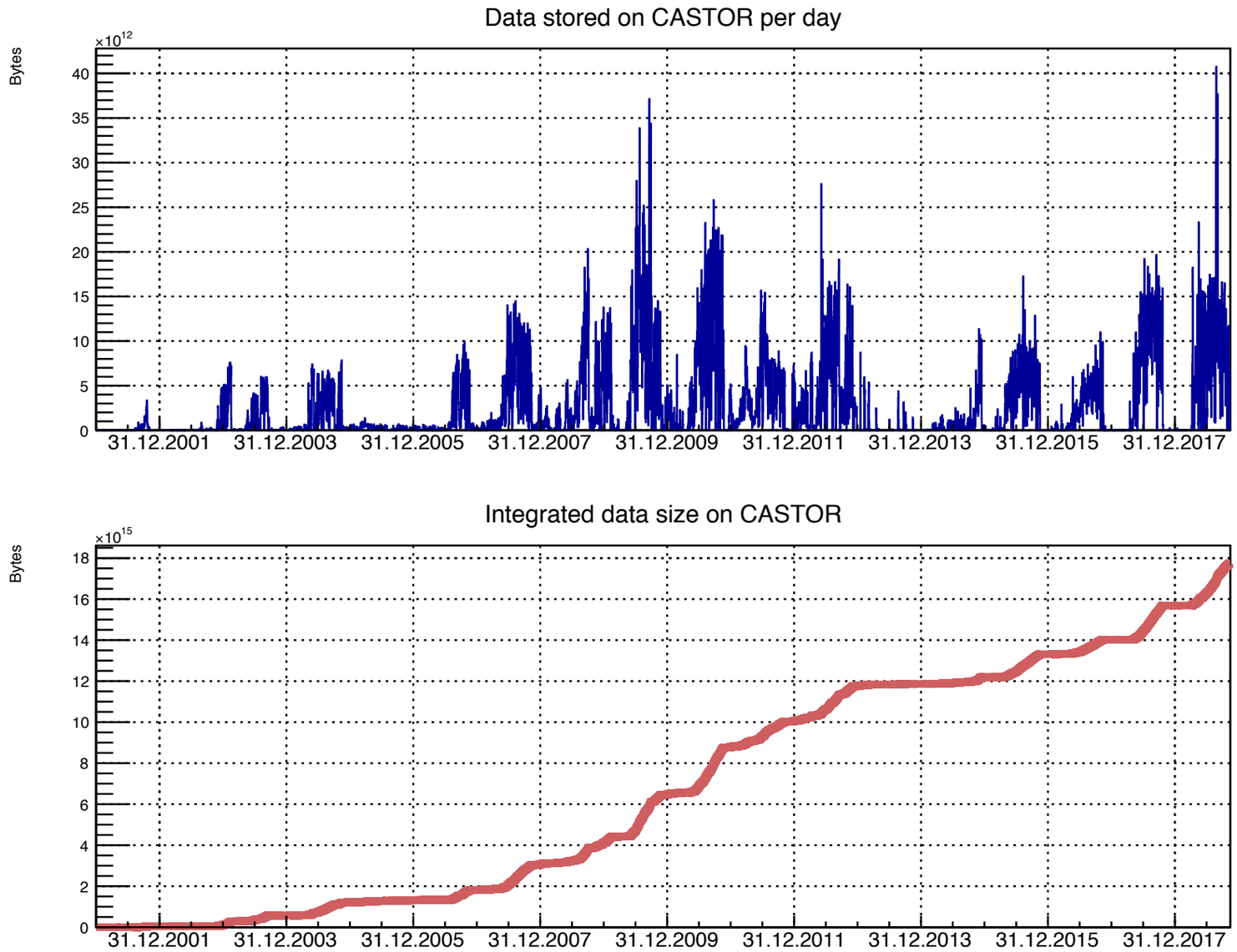
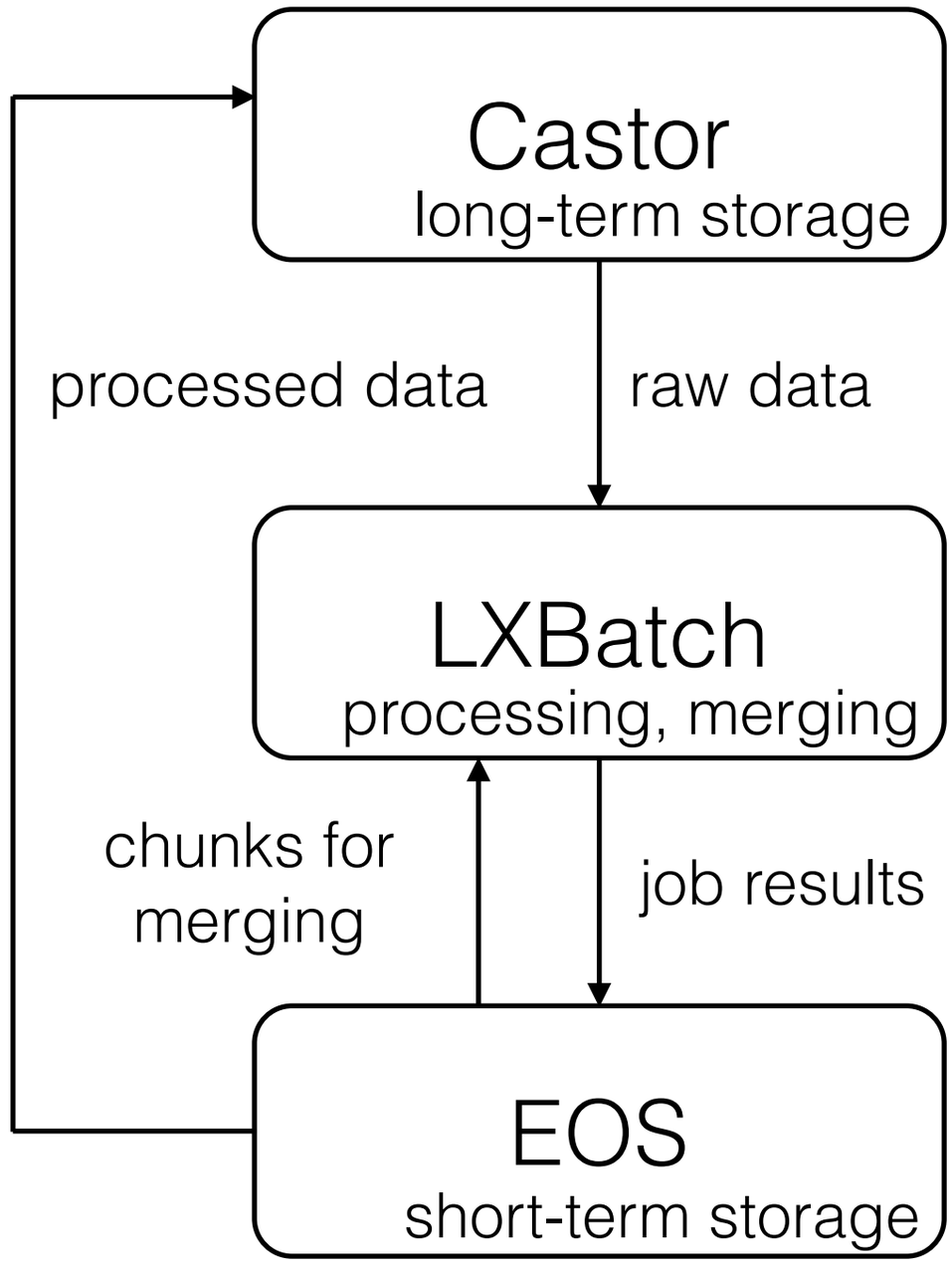


Chart by Sergey Gerasimov

Processing on Grid

Work flow in 2015

- Raw data are stored on Castor (CERN Advanced STORrage manager)
- Data is being requested to be copied from tapes to disks before processing
- Task moves files from Castor to LXBatch for processing
- After processing results are being transferred to EOS for merging or short-term storage or directly to Castor for long-term storage
- Merging, cross checking
- Results are being copied to Castor for long-term storage
- All routines executed under production account at lxplus and use bash commands
- Processing rate: ~**9000** jobs simultaneously



ProdSys components

1. Task requests layer: data files, runs list
2. Job definition layer: automatic
3. Job execution layer: direct submission to LXBatch
4. Workflow management: automatic
5. Data management: automatic
6. Monitoring: summary web page, shell scripts

Data catalog

- Raw and processed files are stored on Castor
- Raw data catalog in Oracle
  - Naming convention: year/period/run/chunk
- ProdSys database as a catalog of processed data
  - Naming convention: year/period/production/run-chunk-processing options

Disk-less computing sites model can be used

ProdSys redesign motivation in 2015

- Change computing site from LSF, which will be decommissioned by the end of 2018, to Condor

    - Even more: get ability to switch computing sites, get more resources, any type, not only LSF

        - Even more: build a system which is able to run jobs on HPCs

- Remove strict connectivity to AFS, which will be replaced by EOS FUSE

- Remove strong connection to Castor, which will be replaced by CTA

We need a Workload Management System

What is WMS?

- WMS — Workload Management System
- Providing a central queue for all users, makes hundreds of distributed sites appear as local
- Hides middleware while supporting diversity and evolution
    - WMS interacts with middleware, users see only high level workflow
    - Automation engines built in WMS, not exposed to users
- Hides variations in infrastructure
    - WMS presents uniform ‘job’ slots to user
    - Easy to integrate grid sites, clouds, HPC sites
- Uses the same system for simulation, data processing and users analysis
- Similar ideas have been implemented in several independent systems developed by LHC experiments: AliEn, Dirac, PanDA

WMS evolution

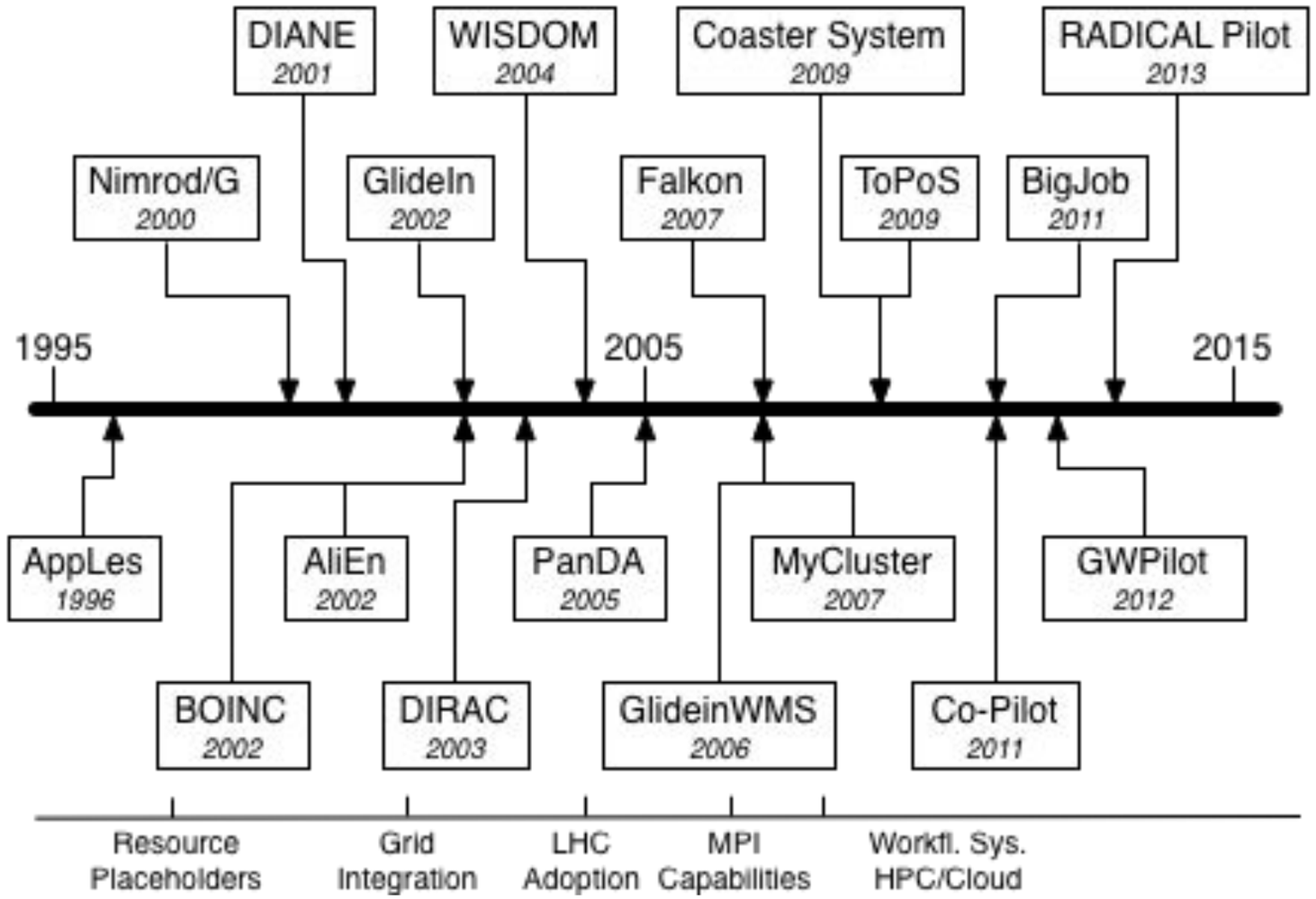


Diagram by Shantenu Jha

Why PanDA?

- The PanDA Production and Distributed Analysis System has been developed by ATLAS to meet requirements of data-driven workload management system for production and distributed analysis processing capable at LHC data processing scale

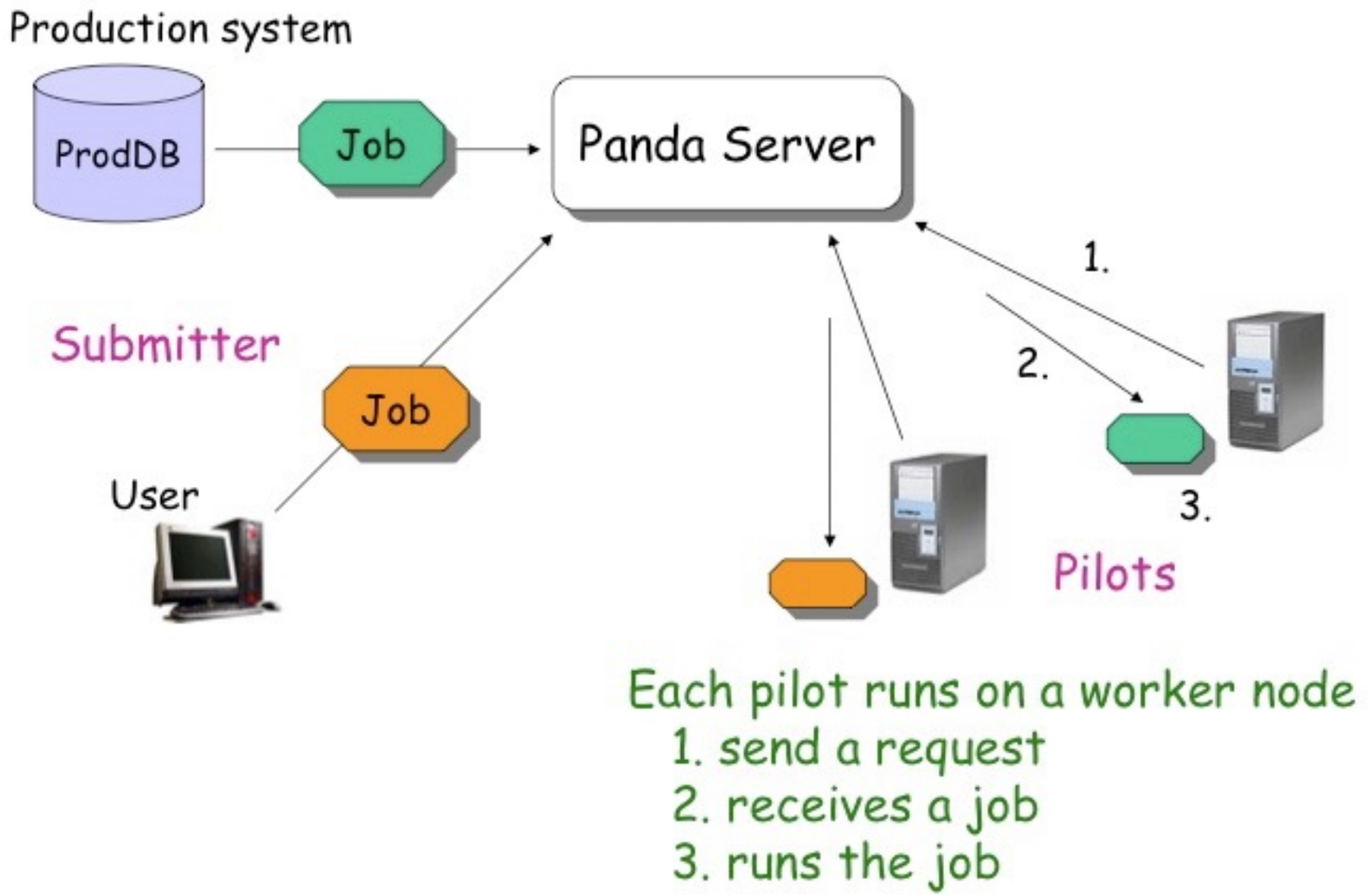
- PanDA manages both user analysis and production jobs via same interface

- PanDA processing rate is 250-300K jobs on ~200 sites every day

- The PanDA ATLAS analysis user community numbers over 1400

- Supports classic Grid computing resources, clouds, HPCs

PanDA job workflow



Schema by Tadashi Maeno

Steps to be done to enable processing through PanDA

- PanDA (DB, Server, APF) instance installation
- Grid environment setup
- COMPASS logic implementation in Pilot code
- Production chain workflow and data flow management software reimplementation
- PanDA monitoring adaptation for COMPASS

No need to use a distributed data management

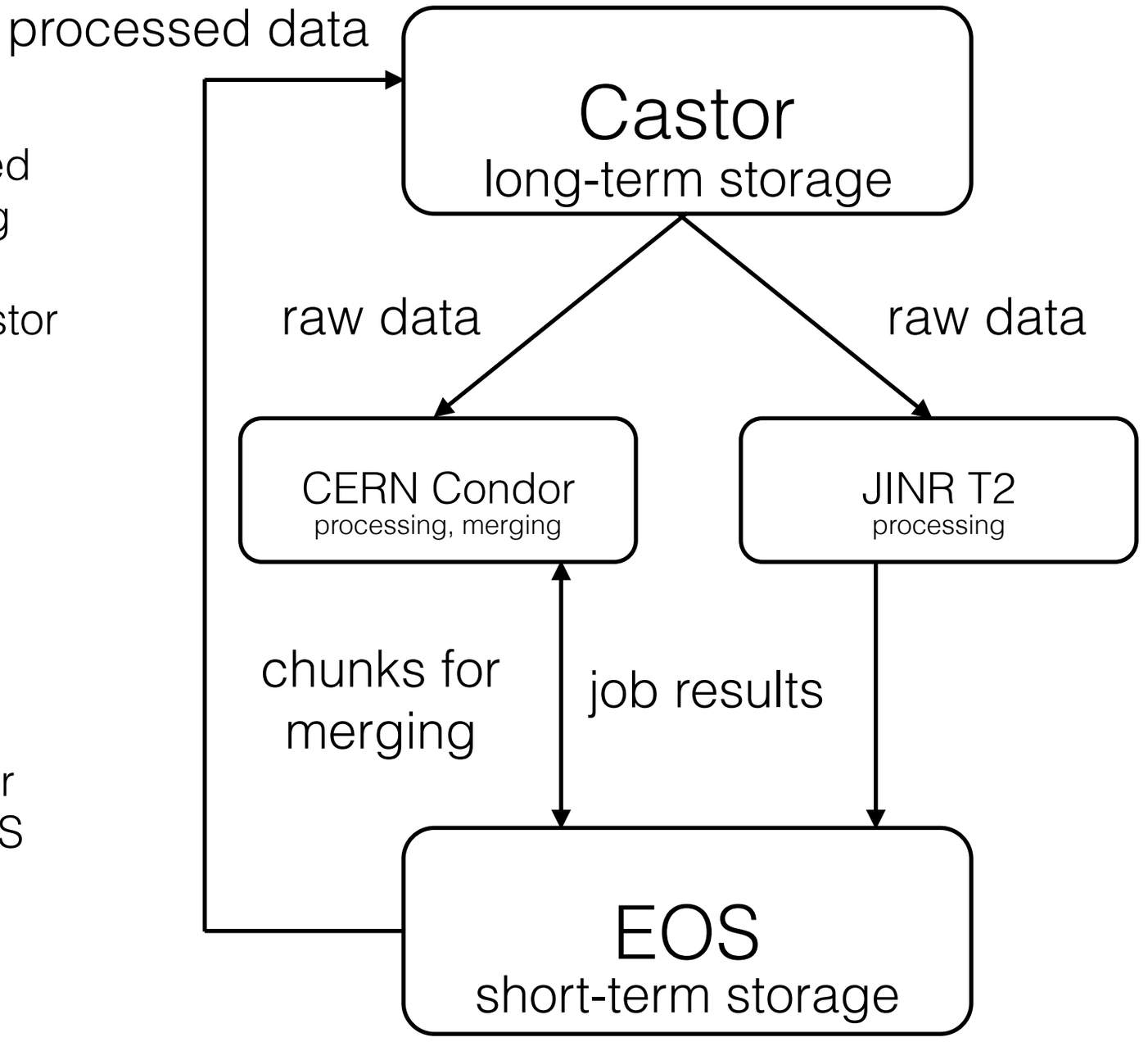
Grid environment

- AFS COMPASS group
  - Production account
- Local batch queue
- EOS directory
- AFS directory to deploy production software

- Virtual organisation
  - Production role
- Computing elements
- EOS storage element
- CVMFS

Work flow in 2017

- Raw data are stored on Castor
- Files are being requested to be copied from tapes to disks before processing
- Task moves files via XRootD from Castor to CERN Condor
- After processing results are being transferred to EOS for merging and short-term storage via XRootD
- Merging is done on CERN Condor
- Results are being copied to Castor for long-term storage via XRootD and FTS
- All management routines work using X509 proxy authentication



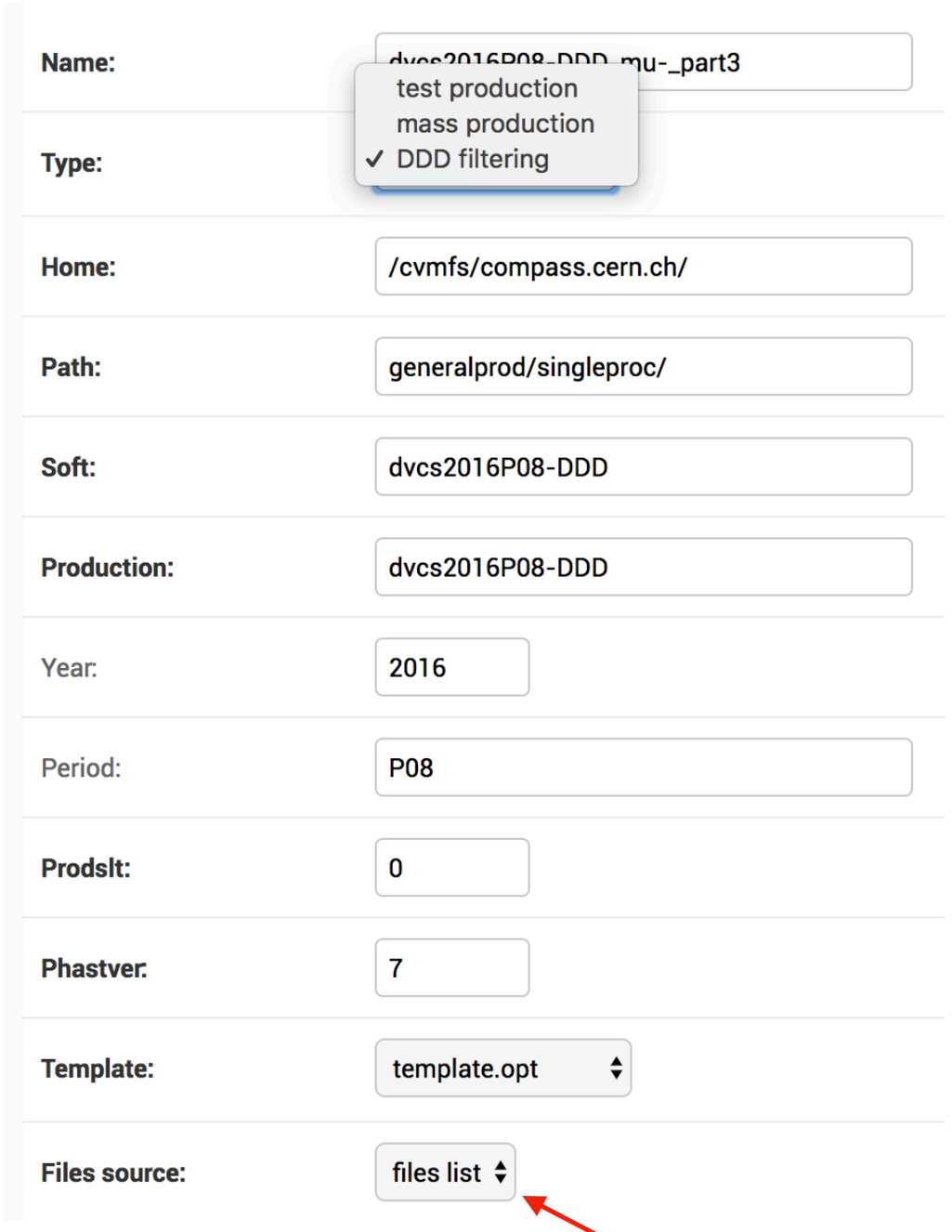
ProdSys components

1. Task requests layer: Web UI
2. Job definition layer: automatic
3. Job execution layer: PanDA
4. Workflow management: automatic
5. Data management: automatic, Web UI
6. Monitoring: a set of Web interfaces

1. Task requests layer

Web UI:

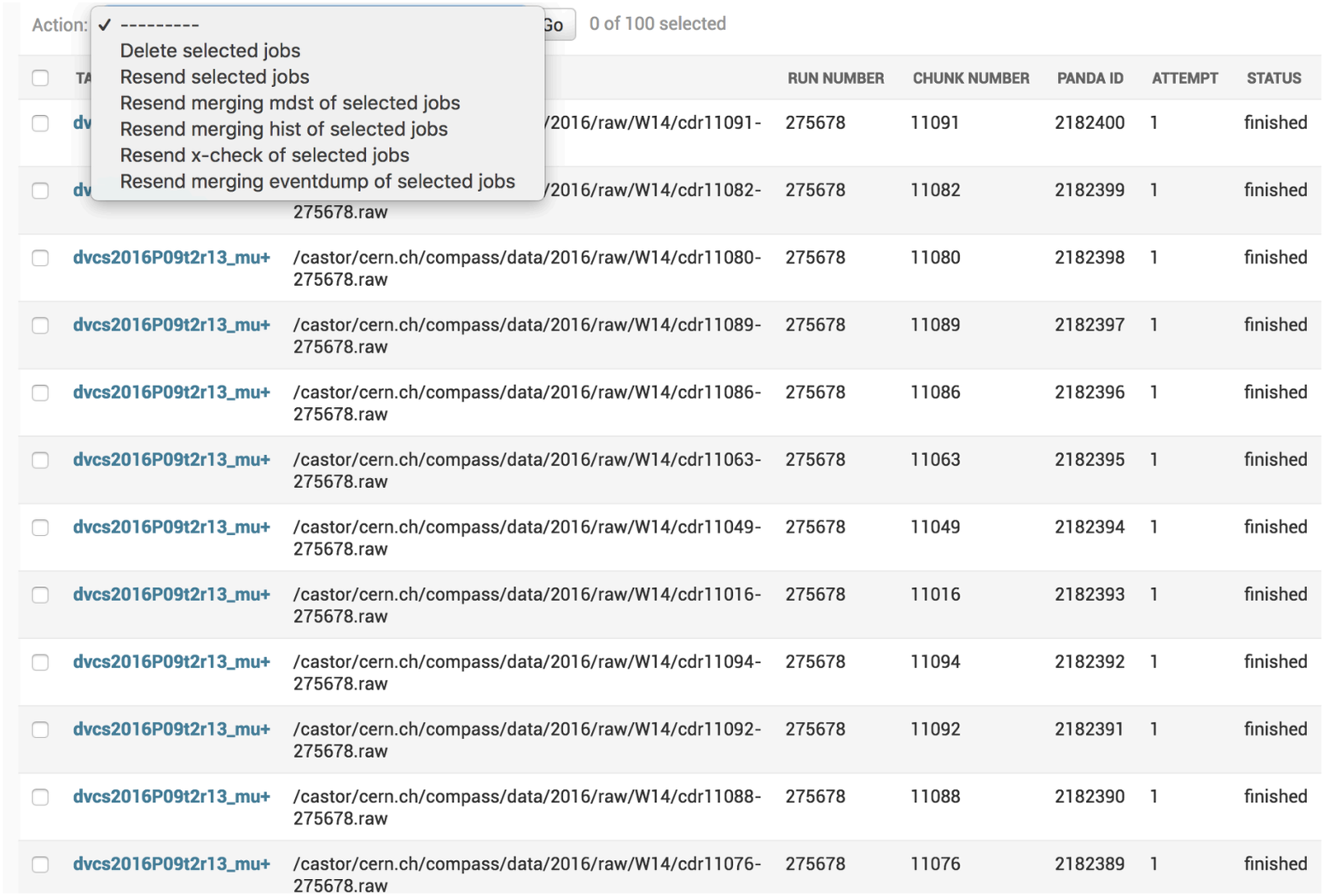
- execution parameters
- paths
- version of software
- list of chunks or runs
- etc.



2. Job definition layer

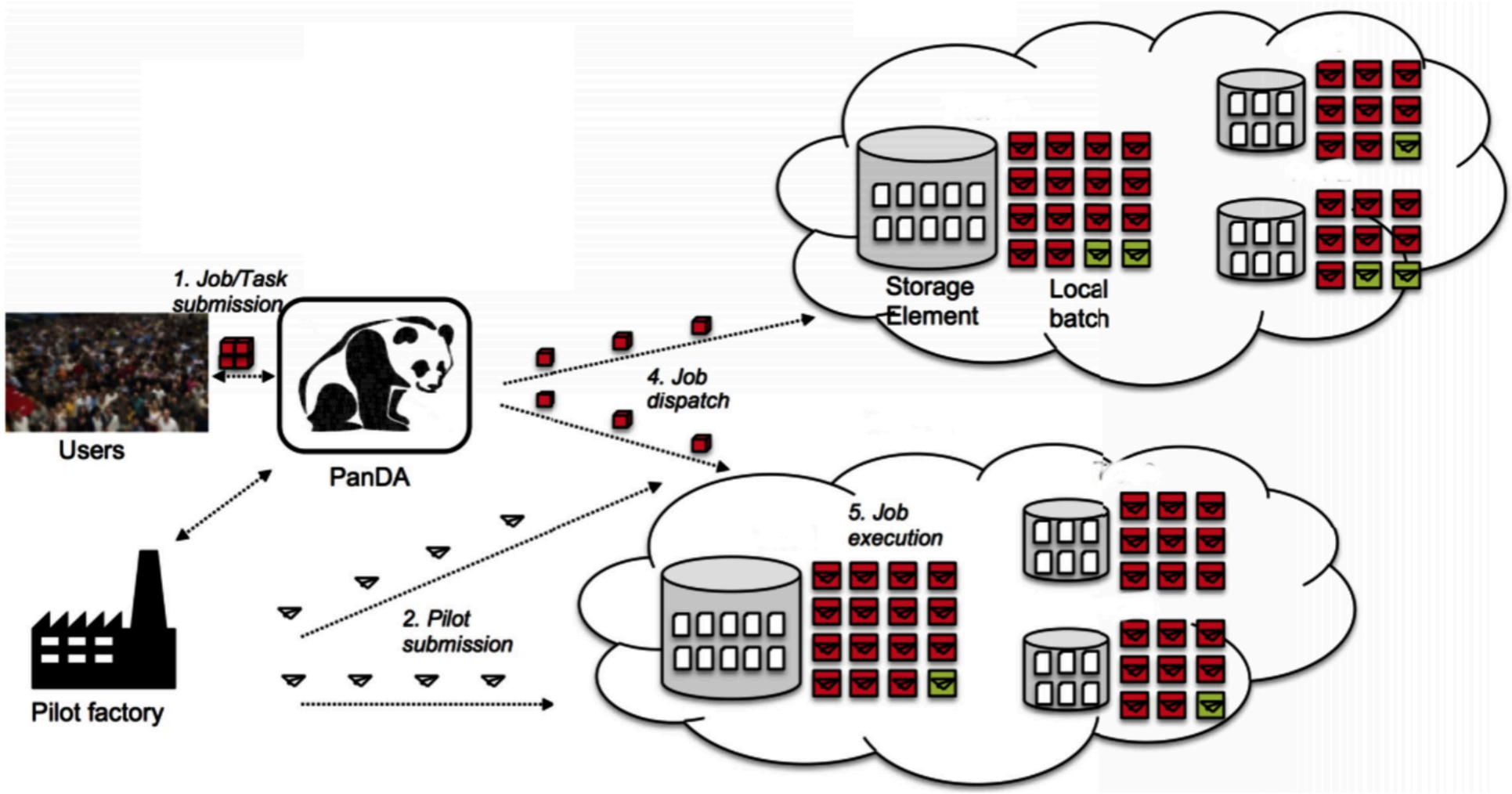
Automatically generates list of jobs for task basing on parameters

Job actions allow to manage any set of selected chunks



TA		RUN NUMBER	CHUNK NUMBER	PANDA ID	ATTEMPT	STATUS
dv	/2016/raw/W14/cdr11091-	275678	11091	2182400	1	finished
dv	/2016/raw/W14/cdr11082-275678.raw	275678	11082	2182399	1	finished
dvcs2016P09t2r13_mu+	/castor/cern.ch/compass/data/2016/raw/W14/cdr11080-275678.raw	275678	11080	2182398	1	finished
dvcs2016P09t2r13_mu+	/castor/cern.ch/compass/data/2016/raw/W14/cdr11089-275678.raw	275678	11089	2182397	1	finished
dvcs2016P09t2r13_mu+	/castor/cern.ch/compass/data/2016/raw/W14/cdr11086-275678.raw	275678	11086	2182396	1	finished
dvcs2016P09t2r13_mu+	/castor/cern.ch/compass/data/2016/raw/W14/cdr11063-275678.raw	275678	11063	2182395	1	finished
dvcs2016P09t2r13_mu+	/castor/cern.ch/compass/data/2016/raw/W14/cdr11049-275678.raw	275678	11049	2182394	1	finished
dvcs2016P09t2r13_mu+	/castor/cern.ch/compass/data/2016/raw/W14/cdr11016-275678.raw	275678	11016	2182393	1	finished
dvcs2016P09t2r13_mu+	/castor/cern.ch/compass/data/2016/raw/W14/cdr11094-275678.raw	275678	11094	2182392	1	finished
dvcs2016P09t2r13_mu+	/castor/cern.ch/compass/data/2016/raw/W14/cdr11092-275678.raw	275678	11092	2182391	1	finished
dvcs2016P09t2r13_mu+	/castor/cern.ch/compass/data/2016/raw/W14/cdr11088-275678.raw	275678	11088	2182390	1	finished
dvcs2016P09t2r13_mu+	/castor/cern.ch/compass/data/2016/raw/W14/cdr11076-275678.raw	275678	11076	2182389	1	finished

3. Job execution layer: PanDA

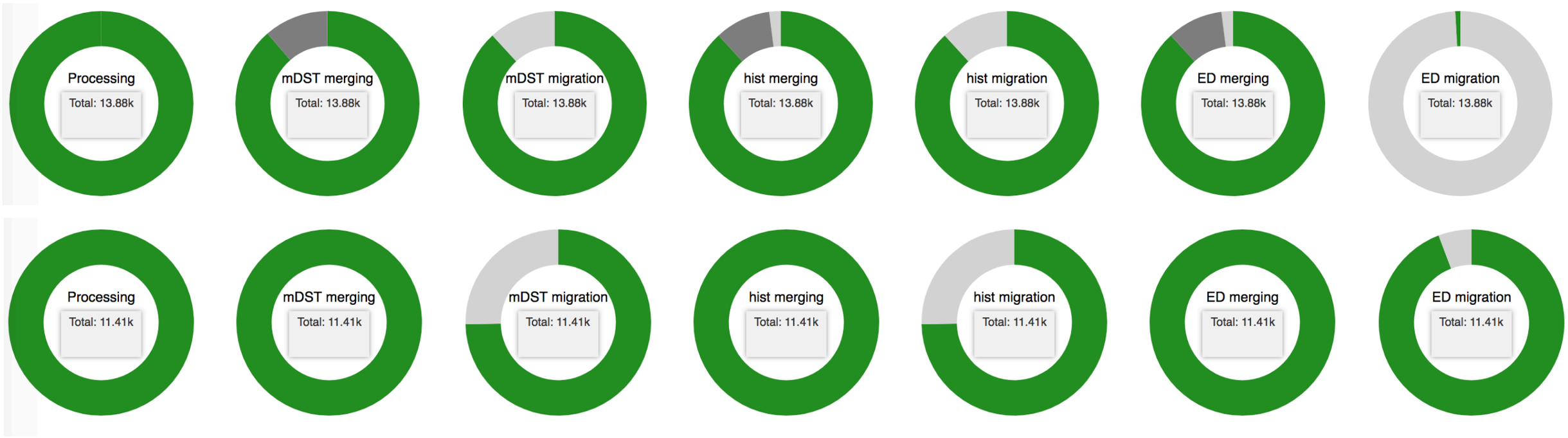


Schema by Misha Borodin

4: Workflow management

Decision making mechanisms guide task from the definition till archive

Each step of each task is managed independently



5: Data management

- Stage-in and stage-out files on Castor
- Number of events in raw files being delivered to ProdSys database, synchronously and asynchronously
- Data management during job execution performed by Pilot
- Job results move to Castor as soon as they are ready
- Job log files are zipped and moved to Castor when task is finished
- Job results and PanDA pilot log files are being removed from EOS when task is finished

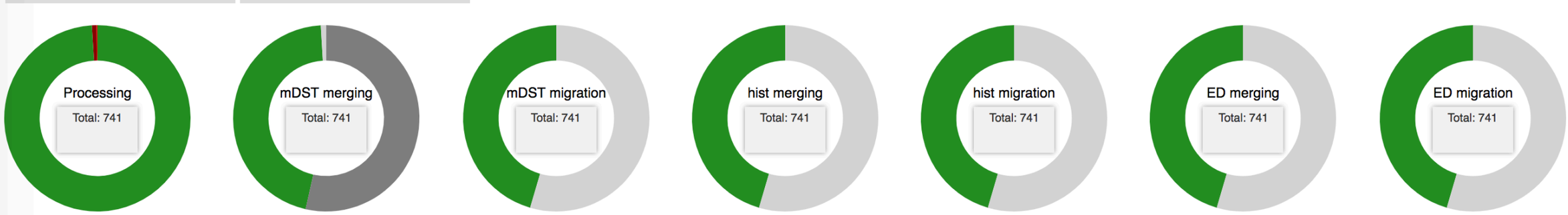
6.1: PanDA Monitoring

Covers all activity during production/task/job lifecycle

COMPASS PanDA | Dash | Tasks | Jobs | Errors | Users | Sites | Search | VO | Help

The summary for the **dvcs2016P09t2MBv3** production started on 06 Nov 2017. The total number of chunks is 741. The average walltime of a finished job is 147 minutes. Built 14:34 Actual version

Tasks for dvcs2016P09t2MBv3 | Jobs for dvcs2016P09t2MBv3



Show 50 entries Search:

Run	Number of chunks	Defined	Sent	Running	Failed	Finished	Status of mDST merging	X-checked	mDST migration	Status of histogram merging	Histogram migration	Status of event dump merging	Event dump migration
275518	404	-	-	-	7	397	-	no	-	-	-	-	-
275603	337	-	-	-	-	337	finished	yes	finished	finished	finished	finished	finished

Showing 1 to 2 of 2 entries Previous 1 Next

6.2: PanDA monitoring

Job attribute summary Sort by count, alpha	
attemptnr (9)	1 (4) 2 (8006) 3 (3468) 4 (1521) 5 (919) 6 (278) 7 (76) 8 (13) 11 (8)
computingsite (1)	CERN_COMPASS_PROD (14293)
destinationse (1)	local (14293)
jobstatus (8)	activated (243) defined (1) failed (2176) finished (7824) holding (164) running (99) sent (3770) starting (16)
minramcount (1)	0-1GB (14293)
priorityrange (2)	1000:1099 (13) 3000:3099 (14280)
prodsourcelabel (1)	prod_test (14293)
produsername (1)	Artem Petrosyan (14293)
taskid (6)	108 (1969) 109 (1606) 110 (1965) 111 (2834) 112 (2226) 113 (3693)
transformation (2)	DDD filtering (14280) merging dump (13)

Overall error summary

Category:code	Attempt list	Nerrors	% of job selection	Sample error description
jobdispatcher:102	jobs	2175	15.22	Sent job didn't receive reply from pilot within 30 min
transformation:1	jobs	1	0.01	Unspecified error, consult log file

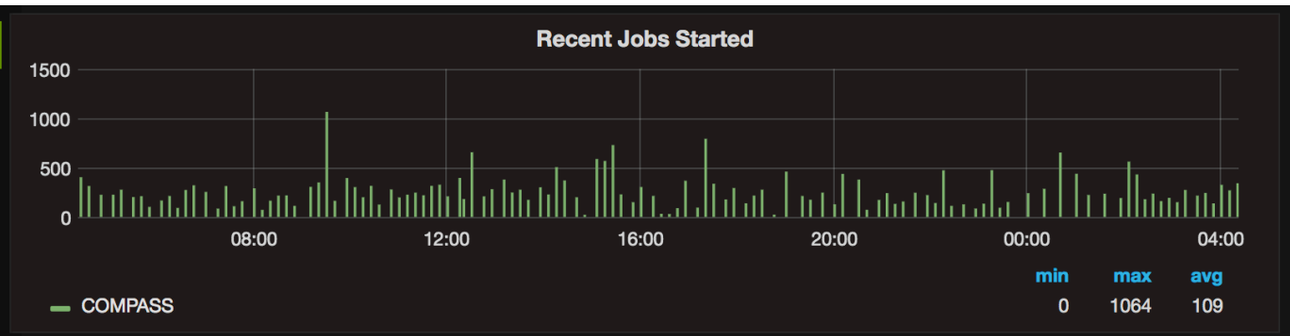
6.3: APF monitoring

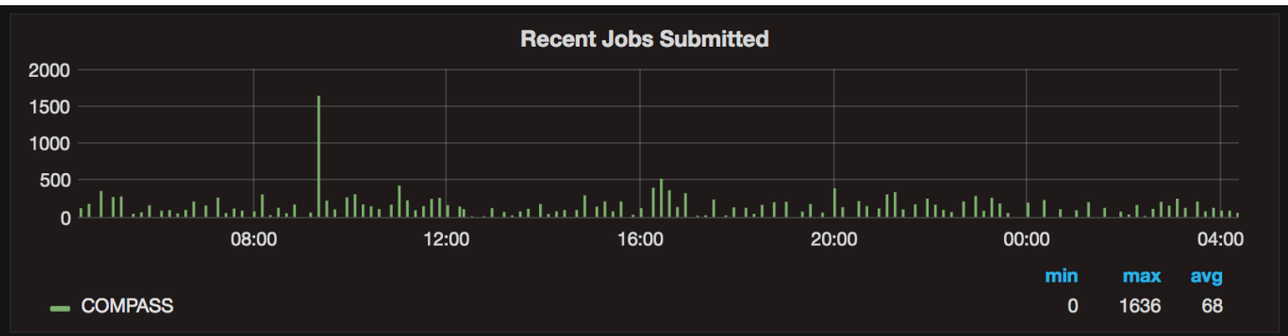
Factory view

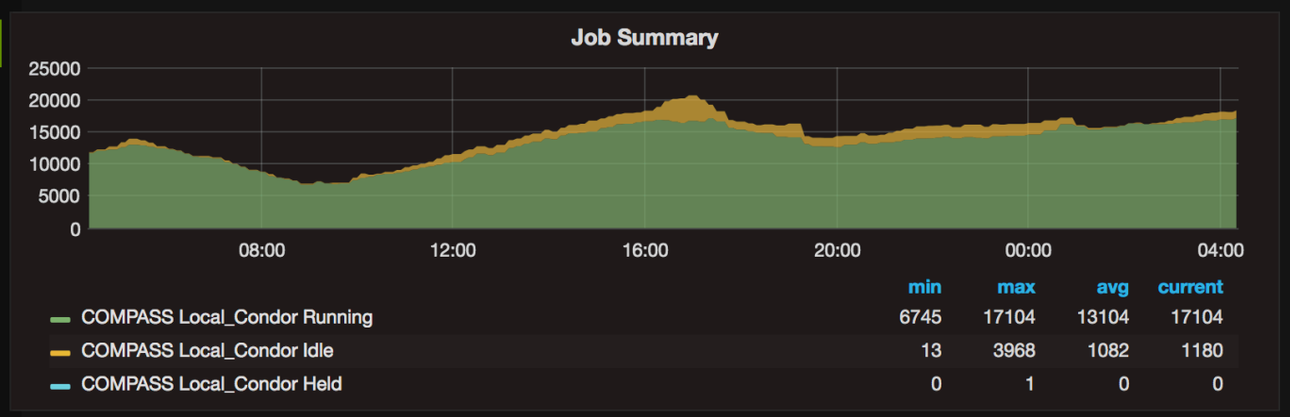
Factory	JINR-pandawms
Version	2.4.9
Last startup	2 days ago
Email	artem.petrosyan@jinr.ru
Activity	86
Links	logs queues.conf

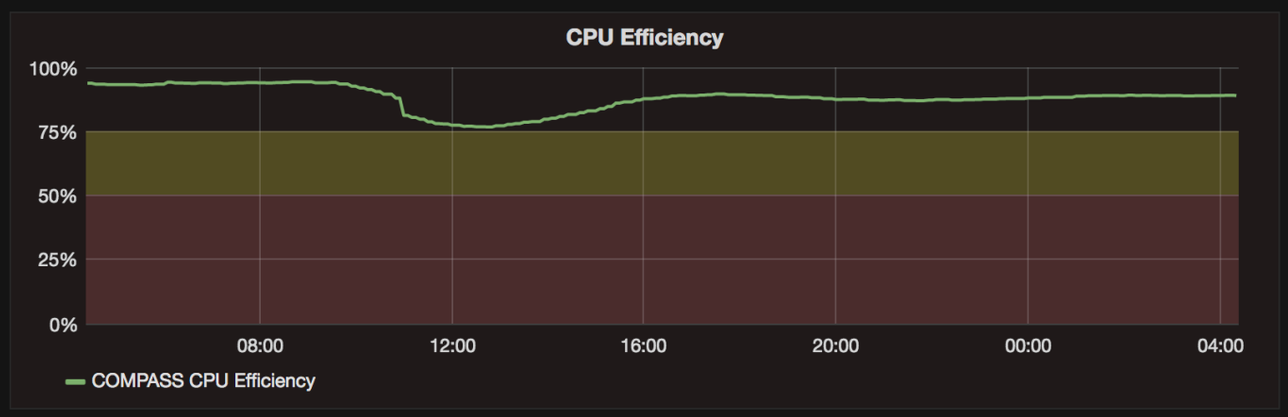
Factory label	last msg
CERN_COMPASS_PROD-ce301-cern-ch	6 mins ago
CERN_COMPASS_PROD-ce302-cern-ch	6 mins ago
CERN_COMPASS_PROD-ce401-cern-ch	6 mins ago
CERN_COMPASS_PROD-ce402-cern-ch	6 mins ago
CERN_COMPASS_PROD-ce403-cern-ch	6 mins ago
CERN_COMPASS_PROD-ce404-cern-ch	6 mins ago
CERN_COMPASS_PROD-ce405-cern-ch	6 mins ago
CERN_COMPASS_PROD-ce406-cern-ch	6 mins ago
CERN_COMPASS_PROD-ce407-cern-ch	6 mins ago
CERN_COMPASS_PROD-ce408-cern-ch	6 mins ago
CERN_COMPASS_PROD-ce503-cern-ch	4 mins ago
CERN_COMPASS_PROD-ce504-cern-ch	seconds ago
CERN_COMPASS_PROD-ce505-cern-ch	1 min ago
CERN_COMPASS_PROD-ce506-cern-ch	3 mins ago
CERN_COMPASS_PROD-ce507-cern-ch	4 mins ago
CERN_COMPASS_PROD-ce508-cern-ch	3 mins ago
CERN_COMPASS_PROD-condorce01-cern-ch	4 mins ago
CERN_COMPASS_PROD-condorce02-cern-ch	5 mins ago
CNAF_COMPASS_PROD-ce04-lcg-cr-cnaf-infn-it	6 mins ago
JINR_COMPASS_PROD-lcgce12-jinr-ru	6 mins ago
JINR_COMPASS_PROD-lcgce21-jinr-ru	6 mins ago
TRIESTE_COMPASS_PROD-ce1-ts-infn-it	6 mins ago

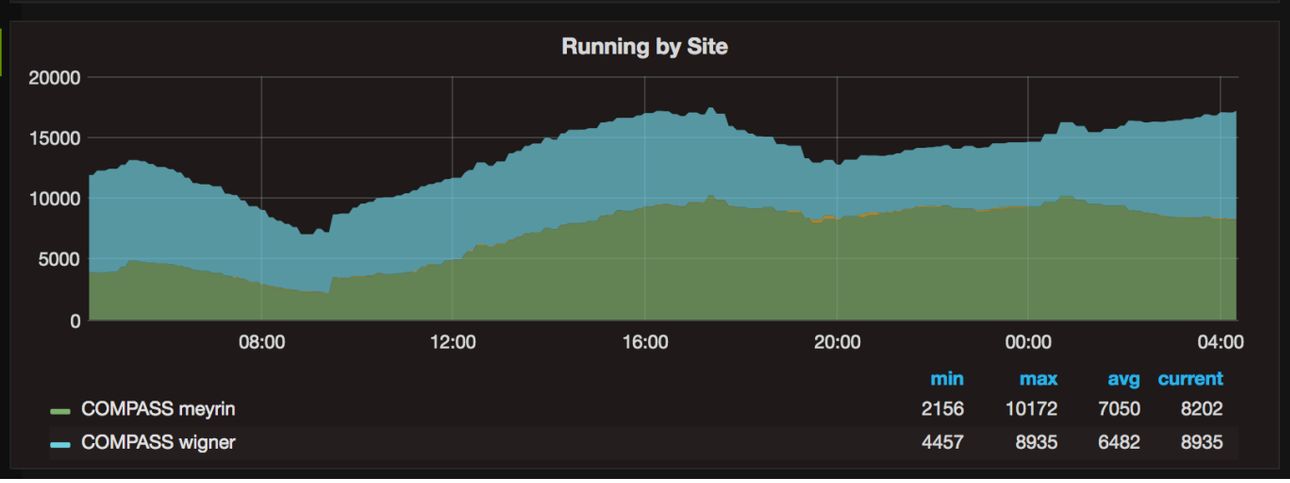
6.4: CERN Condor monitoring

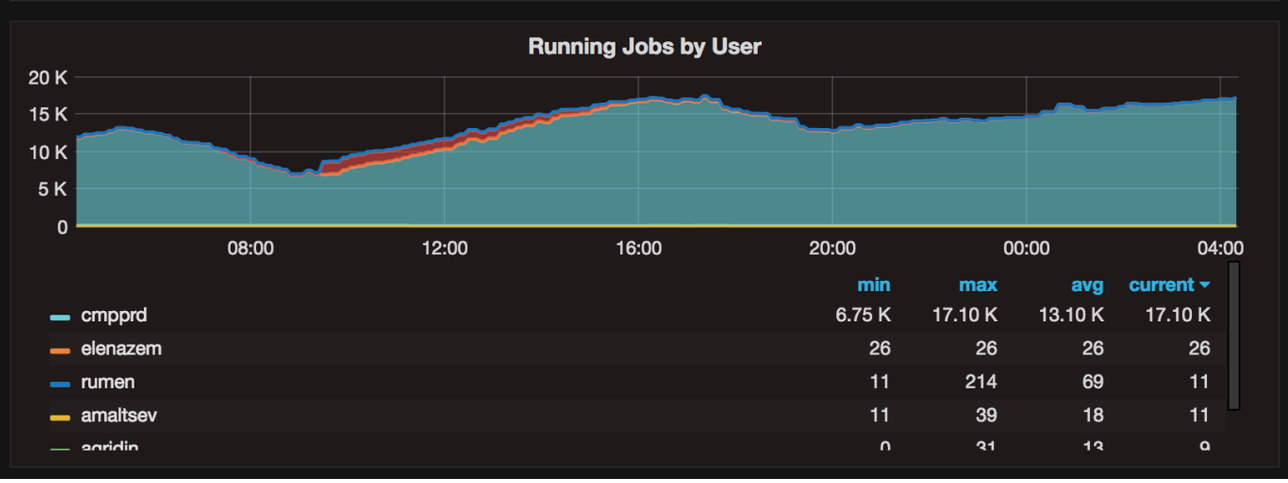












6.5: ProdSys Services Logs

periodic_tasks.check_castor_mdst_status_224.log.2018-10-24	2018-10-25 02:40	1.0M
periodic_tasks.check_castor_mdst_status_225.log	2018-10-26 01:10	655K
periodic_tasks.check_castor_mdst_status_225.log.2018-10-22	2018-10-23 01:40	2.4K
periodic_tasks.check_castor_mdst_status_225.log.2018-10-23	2018-10-24 02:40	89K
periodic_tasks.check_castor_mdst_status_225.log.2018-10-24	2018-10-25 02:40	1.2M
periodic_tasks.check_castor_mdst_status_226.log	2018-10-26 01:10	705K
periodic_tasks.check_castor_mdst_status_226.log.2018-10-23	2018-10-24 00:10	21K
periodic_tasks.check_castor_mdst_status_226.log.2018-10-24	2018-10-25 02:40	1.2M
periodic_tasks.check_castor_mdst_status_230.log	2018-10-29 02:10	65K
periodic_tasks.check_castor_mdst_status_230.log.2018-10-27	2018-10-28 02:40	208K
periodic_tasks.check_castor_mdst_status_231.log	2018-10-28 18:10	137K
periodic_tasks.check_castor_mdst_status_231.log.2018-10-27	2018-10-28 02:40	147K
periodic_tasks.check_castor_mdst_status_232.log	2018-10-28 09:10	58K
periodic_tasks.check_castor_mdst_status_232.log.2018-10-27	2018-10-28 02:40	183K
periodic_tasks.check_castor_status.log	2017-11-01 19:12	507
periodic_tasks.check_job_panda_status.log	2018-11-26 16:00	95K
periodic_tasks.check_job_panda_status.log.2018-11-19	2018-11-20 02:30	169K
periodic_tasks.check_job_panda_status.log.2018-11-20	2018-11-21 02:30	169K
periodic_tasks.check_job_panda_status.log.2018-11-21	2018-11-22 02:30	169K
periodic_tasks.check_job_panda_status.log.2018-11-22	2018-11-23 02:30	169K
periodic_tasks.check_job_panda_status.log.2018-11-23	2018-11-24 02:30	169K
periodic_tasks.check_job_panda_status.log.2018-11-24	2018-11-25 02:30	169K
periodic_tasks.check_job_panda_status.log.2018-11-25	2018-11-26 02:30	169K
periodic_tasks.check_job_panda_status_160_259363.log	2018-08-12 11:30	10K
periodic_tasks.check_job_panda_status_160_259363.log.2018-08-05	2018-08-06 01:58	30K
periodic_tasks.check_job_panda_status_160_259363.log.2018-08-06	2018-08-07 01:28	31K

Task types

- Technical production, test production, mass production
  - One raw file per job, which generates mDST, histogram and event dump, job logs, Pilot logs as output files
  - Once all jobs of one run are finished, merging of mDST begins
  - After merging cross check starts, if successful, merging histograms and event dumps follows, merging of event dumps is being cross checked after finish
  - When merging is finished, files are being migrated to Castor independently for mDST, histogram and event dumps
  - Merging prepares files of 4.2GB size
  - Paths on Castor for mass production and other types of production are different
- DDD filtering
  - Generates only event dumps, they are merged and cross checked

Job types

- Normal
    - File downloads from CASTOR to the computing node
    - After processing results are being transferred to EOS
- Merging
    - Data stages in from EOS
    - Up to 1000 results of normal jobs are merged into one or several files with desired filesize (4Gb)
    - After processing result file are being transferred to EOS
- Cross check
    - Internal job, uses PanDA job metrics
    - Compares files size and number of events in file chunks and in merged file per run

Statuses

- Task statuses
    - Draft, ready, jobs ready, send, running, paused, cancelled, done, archive, archiving, archived
- Job statuses
    - Defined, staging, staged, sent, running, failed, paused, cancelled, finished, manual check is needed
- Job substatuses
    - PanDA status, status merging, status cross check, status merging histos, status merging event dumps, status cross check event dumps, status castor, status castor histos, status castor event dumps, logs deleted, logs archived, status logs castor
- + PanDA job statuses

Stats and performance

- Since August 2017
  - ~3 000 000 chunks of raw data processed
  - ~80 000 000 of events processed
  - ~500TB of merged data produced and migrated to Castor
  - ~6 000 000 jobs processed since August: reco, ddd filtering, merging of mDST, hist and event dumps
- Up to 20 000 of jobs are being processed simultaneously

Summary

- COMPASS Production System provides automated data processing from task definition till archiving

- Features:
  - Production management Web UI allows to define a task, send, follow and manage task at any step during processing
  - Via PanDA layer jobs are being delivered to any type of available computing resource: Condor, LSF, PBS, etc.
  - ProdSys contains 26 independently running management services
  - Rich monitoring

Processing on Blue Waters

Blue Waters System Overview

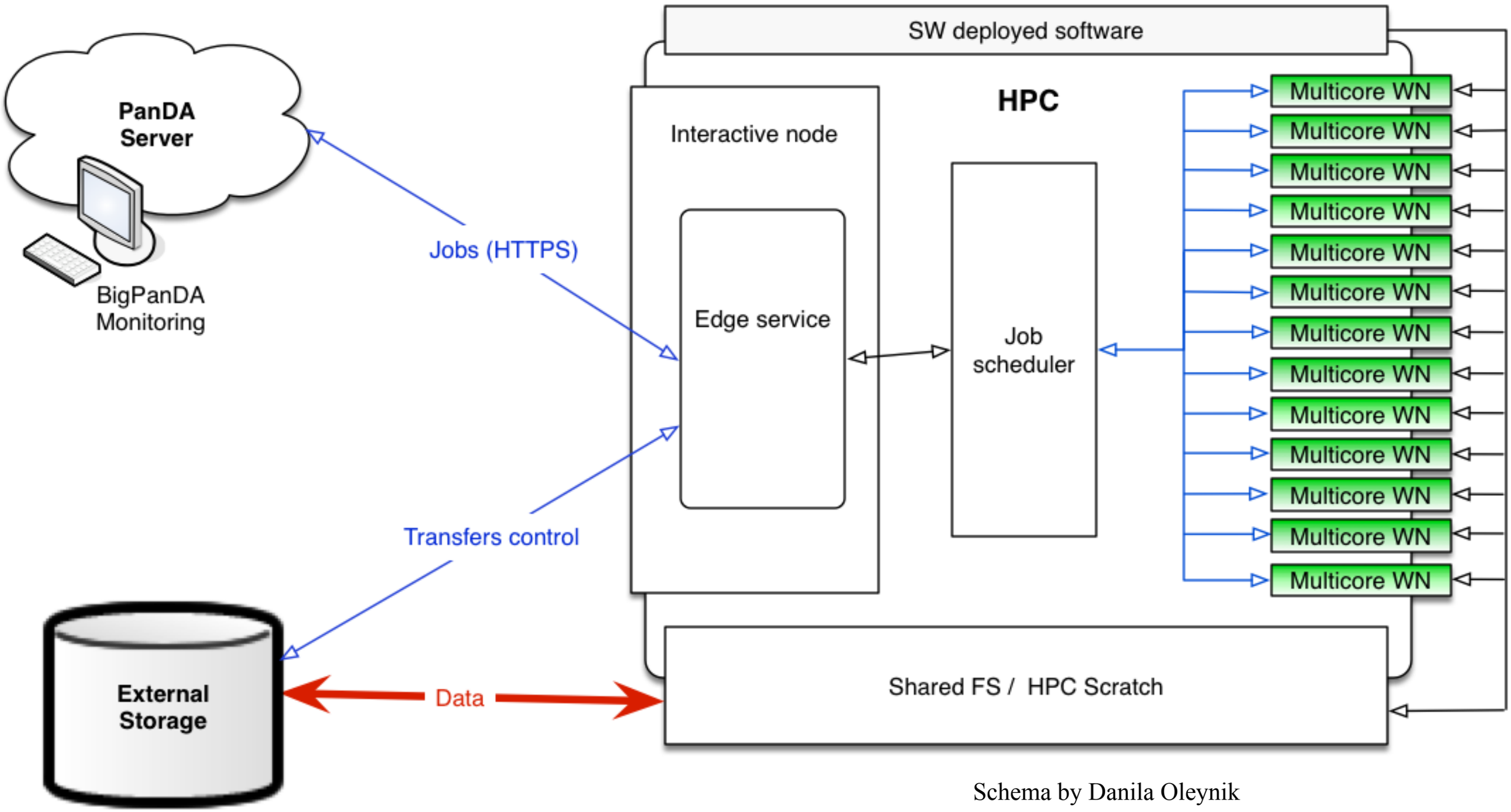
- The Blue Waters system is a Cray XE/XK hybrid machine composed of AMD 6276 "Interlagos" processors (nominal clock speed of at least 2.3 GHz) and NVIDIA GK110 (K20X) "Kepler" accelerators all connected by the Cray Gemini torus interconnect.
- Total Peak Performance: 13.34 PF
- Total System Memory: 1.634 PB
- Total Usable Storage: 26.4 PB
- COMPASS allocation at BW: 9 million node-hours per year

Action items to enable processing on BW via PanDA

- Ordinary Grid site
  - 1 pilot 1 job
  - Pilots delivered by AutoPyFactory
  - Grid environment
- Input from Castor/EOS
- Output to EOS
- CVMFS

- BW
  - 1 pilot N jobs
  - Pilots run in work directory of the prod user (daemon)
  - Proxy delivered via SCP
- Input from local file system
- Output to local file system
- Local SW installation

PanDA for HPC



Schema by Danila Oleynik

Processing on Blue Waters

- Allocation: 9M node hours per year
- Raw data delivered to BW manually via Globus Online
- Production software installed on local file system
- Calibration db runs on each computing node, i.e. per each 32 jobs, first job on the node starts new db instance
- PanDA Multi-Job Pilot is used, extended by COMPASS logic
  - Submission size: each Pilot can run up to 512 jobs on16 nodes
- Task submission, management and monitoring fully integrated into ProdSys UI and PanDA monitoring
- Processing 25-50K jobs, 500-1000 nodes, target is to process 100-150k of jobs

Solved issues

- PanDA server was upgraded in order to increase jobs dispatch rate from 1 per minute to 500 per minute in bulk mode

- Infrastructure changes: from PanDA server, DB and AutoPyFactory on one machine to dedicated server for each service: PanDA server, DB, AutoPyFactory

- Pilots are consuming CPU resources and, when run on login node, being removed by process watcher. In order to get rid of that, pilots are now run on a MOM node, shared node for submissions management

- Archiving of logs at Pilot side was removed in order to reduce CPU consumption

- COMPASS calibration database has to run with jobs on the same node since there is no commutation between worker nodes during execution

Jobs submission tuning

- Pilot can work stable with 512 jobs
- If PanDA server replies that there is no jobs, smaller submission is prepared
- Production jobs run up to 18 hours, depending on number of events in the raw file
- Merging of production job results run 1 hour
- Merging of histograms runs 30 hour
- Merging of event dumps runs less than 30 minutes
- In order to avoid requesting excessive resources, three queues were defined: long for processing, shorter for merging of job results and short for histogram and event dumps merging

System performance

Job attribute summary Sort by count, alpha	
attemptnr (8)	1 (18) 4 (1913) 5 (2823) 6 (7831) 7 (11104) 8 (10595) 9 (3343) 10 (708)
computingsite (1)	BW_COMPASS_MCORE (38335)
destinationse (1)	local (38335)
jobstatus (7)	activated (4292) failed (4) finished (6679) holding (65) running (25201) starting (2093) transferring (1)
minramcount (1)	0-1GB (38335)
priorityrange (2)	1000:1099 (18) 2000:2099 (38317)
prodsourcelabel (1)	prod_test (38335)
production (1)	dy2015W07t5BW (38317)

Submissions tuning

3885773.bw	petrosya	normal	SAGA-Python-PBSJ	29778	16	512	--	24:00:00	R	07:40:07
3885779.bw	petrosya	normal	SAGA-Python-PBSJ	17154	16	512	--	24:00:00	R	02:02:20
3887209.bw	petrosya	normal	SAGA-Python-PBSJ	22097	16	512	--	18:00:00	R	15:51:38
3888162.bw	petrosya	normal	SAGA-Python-PBSJ	32692	16	512	--	18:00:00	R	05:18:11
3888276.bw	petrosya	normal	SAGA-Python-PBSJ	--	16	512	--	18:00:00	Q	--
3888278.bw	petrosya	normal	SAGA-Python-PBSJ	--	16	512	--	18:00:00	Q	--
3888281.bw	petrosya	normal	SAGA-Python-PBSJ	--	16	512	--	18:00:00	Q	--
3888282.bw	petrosya	normal	SAGA-Python-PBSJ	--	16	512	--	18:00:00	Q	--
3888283.bw	petrosya	normal	SAGA-Python-PBSJ	--	16	512	--	18:00:00	Q	--
3888286.bw	petrosya	normal	SAGA-Python-PBSJ	--	16	512	--	18:00:00	Q	--
3888289.bw	petrosya	normal	SAGA-Python-PBSJ	--	16	512	--	18:00:00	Q	--
3888290.bw	petrosya	normal	SAGA-Python-PBSJ	--	16	512	--	18:00:00	Q	--
3888291.bw	petrosya	normal	SAGA-Python-PBSJ	--	16	512	--	18:00:00	Q	--
3888292.bw	petrosya	normal	SAGA-Python-PBSJ	--	16	512	--	18:00:00	Q	--
3888295.bw	petrosya	normal	SAGA-Python-PBSJ	--	16	512	--	18:00:00	Q	--
3888297.bw	petrosya	normal	SAGA-Python-PBSJ	--	16	512	--	18:00:00	Q	--
3888299.bw	petrosya	normal	SAGA-Python-PBSJ	--	16	512	--	18:00:00	Q	--
3888300.bw	petrosya	normal	SAGA-Python-PBSJ	--	16	512	--	18:00:00	Q	--
3888301.bw	petrosya	normal	SAGA-Python-PBSJ	--	16	512	--	18:00:00	Q	--
3888304.bw	petrosya	normal	SAGA-Python-PBSJ	--	16	512	--	18:00:00	Q	--
3888307.bw	petrosya	normal	SAGA-Python-PBSJ	--	16	512	--	18:00:00	Q	--
3888308.bw	petrosya	normal	SAGA-Python-PBSJ	--	16	512	--	18:00:00	Q	--
3888309.bw	petrosya	normal	SAGA-Python-PBSJ	--	16	512	--	18:00:00	Q	--
3888339.bw	petrosya	normal	SAGA-Python-PBSJ	--	16	512	--	18:00:00	Q	--
3888340.bw	petrosya	normal	SAGA-Python-PBSJ	--	16	512	--	18:00:00	Q	--
3888341.bw	petrosya	normal	SAGA-Python-PBSJ	--	16	512	--	18:00:00	Q	--
3888346.bw	petrosya	normal	SAGA-Python-PBSJ	--	16	512	--	18:00:00	Q	--
3888349.bw	petrosya	normal	SAGA-Python-PBSJ	--	16	512	--	18:00:00	Q	--
3888356.bw	petrosya	normal	SAGA-Python-PBSJ	--	16	512	--	18:00:00	Q	--
3888358.bw	petrosya	normal	SAGA-Python-PBSJ	18518	16	512	--	18:00:00	R	00:14:45
3888368.bw	petrosya	normal	SAGA-Python-PBSJ	--	16	512	--	18:00:00	Q	--
3888370.bw	petrosya	normal	SAGA-Python-PBSJ	--	16	512	--	18:00:00	Q	--
3888372.bw	petrosya	normal	SAGA-Python-PBSJ	--	16	512	--	18:00:00	Q	--
3888373.bw	petrosya	normal	SAGA-Python-PBSJ	--	16	512	--	18:00:00	Q	--

CPU consumption by Pilots

```
top - 03:26:03 up 21 days, 11:27,  1 user,  load average: 16.98, 18.98, 19.47
Tasks: 677 total,  19 running, 658 sleeping,   0 stopped,   0 zombie
Cpu(s): 45.1%us,  9.8%sy,  0.0%ni, 44.9%id,  0.2%wa,  0.0%hi,  0.1%si,  0.0%st
Mem:     64624M total,    46480M used,    18143M free,       13M buffers
Swap:        0M total,        0M used,        0M free,    28803M cached

  PID USER      PR  NI  VIRT  RES  SHR S   %CPU %MEM    TIME+  COMMAND
22317 petrosya  20   0 1239m  60m 6520 R    101  0.1  16:51.14 python
28517 petrosya  20   0 1239m  60m 6516 R    100  0.1  31:27.62 python
30344 petrosya  20   0 1239m  60m 6516 R    100  0.1  36:54.66 python
 3184 petrosya  20   0 1239m  60m 6520 R    100  0.1  41:00.07 python
28519 petrosya  20   0 1240m  60m 6536 R    100  0.1  17:54.20 python
30225 petrosya  20   0 1303m  59m 6520 R    100  0.1  67:38.51 python
  594 petrosya  20   0 1302m  60m 6516 R    100  0.1 135:29.71 python
 6902 petrosya  20   0 1239m  61m 6520 R    100  0.1 134:55.77 python
17423 petrosya  20   0 1303m  59m 6520 R    100  0.1 100:24.62 python
17962 petrosya  20   0 1239m  60m 6524 R    100  0.1  45:07.43 python
20174 petrosya  20   0 1239m  60m 6516 R    100  0.1  31:30.95 python
16537 petrosya  20   0 1240m  59m 6520 R     99  0.1  15:34.36 python
 6937 petrosya  20   0 1240m  60m 6520 R     92  0.1  58:02.83 python
 7532 petrosya  20   0 1231m  54m 6496 R     57  0.1   9:18.13 python
 4950 petrosya  20   0 1293m  51m 6496 R     34  0.1  62:21.43 python
 6906 petrosya  20   0 1229m  52m 6500 R     33  0.1  57:43.53 python
 7609 petrosya  20   0 1293m  51m 6496 S     30  0.1  53:56.00 python
 5813 petrosya  20   0 1290m  48m 6508 S     20  0.1   5:24.34 python
31766 petrosya  20   0 1293m  52m 6496 S     12  0.1  55:38.07 python
10889 petrosya  20   0 1293m  51m 6496 S     11  0.1  58:02.88 python
23805 petrosya  20   0 1293m  51m 6496 S     10  0.1  60:55.16 python
28951 petrosya  20   0 1293m  51m 6500 S     10  0.1  59:03.84 python
22460 petrosya  20   0 1293m  51m 6496 R      8  0.1  64:52.34 python
18594 root       0 -20     0    0    0 S      7  0.0 148:34.66 kgnilnd_sd_00
18595 root       0 -20     0    0    0 S      7  0.0 147:25.54 kgnilnd_sd_01
31351 petrosya  20   0 1290m  48m 6508 S      7  0.1   5:45.84 python
18596 root       0 -20     0    0    0 S      7  0.0 148:21.28 kgnilnd_sd_02
27177 petrosya  20   0 1226m  49m 6508 S      7  0.1   5:24.06 python
25964 petrosya  20   0 1291m  49m 6508 S      6  0.1   5:44.03 python
32322 petrosya  20   0 1290m  49m 6508 S      6  0.1   6:00.73 python
13006 petrosya  20   0 1291m  50m 6508 R      6  0.1   5:25.04 python
14155 petrosya  20   0 1291m  49m 6512 S      5  0.1   5:07.45 python
20503 petrosya  20   0 1293m  52m 6496 S      5  0.1  56:00.92 python
```

Summary

- ProdSys runs COMPASS production jobs via PanDA on Blue Waters

- Environment for automated data processing on BW was prepared and runs reliably in daemon mode

- Further development to reach 150K of running jobs

  - Upgrade to PanDA Harvester will allow to consume more resources with higher level of stability and efficiency

Infrastructure overview

- PanDA server, MySQL, Monitoring, AutoPilotFactory/Harvester, Production System deployed in Dubna at JINR cloud service
- Condor CE at CERN
- PBS CE at JINR
- Blue Waters HPC at Urbana Champaign
- EOS SE at CERN
- Castor at CERN
- PerfSonar service at JINR cloud network segment to monitor network connectivity between JINR and CERN
- CRIC information system at CERN

Conclusions

- COMPASS situation of high dependency on Castor, AFS, LSF, etc. was solved
- New system, which provides distributed data processing on any type of computing resources, was built
- Software components, developed for experiment on LHC, allow to construct from them production system with necessary set of characteristics
- Key areas of effort
  - Management: Grid environment
  - Development: workflow management services — each project is unique and there is no out of the box solution to describe all workflow nuances

Acknowledgements

- COMPASS: Elena Zemlyanichkina, Vladimir Frolov, Sergey Gerasimov, Andrea Bressan, Yann Bedfer, Bakur Parsamyan
- PanDA Pilot, Multi-Job Pilot, Harvester: Danila Oleynik
- PanDA: Tadashi Maeno, Fernando Barreiro Megino, Ruslan Mashinistov, Pavlo Svirin
- PanDA Monitoring for COMPASS: Tatiana Korchuganova
- Certificate authentication for Django dev, instances deployment at JINR: Daniil Malevanniy
- Blue Waters COMPASS team: Marco Meyer, Riccardo Longo, Caroline Riedl
- CRIC: Alexey Anisenkov
- JINR Cloud: Nikolay Kutovskiy, Alexander Baranov
- CE admins, experts: Ben Jones, Valery Mitsyn, Craig Steffen

Thanks!