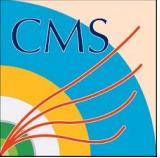# Agent Monitoring Overview
WMStats, WMArchive, WMAgent Monitor, Dashboard

Seangchan Ryu(FNAL)
on behalf of WMCore team

Overview

1. Summary of monitoring applications
2. Data sources
3. Data contents
4. Usage
5. Future plans

Category of current monitoring tools

1. Monitor for smooth data operation (WMStats, WMAgent Monitor, Dashboard)
   a. Monitors system healthiness
   b. Spots the problems as early as possible
   c. Provides helpful information for the cause of the problem
   d. Provides validation on data produced
   e. Delivers the information in real time (or close to real time)
   f. Short term transient data
2. Monitor for statistics, analytics and reports (WMArchive, Dashboard)
   a. Provides different perspective of data spanning a longer period of time
   b. Helps to spot hidden problems which couldn't be detected in short term
   c. Provides statistics that help determine long term strategy
   d. Long term permanent data

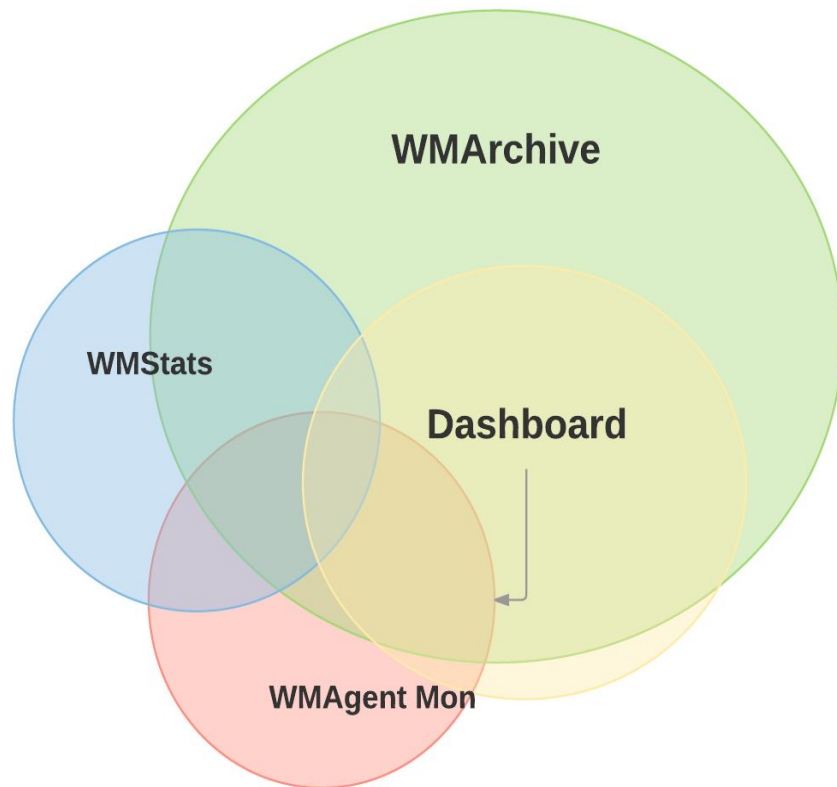|	WMStats	WMArchive	Dashboard	WMAgent Mon
Usage	Trouble shooting, Alarm,	Stats, Analytics	Stats	Stats.
Data lifetime	Transient (only keeps the current snapshot)	Permanent	Semi Permanent?	Permanent
Contents	Job summary by Tasks. Partial FWJRs for only failed jobs All the request properties	All the FWJRs	Task, Job, NCore, exit code, site, cpu time. Wall clock time	Job summary by WMAgent, Site, WQ
Delay	~10 min	~1hour - 1day	~10 min?	~10 min

Data sources

1. WMStats: (every 10 min)
   a. Data from job summary and subset of FWJR for failed jobs
2. WMArchive: (every 10 min),
   a. Bulk FWJR update from couchdb
3. WMAgent Monitor:(every 10 min - still not in production)
   a. Work overview (also distributed by site, priority, status) from global workqueue couchDB
   b. Work overview from WMAgent (both local workqueue and WMBS)
4. Dashboard:
   a. From ReqMgr2, WMAgent and worker node when job status is changed
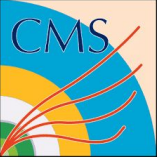
Monitoring Applications (content and source)

1. WMStats: summaries of job statistics by request/task and by agent, plus subset of FWJRs for only failed jobs - local WMAgent CouchDB - can lose the data

2. WMArchive: collection of all the framework job report - local WMAgent CouchDB - can lose the data

3. Dashboard: job stats (partially from by requests (keeps the data) - live update, when job is changing its status - can lose the data

4. WMAgent Monitor: job statistics by agents (not by task)- from WMAgent WMBS - consistent

Data contents

WMArchive

WMStats

Dashboard

WMAgent Mon

1. **WMArchive**: all data (fwjr + agent/request meta data)

   https://github.com/dmwm/WMArchive/blob/master/src/python/WMArchive/Schemas/FWJRProduction.py

   Unique data: Performance, (CPU, Read, Write etc) - step level information

   Not available Data: Current running jobs. WMArchive only collects the data when jobs are finished. some of the Request Properties, (Total events/lumis, output dataset, priority), WQ elements related to jobs,

2. **WMStats**: (summary + all failed job reports) - transient

   Unique data: Request properties, output datasets, Total events/lumis

   Not available data: individual job info for the successful job

   Common data: site, job status, few performance matics, event, lumis, priority, exitcode information

3. **WMAgent Monitor**: (summary)

   Unique data: WQE information, job with priority by site and agent, site threshold information.

   Not available Data: some of the Request Properties, (Total events/lumis, Output dataset, priority), data by request and task.

   The information can be used for the job submission decision making for balancing work.

4. **Dashboard**: (all data)

Unique data: Site pledge, CPUBound, IOBound and site status for production (not from wmagent)

Not available data: performance matrix

Common data: site, job status, few performance matics, exitcode information

Data inconsistency

1. Different data sources - couch db, wmbs, live data
2. Time delay of data update
3. Missing data - Only WMBS is the consistent data source. (If couch updates fails, won't be retried)
4. Difference in job counting - WMArchive counts all the the jobs including retries. Other applications consider retry jobs as the same job. (i.e. WMArchive counts 2 job failures and one successful job if a job is retried 3 times and succeed eventually. In the same case, WMStats and Dashboard count that as 1 successful job).
5. Error code difference - Dashboard only record errors in the worker node. WMStats/WMArchive preserve all the errors in different layer. (But only provide search for top level errors)

Except 4, 5 cases, the difference of job stats shouldn't be excessive.

Usage of the different monitors

WMStats: Agent Monitoring

WMStats: Error/Warning logs by request

| active request | request alert (653, 14) | agent info (1) | error logs (48) | search | 2017/11/14 (Tue) 15:46:07 UTC |

▼ pdmvserv_task_HIG-RunIISummer15wmLHEGS-01969__v1_T_171108_133552_3969

- **agent**: vocms0253.cern.ch
- **thread**: WorkQueue
- **type**: agent-error
- **update time**: 2017/11/13 (Mon) 11:14:59 UTC
- **error message**:

    pdmvserv_task_HIG-RunIISummer15wmLHEGS-01969__v1_T_171108_133552_3969, None:
    creating subscription failed in LQ:
    (IntegrityError) ORA-00001: unique constraint (CMS_WMBS_PROD12.WMBS_FILESET_UNIQUE) violated
    'INSERT INTO wmbs_fileset (id, name, last_update, open)\n   VALUES (wmbs_fileset_SEQ.nextval, :NAME, :LAST_UPDATE, :OPEN)' {'OPEN': 1, 'NAME':
    u'pdmvserv_task_HIG-RunIISummer15wmLHEGS-01969__v1_T_171108_133552_3969-HIG-RunIISummer15wmLHEGS-01969_0-7ef8f14641e39205fe91e3efce19f79b', 'LAST_UPDATE':
    1510571663}

- **id**: 0125c0c7f4470ee899c47c01a9cfcb89

▶ pdmvserv_task_HIG-RunIISummer15wmLHEGS-01970__v1_T_171108_133702_6213
▶ pdmvserv_task_HIG-RunIISummer15wmLHEGS-01986__v1_T_171108_134054_511
▶ pdmvserv_task_HIG-RunIISummer15wmLHEGS-01837__v1_T_171108_145833_4921
▶ pdmvserv_task_HIG-RunIISummer15wmLHEGS-01997__v1_T_171108_134314_5199
▶ pdmvserv_task_HIG-RunIISummer15wmLHEGS-01830__v1_T_171108_143021_4
▶ pdmvserv_task_HIG-RunIISummer15wmLHEGS-01857__v1_T_171108_143639_1343
▶ pdmvserv_task_HIG-RunIISummer15wmLHEGS-01927__v1_T_171108_135343_5319
▼ pdmvserv_task_HIG-RunIISummer15wmLHEGS-01991__v1_T_171108_134217_2216

WMStats Alarms

| | EGM-RunIIFall17GS-00010_0 | jobcooloff | T2_UK_SGrid_RALPP | 99303 | 1 | NoJobReport | |
| | EGM-RunIIFall17GS-00010_0 | jobcooloff | T2_US_MIT | 99303 | 4 | NoJobReport | |

Showing 1 to 25 of 8,179 entries

Previous 1 **2** 3 4 5 ... 328 Next

`1`

- **Job Name:** 2b82925a-c89f-11e7-9a87-02163e018570-867
- **WMBS job id:** 2280623
- **Workflow:** pdmvserv_task_EGM-RunIIFall17GS-00010__v1_T_171013_161526_3751
- **Task:** /pdmvserv_task_EGM-RunIIFall17GS-00010__v1_T_171013_161526_3751/EGM-RunIIFall17GS-00010_0
- **Status:** jobcooloff
- **Input dataset:**
- **Site:** T2_TW_NCHC
- **Agent:** vocms0253.cern.ch
- **ACDC URL:** https://cmsweb.cern.ch/couchdb/acdcserver
- ▶ **State Transition**
- **Exit code:** 99109
- **Retry count:** 0
- **Errors:**
    - logArch1
        - **LogArchiveFailure (Exit Code: 60307)**

            - `<@========== WMException Start ==========@>`
              `Exception Class: StageOutFailure`
              `Message: Failure for fallback stage out:`
              `<@========== WMException Start ==========@>`
              `Exception Class: StageOutError`
              `Message: 2017-11-14T15:13:01 : Command exited non-zero ExitCode:127`
              `Output: (stdout: /bin/sh: line 2: srmcp: command not found`
              `srmcp exit status:`
              `ERROR: srmcp exited with`
              `Cleaning up failed file:`

              `stderr: cat: /srv/job/WMTaskSpace/logArch1/srm.report.929: No such file or directory`
              `/bin/sh: line 12: srmrm: command not found`
              `)`
              `        ErrorCode : 60311`
              `        ModuleName : WMCore.Storage.StageOutError`

14

WMStats: Request Summary

- **progress:**
  - /TTJets_TuneCP5_13TeV-amcatnloFXFX-pythia8/RunIIFall17wmLHEGS-93X_mc2017_realistic_v3-v2/GEN-SIM: event: 21.8%, lumi: 61.5%
  - /TTJets_TuneCP5_13TeV-amcatnloFXFX-pythia8/RunIIFall17wmLHEGS-93X_mc2017_realistic_v3-v2/LHE: event: 60.5%, lumi: 60.5%
- **output events:** 0
- **queued (first):** 779
- **queued (retried):** 714
- **cooloff jobs:** 3
- **pending:** 21148
- **running:** 500
- **failure:** 26802
- **success:** 119102
- ▶ **Sites**
- ▼ **Skipped Summary**
  - **TOP-RunIIFall17wmLHEGS-00010_0MergeRAWSIMoutput:** T2_CH_CERN_HLT-skippedFiles:4, T2_FR_GRIF_LLR-skippedFiles:2, T2_CH_CERN-skippedFiles:2, T1_UK_RAL-skippedFiles:3, T2_FR_GRIF_IRFU-skippedFiles:1, T2_UK_London_Brunel-skippedFiles:4, T2_UK_SGrid_RALPP-skippedFiles:5, T2_US_Nebraska-skippedFiles:1, T1_ES_PIC-skippedFiles:3, T2_UK_London_IC-skippedFiles:4, T1_US_FNAL_Disk-skippedFiles:1, T2_US_Vanderbilt-skippedFiles:2,
  - **TOP-RunIIFall17wmLHEGS-00010_0MergeLHEoutput:** T2_CH_CERN_HLT-skippedFiles:1, T2_FR_GRIF_IRFU-skippedFiles:3, T2_FR_GRIF_LLR-skippedFiles:24, T1_US_FNAL_Disk-skippedFiles:13, T2_UK_London_IC-skippedFiles:10, T2_UK_London_Brunel-skippedFiles:21, T1_UK_RAL-skippedFiles:3, T2_UK_SGrid_RALPP-skippedFiles:58, T1_ES_PIC-skippedFiles:3, T2_US_Vanderbilt-skippedFiles:4,

WMArchive: Stats/Analytics
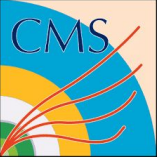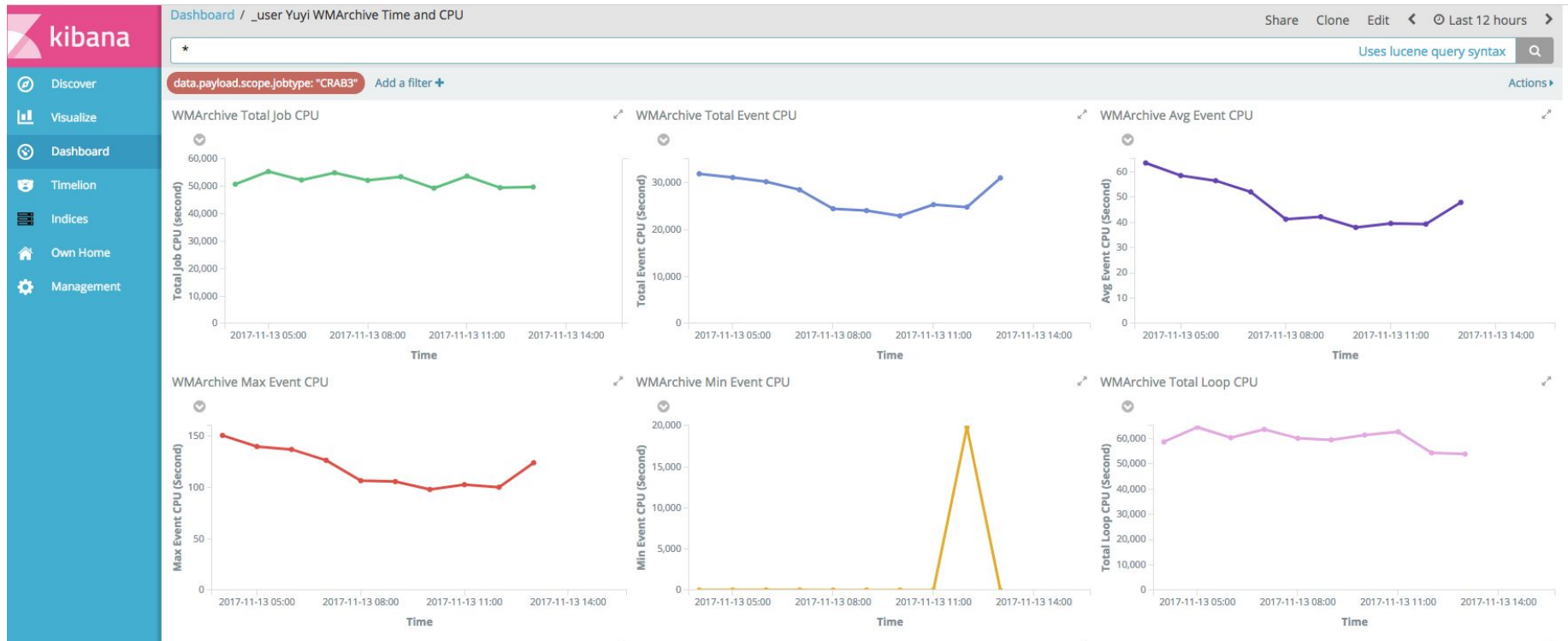

16

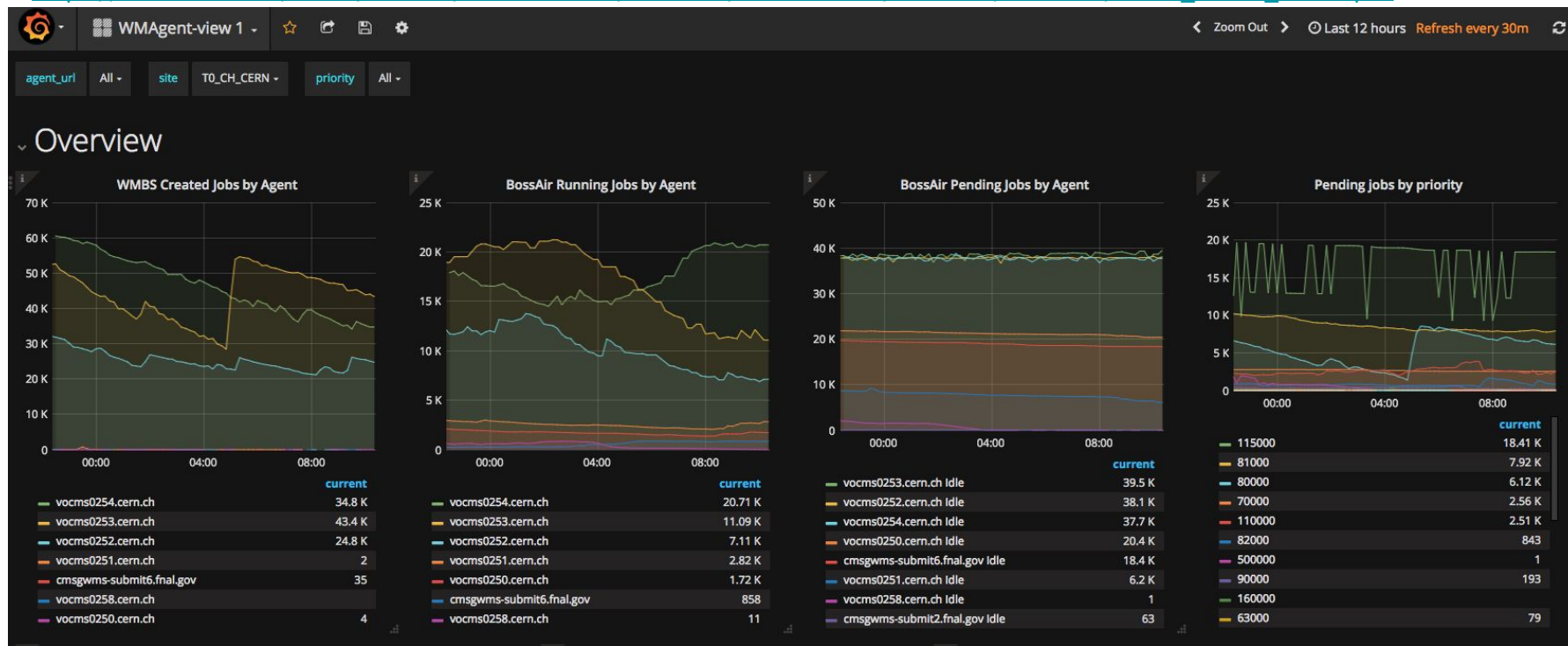# WMArchive: Jobs per sites
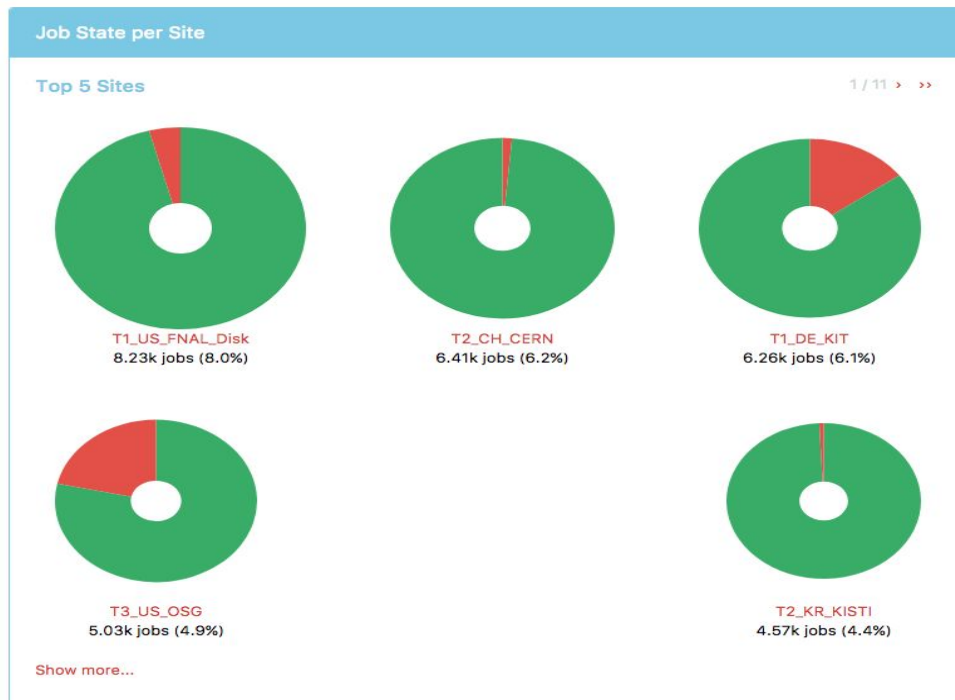
WMArchive: performance matrix

WMAgent Monitor (Stats/Alarms in short term)
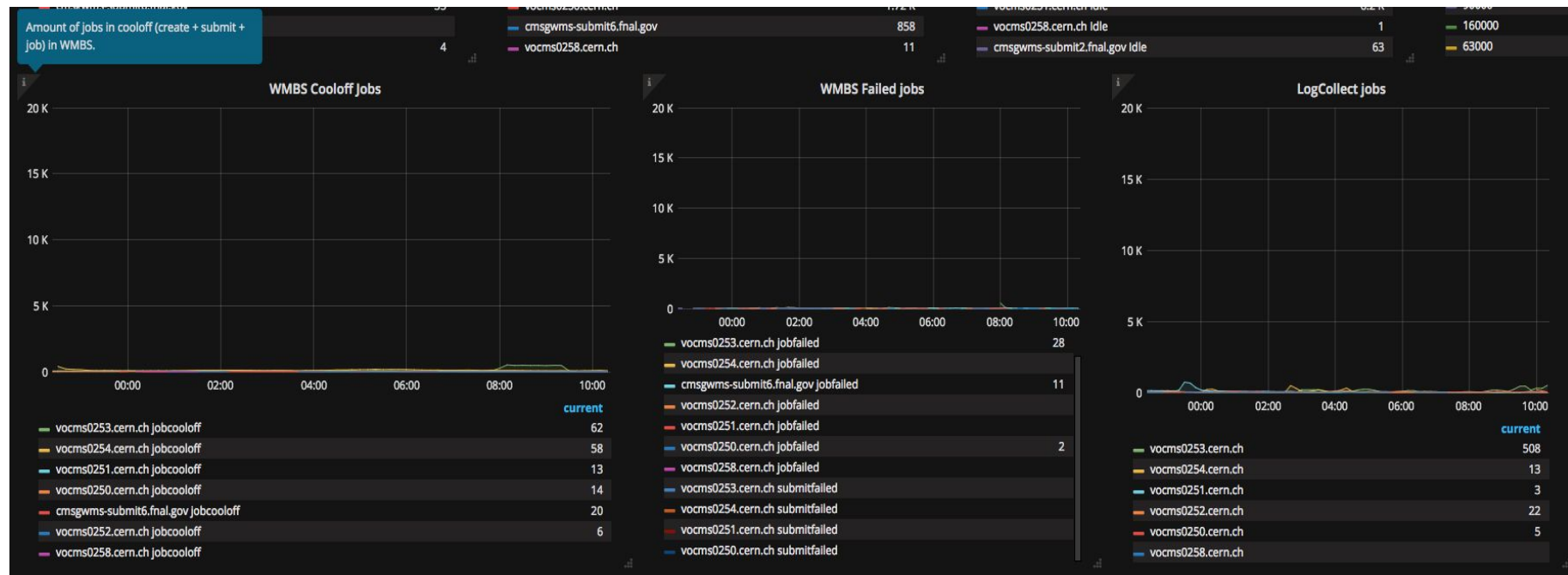
See talk from Fall16 O&C:

https://indico.cern.ch/event/581102/contributions/2369013/attachments/1369212/2075742/WMA_Monit_Fall16.pdf
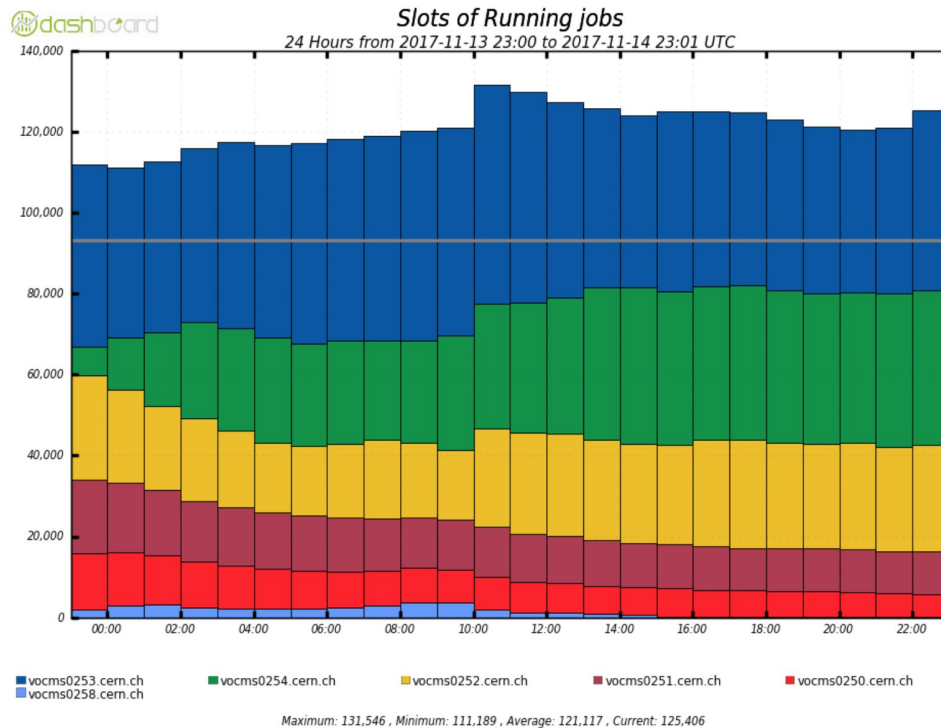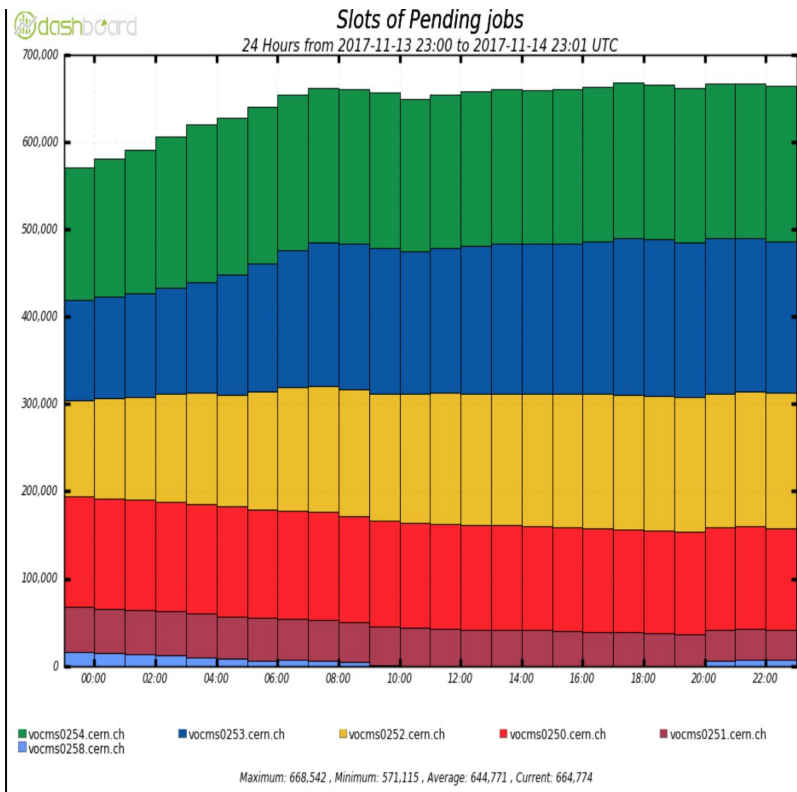
WMAgent Monitor (Stats/Alarms in short term)

WMAgent Monitor (Stats/Alarms in short term)

Dashboard (Agent monitor)

Slots of Pending jobs
24 Hours from 2017-11-13 23:00 to 2017-11-14 23:01 UTC

Maximum: 668,542 , Minimum: 571,115 , Average: 644,771 , Current: 664,774

Slots of Running jobs
24 Hours from 2017-11-13 23:00 to 2017-11-14 23:01 UTC

Maximum: 131,546 , Minimum: 111,189 , Average: 121,117 , Current: 125,406
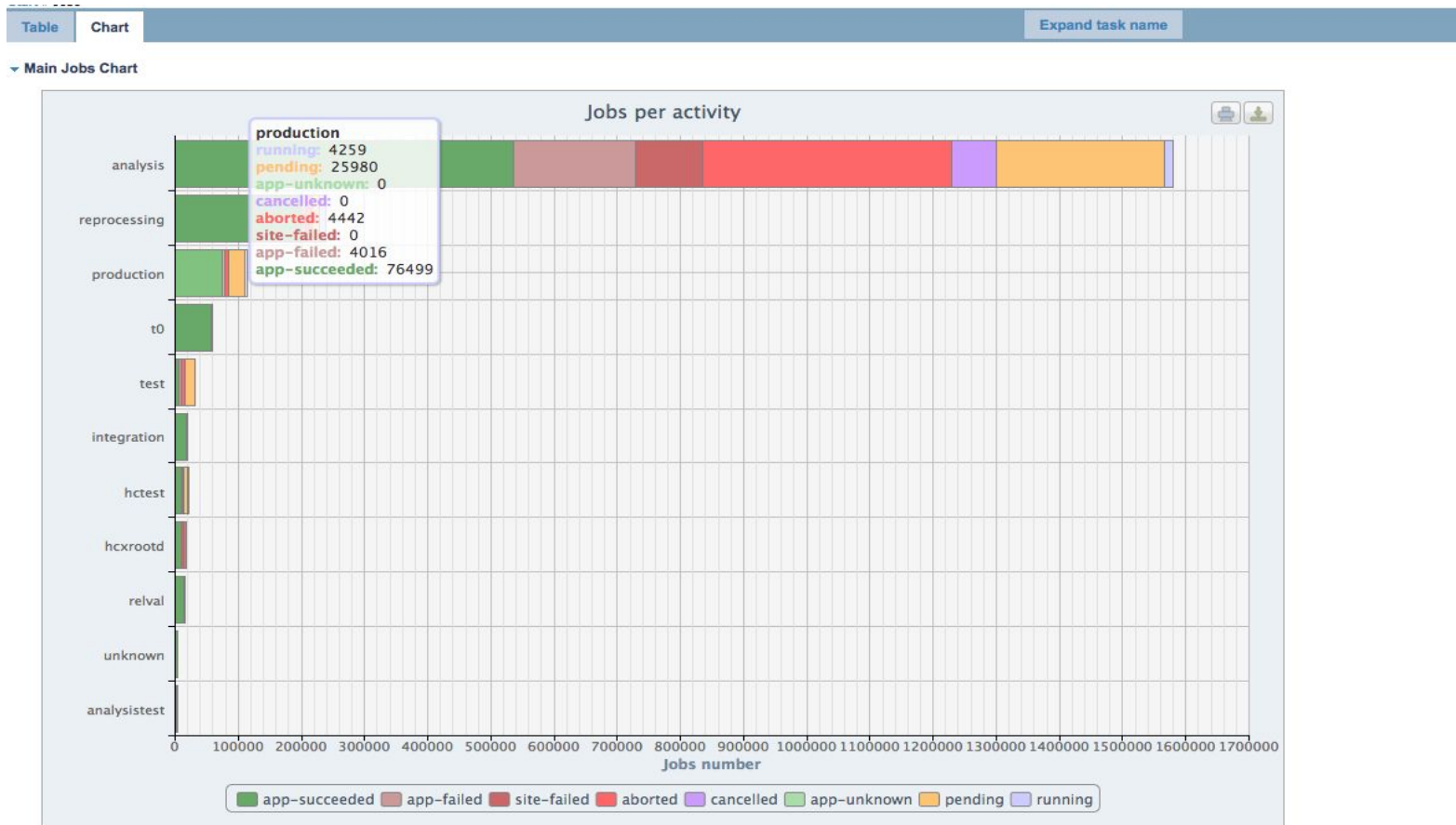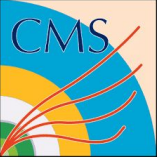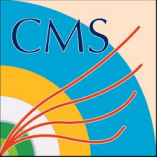
Dashboard (request monitor)

Other monitors (Unified, Prodmon, GMSMon, ElasticSearch)

Some of these tools can be integrated with other applications and should get an official support. (not a complete list)
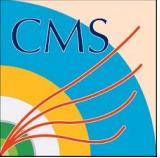
1. Unified monitoring: Data operation supporting tool, Identify the problems and report. Also provide link for all the relevant monitoring applications (https://vocms049.cern.ch/unified/)

2. Prodmon: Provide statistics for overall request in different catagories in given time period (https://dmytro.web.cern.ch/dmytro/cmsprodmon/)

3. GMSMon: Provide the statistics base on the data from condor global pool (https://cms-gwmsmon.cern.ch/)

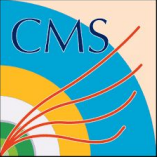4. ElasticSearch CMS: https://es-cms.cern.ch/kibana/app/kibana

Future plans

1. Need to survey currently used monitoring tools and provide single entry point with a description about scope and usage for each tools.

2. Identify duplicate features and combine/streamline tools for better support (long term)

3. Automate some of the data operational work. (Identify problems, report to the proper parties with helpful information)

Midterm development plan

1. Improve alarms in WMStats: (stuck workflow, failed workflow, site failure - port from unified monitor)

2. Add more error/warning logs during the lifetime of the request in wmstats

3. Provide access to the log files (condor log, error, out)

4. Improve debugging by retrieving more accurate information for failed job logs

5. Add the monitor with site perspective. (WMAgent Monitor - data is already collected)

6. WMArchive plan - next presentation

References

1. WMStats
   https://cmsweb.cern.ch/wmstats/
2. WMArchive
   https://cmsweb.cern.ch/wmarchive/web/performance
   https://monit-kibana.cern.ch/kibana/goto/5390243b9703ef913d61e9bc91a0c125
   https://monit-kibana.cern.ch/kibana/goto/8fc609f77944a1f378852cd65a0961f5
3. WMAgent monitor
   https://monit-grafana.cern.ch/dashboard/db/wmagent-view-1?orgId=11&from=now-1d&to=now&refresh=30m
4. Dashboard
   https://dashb-cms-job.cern.ch/dashboard/templates/web-job2/
4. CSP report instruction
   https://twiki.cern.ch/twiki/bin/view/CMS/CompOpsPRCSP