Jenkins & CMS Build Infrastructure

Giulio Eulisse

Problem

- CMS Offline SW (CMSSW) consists of 6M SLOCs, roughly 150 active users per month.

- Same amount of code from roughly 100 externals, which we keep under our control, from gcc / glibc up (the only things we do not compile ourself are basically the kernel and X11).

- 13 active release cycles, 4 active architectures. Up to 40 releases / day.

- All of this using Git & GitHub: as part of the migration we agreed to tests PR for users in order to have them surrender their “commit everywhere” rights. 400 PRs per month need to be tested (sometimes multiple times).

Driving CI: Jenkins

- An extensible open source Continuos Integration (CI) server: http://jenkins-ci.org. Opensource, Java based, but works well also for C++ projects. Literally hundreds of plugins and wide industry support.

- Initially used just to build releases, now used as a general batch system for release engineering and deployment:
    - Maintenance jobs (clean up jobs, web page generation)
    - Day by day integration builds, validation & QA
    - Pull Request testing & QA
    - (Semi) Automated release building
- Scheduling jobs on a 250 cores, CERN OpenStack, cluster.

New Item
People
Build History
Manage Jenkins
Credentials
My Views
Disk usage
Job Config History
Scriptler
Jenkins 100K

Build Queue (10)

CMS Github Bot
ib-schedule-pr-tests
update-github-pages
ib-any-integration
ib-any-integration
ib-any-integration
ib-any-integration
ib-any-integration
ib-any-integration
ib-any-integration
Backup Jenkins

Build Executor Status

master

1 Idle
2 Idle

add description

All | Build Status | IB Building Status | IB Pipeline | IBs | Integration Status | Plumbing status | +

S	W	Name ↓	Last Success	Last Failure	Last Duration	Built On
		afs-install-package	N/A	N/A	N/A	
		Backup Jenkins	4 mo 18 days - #449	N/A	35 sec	
		baseline-ib-results-testing	N/A	1 day 11 hr - CMSSW_7_5_X_2015-03-16-0200 -- slc6_amd64_gcc491	26 min	cmsbuild19
		build-any-ib	4 hr 58 min - CMSSW_7_5_X_2015-03-18-1400 - slc6_amd64_gcc491	4 hr 58 min - CMSSW_7_5_DEVEL_X_2015-03-18-1400 - slc6_amd64_gcc491	1 hr 12 min	cmsbuild31
		build-release	9 hr 52 min - #CMSSW_7_3_5 - #slc6_amd64_gcc491 Issue #8372 ONLY TOOLCONF:false	11 hr - #228	7 hr 13 min	cmsbuild34
		build-release-testing	8 days 7 hr - #CMSSW_7_1_50 - #slc6_amd64_gcc481 Issue #6375 ONLY TOOLCONF:false	8 days 9 hr - #31	41 sec	cmsbuild19
		BUILDEXT	1 yr 3 mo - cms-sw:IB/CMSSW_7_0_X/root6 EXT for slc5_amd64_gcc481	1 yr 3 mo - cms-sw:IB/CMSSW_7_0_X/root6 EXT for slc5_amd64_gcc481	1 hr 11 min	lxbuild169
		cleanup-auto-build	2 days 11 hr - #CMSSW_6_2_0_SLHC25 - #slc6_amd64_gcc472 Issue #8269	5 days 6 hr - #CMSSW_6_2_0_SLHC25 - #slc6_amd64_gcc472 Issue #8230	5 min 5 sec	
		cleanup-cmssdt	20 hr - #540	1 yr 1 mo - #131	12 min	cmssdt
		cleanup-elasticsearch	8 min 5 sec - #12158	1 day 9 hr - #12024	37 sec	mesos-jenkins-a02286e5-bc8d-41ac-85b0-3fa56a1ce486
		cleanup-tags	20 hr - #210	N/A	15 min	
		CMS Github Bot	3 min 41 sec - #160237	1 day 10 hr - #159906	2 min 47 sec	mesos-jenkins-258b8055-2220-42a2-b4c5-bfbeb5e4302f
		cms-bot-testing	2 mo 3 days - #8	N/A	18 sec	
		CMSSW Valgrind tests	1 yr 6 mo - CMSSW_7_0_X_2013-09-10-0200	1 yr 6 mo - CMSSW_7_0_X_2013-08-28-0200	1 hr 0 min	vocms13

Driving CI: Jenkins

- Simplifies the creation of complex workflows. Jobs have dependencies and can trigger different behavior depending on results of the previous step.

- Simplifies access to logs and keeps history of what happens.

- Simplifies access to build infrastructure to newcomers. Pointing them to a Jenkins “Job” page seems to be much easier than having them look at some script.

- Provides scheduling, structuring and monitoring of jobs.

Driving CI: cms-bot

- Actual payload scripts are actually maintained in a GitHub repository rather then in Jenkins itself.
- For “multistage” behavior, we use comments in GitHub issues to drive integration.
  - Pull request approval process is updated by a Jenkins job which keeps track of the +1 / -1 by coordinators in GitHub PR comments.
  - Automated release building steps (request, build, upload, announce) are also tracked as comments in a GitHub Issue.
- Using GitHub comments as state tracker for our bot allows us to avoid a private integration state tracking DB.
- Using GitHub PR labels to show current state for various PRs.

134 Open ✔ 7,965 Closed Author Labels Milestones Assignee Sort

HCAL Physim pedestal2015 ✖ — 12
alca-pending comparison-pending operations-pending orp-pending pending-signatures tests-rejected
#8350 opened a day ago by kodolova Next CMSSW_7_5_X

bsunanda:Run2-alca6 Make the last changes for AlCaReco producer and tests for IsoTrack trigg.. ✔ — 16
alca-pending comparison-available dqm-pending orp-pending pending-signatures reconstruction-approved tests-approved
#8347 opened 2 days ago by bsunanda Next CMSSW_7_5_X

Improved error reporting in ConfigToolBase (75X) ✔ — 3
comparison-available core-pending orp-pending pending-signatures tests-approved
#8340 opened 2 days ago by ferencek Next CMSSW_7_5_X

Add variables to StoppedParticles 71X ✔ — 6
comparison-pending fully-signed orp-pending simulation-approved tests-approved
#8339 opened 2 days ago by jalimena Next CMSSW_7_1_X

StoppedParticles Bug Fix for 71X ✔ comparison-available fully-signed orp-pending simulation-approved tests-approved — 9
#8337 opened 2 days ago by jalimena Next CMSSW_7_1_X

IPProducer update: Part 2 (75X) ✖ comparison-pending fully-signed orp-pending reconstruction-approved tests-started — 10
#8329 opened 2 days ago by ferencek Next CMSSW_7_5_X

Removing Duplications of some DD Algos ✔ — 3
comparison-available geometry-pending orp-pending pending-signatures tests-approved
#8327 opened 2 days ago by boudoul Next CMSSW_7_5_X

Harmonize the use of reco::TrackBase::TrackAlgorithm enumeration ✔ — 9
analysis-pending comparison-available dqm-pending orp-pending pending-signatures reconstruction-approved tests-approved
#8315 opened 3 days ago by makortel Next CMSSW_7_5_X

update list of paths monitored by the Higgs HLT DQM (backport PR#8313) ✔ — 5
comparison-available dqm-pending orp-pending pending-signatures tests-approved
#8314 opened 3 days ago by HuguesBrun Next CMSSW_7_4_X

update list of paths monitored by the Higgs HLT DQM ✔ — 5
comparison-available dqm-pending orp-pending pending-signatures tests-approved
#8313 opened 3 days ago by HuguesBrun Next CMSSW_7_5_X

Build CMSSW_7_3_5 #8372

Open **davidlange6** opened this issue 14 hours ago · 22 comments

davidlange6 commented 14 hours ago — Owner

Integrating bug fixes and requests for CRAFT operations

cmsbuild commented 14 hours ago — Owner

Request received. I will start to build the release after one of the following approve the issue: **@nclopezo**, **@ktf**, **@smuzaffar**, **@degano**, **@davidlange6**. You can do this by writing "+1" in a comment.
You can also ask me to begin to build cmssw-tool-conf first (Cannot be done for patch releases). To do this write "build cmssw-tool-conf" in a comment. I will start to build cmssw-tool-conf and then wait for the "+1" to start the build of the release.

@smuzaffar, **@nclopezo**, **@davidlange6**, **@degano** you requested to watch the automated builds for CMSSW_7_3_X

cmsbuild added the **build-pending-approval** label 14 hours ago

davidlange6 commented 13 hours ago — Owner

+1

cmsbuild commented 13 hours ago — Owner

Release created: https://github.com/cms-sw/cmssw/releases/tag/CMSSW_7_3_5

cmsbuild added **slc6_amd64_gcc491-build-queued** **osx108_amd64_gcc481-build-queued** **slc6_amd64_gcc481-build-queued** and removed **slc6_amd64_gcc491-tool-conf-ok** **osx108_amd64_gcc481-tool-conf-ok** **slc6_amd64_gcc481-tool-conf-ok** labels 13 hours ago

Dealing with load

- Some tasks are bigger then others: non-homogeneous load => static partitioning of build machines hits us.
    - One alternative is to create fake builders, so that short lived, trivial tasks are executed “out of band” WRT long tasks. This leads to maintenance burden.
    - **Apache Mesos** to the rescue. Jenkins can create differently sized slaves on a Mesos Cluster, dynamically distributing payloads. This solution has also the advantage that Mesos is really a must for long running services (e.g. Elasticsearch) as well.

CMS Build Infrastructure

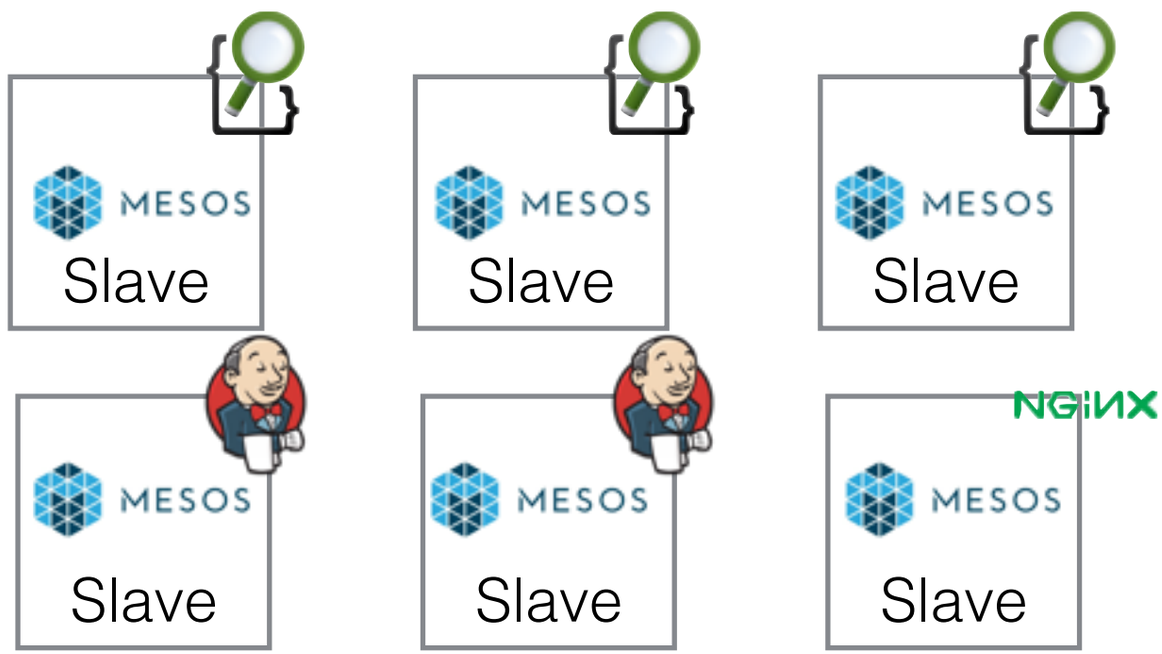
Frontend with HA setup using CERN LB DNS, nginx for SSL termination and authorisation, haproxy for traffic routing and SSO backend

Resource arbitration via 3-way redundant **Apache Mesos** setup, using different OpenStack zones, leader election via Zookeeper (1 dead master resilience)

Service Discovery via DNS, populated with A and SRV records discovery by mesos registry information (ala Consul) or using **Marathon Framework** REST API.

Services run on undifferentiated CPU boxes, either running on the bare OS or running inside **Docker**

Services which we run varies from **Jenkins** build slave, to **web server backends** or **Elasticsearch**



All the services are being restarted automatically by Marathon whenever they die on machines that offer a compatible set of resources. Looking forward dynamic resource allocation (i.e. persistent disk storage on slaves) to simplify setup even further.

Dealing with multiple archs

- Another problem is that sometimes we need to provide support for both SLCX and SLC(X+1).

- Sometimes is desirable to migrate infrastructure to a new platform before Offline SW provides support for it.

- Again, Apache Mesos and Docker come to the rescue. Builders are created on the fly, using a special Docker container (e.g. based on SLC5). We can therefore decouple migration of the infrastructure from the migration of job environment.

Monitoring

- Jenkins provides some minimal level of monitoring / logging:
  - Job statistics
  - Slave statistics
  - Build log (possibly parsed via various plugins)
- Now that Mesos allows us to maintain an Elasticsearch cluster for “free”, we have started to push more and more information into it to allow data mining via the Opensource “Kibana” dashboard.

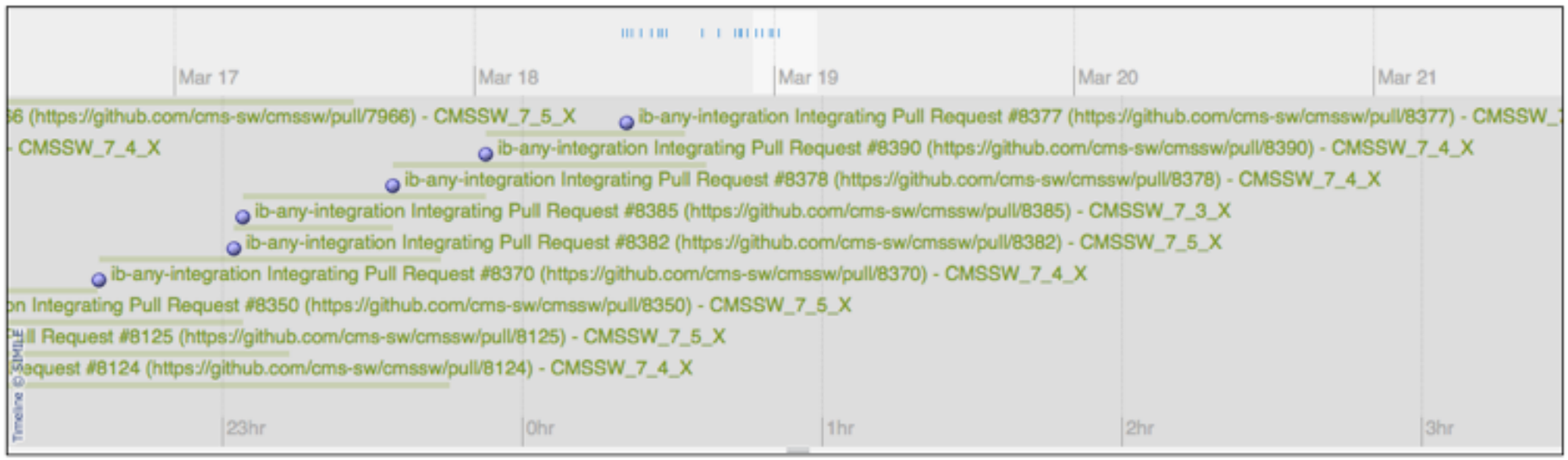
Monitoring: Jenkins

- Back to Dashboard
- Status
- Changes
- Workspace
- Build with Parameters
- Delete Project
- Configure
- Parameterized Builds Report
- Rebuild Last
- Job Config History
- GNU Compiler Warnings

Build History trend

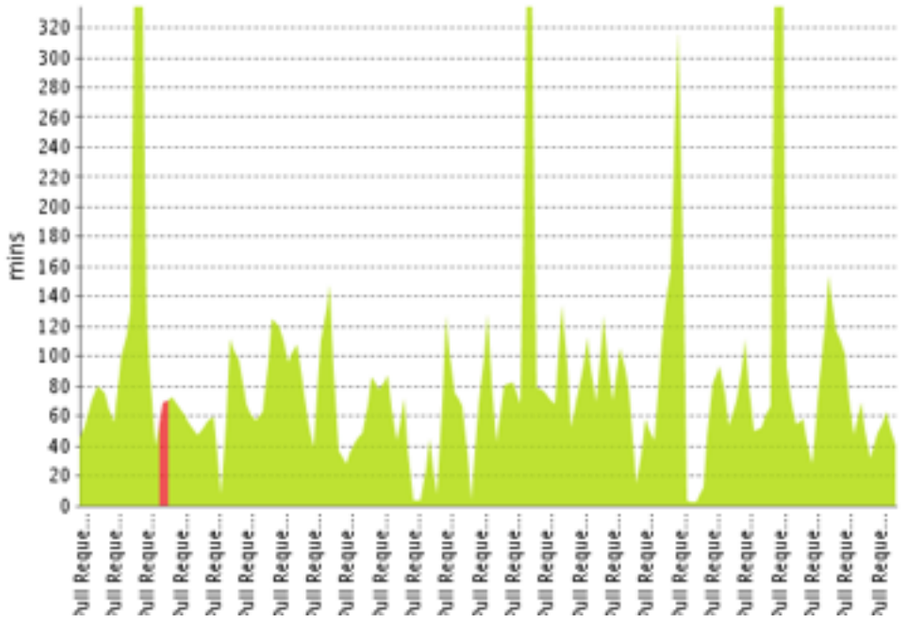
RSS for all RSS for failures

Timeline

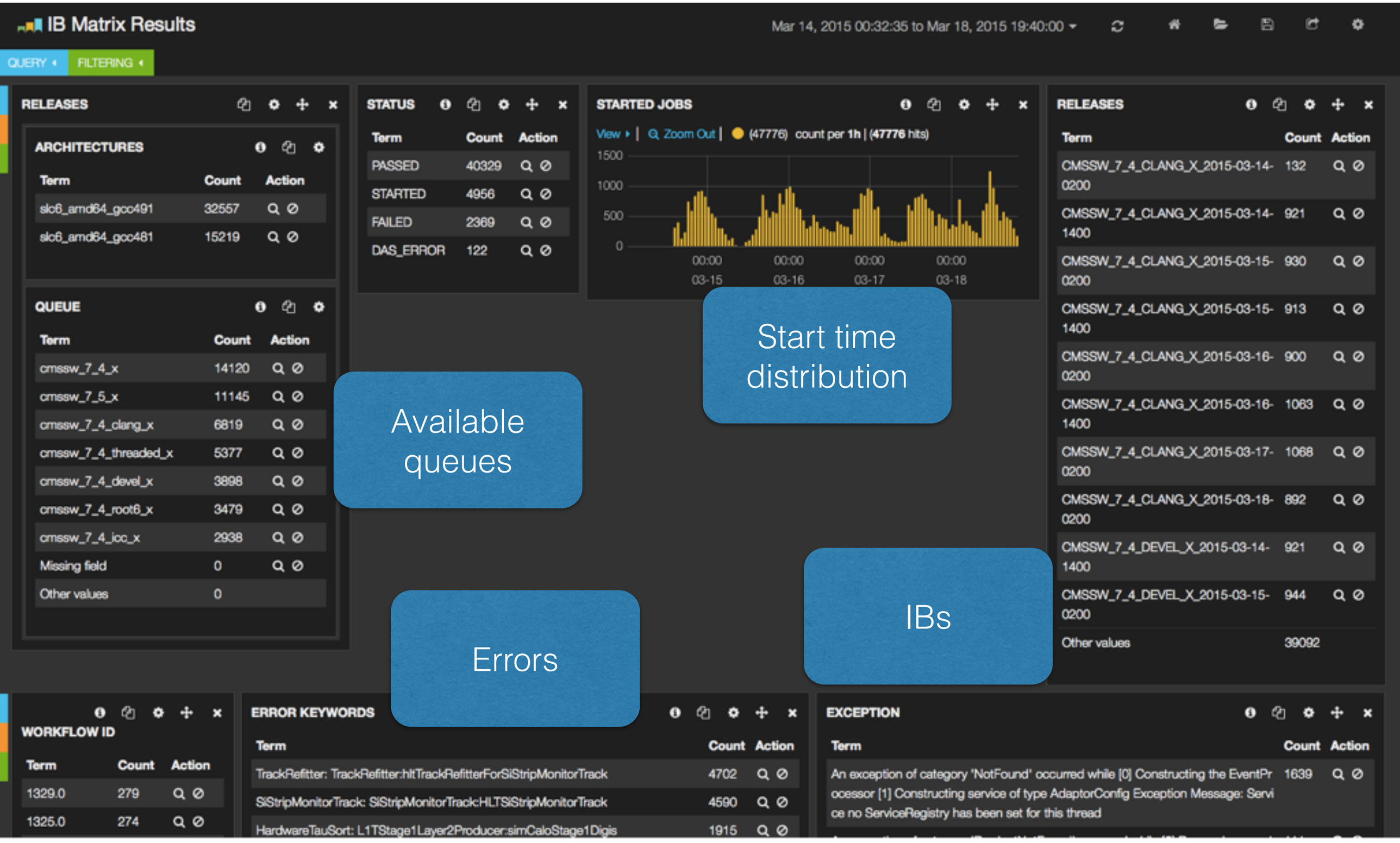


Build Time Trend

Build ↑	Duration	Slave
Integrating Pull Request #8377 (https://github.com/cms-sw/cmssw/pull/8377) - CMSSW_7_4_X	31 min and counting	cmsbuild21
Integrating Pull Request #8390 (https://github.com/cms-sw/cmssw/pull/8390) - CMSSW_7_4_X	40 min	cmsbuild20
Integrating Pull Request #8378 (https://github.com/cms-sw/cmssw/pull/8378) - CMSSW_7_4_X	1 hr 2 min	cmsbuild15
Integrating Pull Request #8385 (https://github.com/cms-sw/cmssw/pull/8385) - CMSSW_7_3_X	48 min	cmsbuild20
Integrating Pull Request #8382 (https://github.com/cms-sw/cmssw/pull/8382) - CMSSW_7_5_X	31 min	cmsbuild15
Integrating Pull Request #8370 (https://github.com/cms-sw/cmssw/pull/8370) - CMSSW_7_4_X	1 hr 8 min	cmsbuild18
Integrating Pull Request #8350 (https://github.com/cms-sw/cmssw/pull/8350) - CMSSW_7_5_X	46 min	cmsbuild18
Integrating Pull Request #8125 (https://github.com/cms-sw/cmssw/pull/8125) - CMSSW_7_5_X	1 hr 41 min	cmsbuild20
Integrating Pull Request #8124 (https://github.com/cms-sw/cmssw/pull/8124) - CMSSW_7_4_X	1 hr 57 min	cmsbuild30



Monitoring: ES & Kibana



Presenting results

- Results for general consumption are created as static views from different sources (Elasticsearch, log tarballs on AFS, etc) and published routinely (via Jenkins) as “GitHub Pages”.
  - Allows us to exploit the GitHub CDN for most of the web pages. Big data files still kept at CERN.
  - Allows us to keep historical information at minimal cost (do not underestimate git ability to compress very similar, day by day, data).

CMSSW_7_4_X_2015-03-18-1400

IB Tag

Architectures	Builds	Unit Tests	RelVals	Other Tests	Q/A
slc6_amd64_gcc491 Patch from 2015-03-16-0200	See Details	Unknown			
slc6_amd64_gcc481 Patch from 2015-03-15-0200	See Details	See Details	Not complete	See Details	

See comparison with CMSSW_7_4_X_2015-03-18-0200 on GitHub

- #8364 from wmtan: Port back from ROOT6 IB for commonality (Sim)
- #8365 from wmtan: Port back from ROOT6 IB for commonality (L1)

CMSSW_7_4_X_2015-03-18-0200

IB Tag
Static Analyzer
Modules to thread unsafe statics
Modules to thread unsafe EventSetup products
HLT Validation
Relvals Exceptions Summary

Architectures	Builds	Unit Tests	RelVals	Other Tests	Q/A
slc6_amd64_gcc491 Patch from 2015-03-16-0200	2 Warnings	See Details	Pass: 1056 Fail: 15	See Details	
slc6_amd64_gcc481 Patch from 2015-03-15-0200	See Details	See Details	Pass: 1075 Fail: 8	See Details	

See comparison with CMSSW_7_4_X_2015-03-17-1400 on GitHub

- #8301 from venturia: Fix in two scripts to monitor the SiStrip DB objects
- #8344 from slava77: trim hep top tagger combinatorics
- #8361 from wmtan: Port back from ROOT6 IB for commonality (Core)
- #8363 from wmtan: Port back from ROOT6 IB for commonality (DQM)
- #8343 from cms-tsg-storm: CCC PSet migration with HLT for 74X
- #8208 from arizzi: MINIAOD: reducing packedMETunceratinty to float16 on top of 8159

CMSSW_7_4_X_2015-03-17-1400

IB Tag

Architectures	Builds	Unit Tests	RelVals	Other Tests	Q/A
slc6_amd64_gcc491 Patch from 2015-03-16-0200	1 Warnings	See Details	Pass: 1055 Fail: 15	See Details	
slc6_amd64_gcc481	See Details	See Details	Pass: 1065 Fail: 12	See Details	

CMSSW_7_5_X

Back to IB pages

CMSSW_7_5_X_2015-03-18-0200

See all results for CMSSW_7_5_X_2015-03-18-0200

Exception	Workflows

```
An exception of category 'Configuration' occurred while
   [0] Processing run: 1 lumi: 47 event: 4601
   [1] Running path 'digitisation_step'
   [2] Calling event method for module MixingModule/'mix'
Exception Message:
RootInputFileSequence::readOneRandom(): no input files specified for seconda
ry input source.
```

slc6_amd64_gcc481:

```
202.0
```

```
An exception of category 'LogicError' occurred while
   [0] Processing run: 1
   [1] Calling global beginRun for module Pythia8HadronizerFilter/'generator
'
Exception Message:
::getByToken: An attempt was made to read a Run product before endRun() was
called.
The index of the token was 1.
```

slc6_amd64_gcc481:

```
521.0, 523.0, 519.0, 525.0, 513.0,
516.0, 528.0
```

```
An exception of category 'ProductNotFound' occurred while
   [0] Processing run: 1 lumi: 1 event: 1
   [1] Running path 'reconstruction_step'
   [2] Calling event method for module GoodSeedProducer/'trackerDrivenElectr
onSeedsTmp'
Exception Message:
Principal::getByToken: Found zero products matching all criteria
```

slc6_amd64_gcc481:

```
400.0, 401.0, 25400.2, 25400.0, 254
01.0
```