

Bringing the ATLAS HammerCloud setup to the next level with containerization

Benjamin Rottler, Michael Böhler, Günter Duckeck
Alexander Lory, Christoph Anton Mitterer,
Jaroslava Schovancova

on behalf of the ATLAS Computing Activity



CHEP 2023 – Norfolk, Virginia, USA
09 May 2023



FSP ATLAS
Erforschung von
Universum und Materie

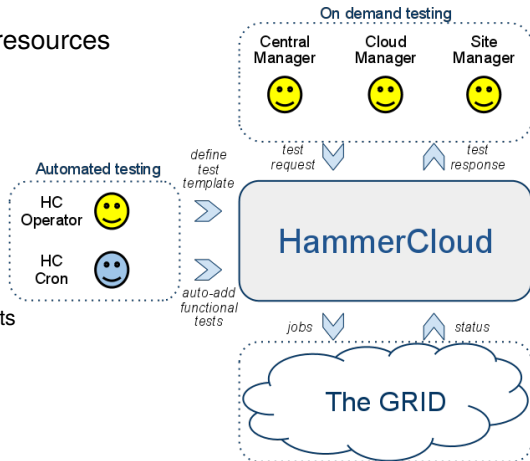


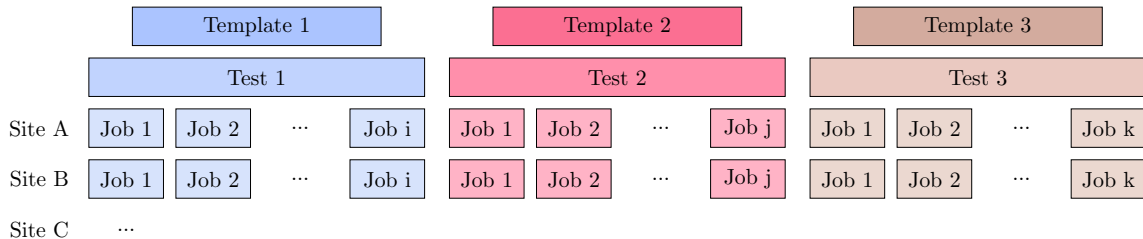
Bundesministerium
für Bildung
und Forschung



- ▶ framework to test, commission, benchmark WLCG resources
- ▶ testing with **full chain** jobs
 - ▶ same environment as the “real” analysis/production jobs
- ▶ part of the ATLAS distributed computing operations (ADC Ops) automation suite
 - ▶ automatic exclusion & recovery of the resources
 - ▶ spot system-wide outages & alerting
 - ▶ spot issues with the next generation clients / environments
 - ▶ commission new resources & distributed system components
 - ▶ benchmark resources & workflows

→ <http://hammercloud.cern.ch/hc/app/atlas/>

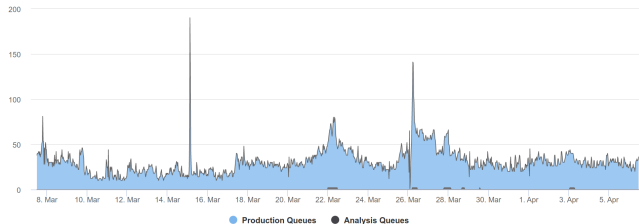




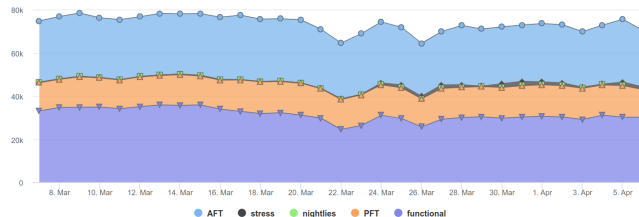
- ▶ template: defines test type (sites, workload, auto-exclusion)
- ▶ test runs 24 hours on all sites defined in the template
- ▶ jobs are continuously sent to each site until the end of the test

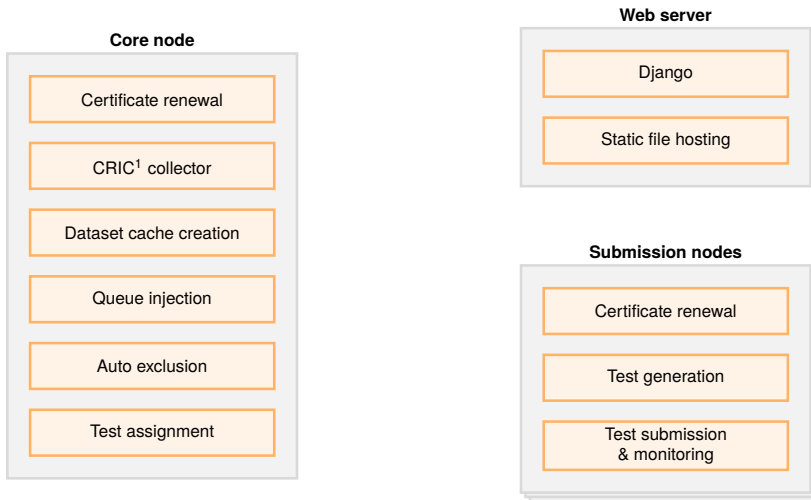
- ▶ spots issues with infrastructure
- ▶ helps with commissioning and R&D
 - ▶ resources
 - ▶ components of distributed systems
 - ▶ new approaches to data access/management
 - ▶ GPUs
- ▶ in total 60k+ jobs per day

Number of excluded queues per day between March 7 and April 6



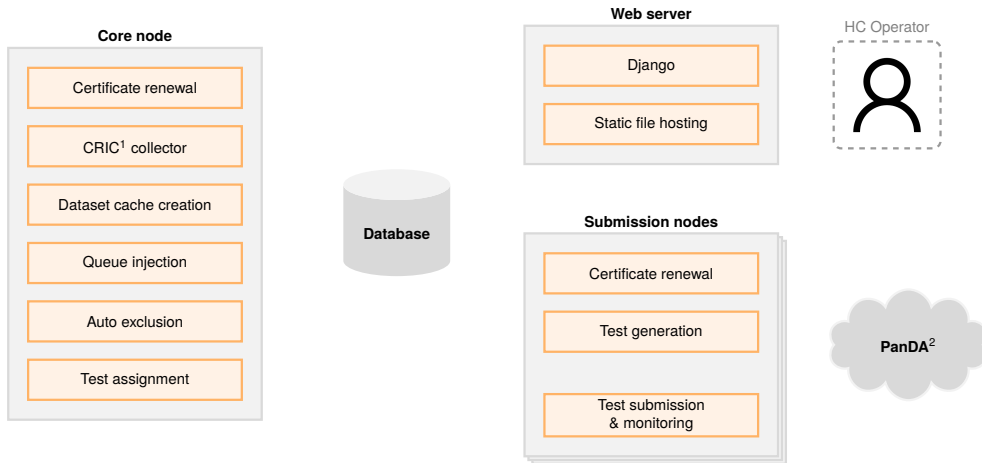
Number of submitted jobs per day between March 7 and April 6





- ▶ three types of nodes: core (1x), web server (1x), submission nodes (multiple)

¹Computing Resource Information Catalog

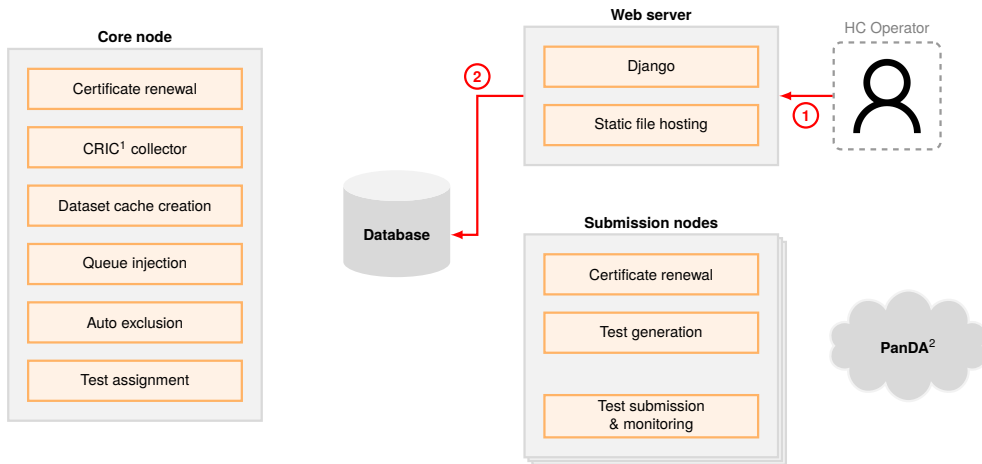


¹Computing Resource Information Catalog

Benjamin Rottler

²Production and Distributed Analysis – workload management system

CHEP 2023 – ATLAS HammerCloud Containerization



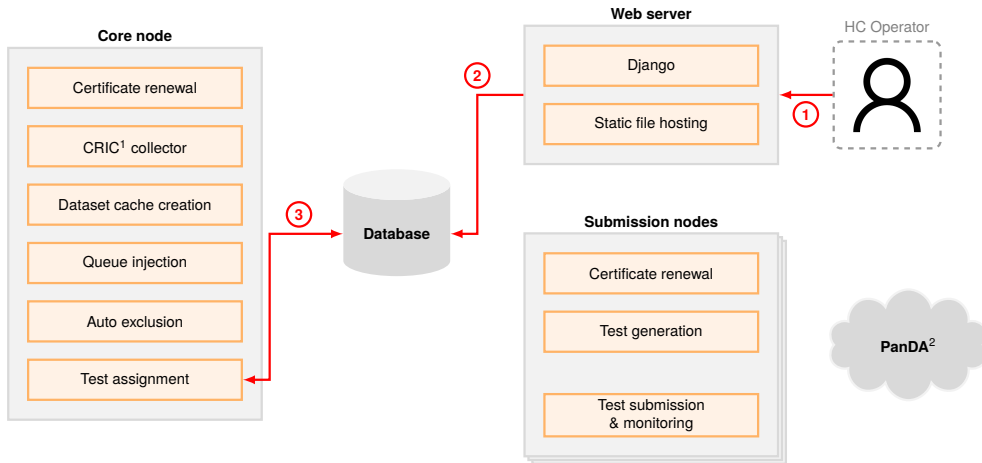
1.+2. create test

¹Computing Resource Information Catalog

Benjamin Rottler

²Production and Distributed Analysis – workload management system

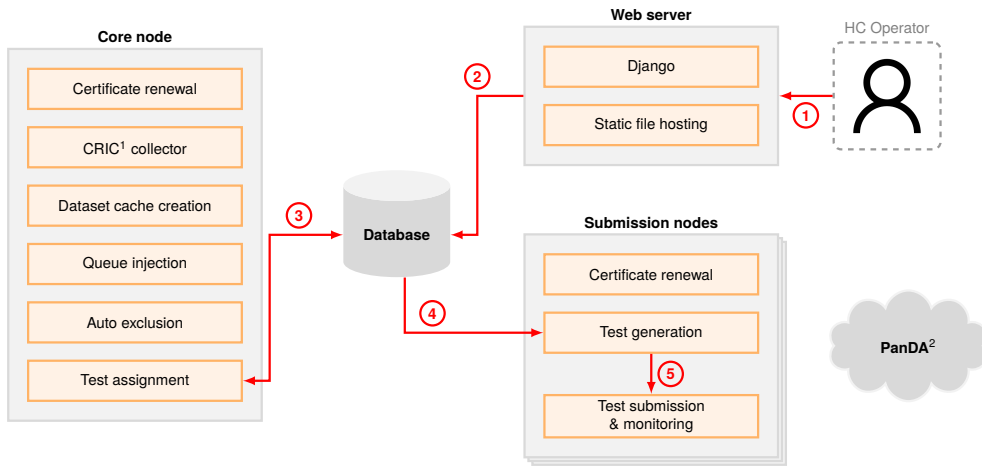
CHEP 2023 – ATLAS HammerCloud Containerization



1.+2. create test

3. assign test to submission node

¹Computing Resource Information Catalog ²Production and Distributed Analysis – workload management system



1.+2. create test

3. assign test to submission node

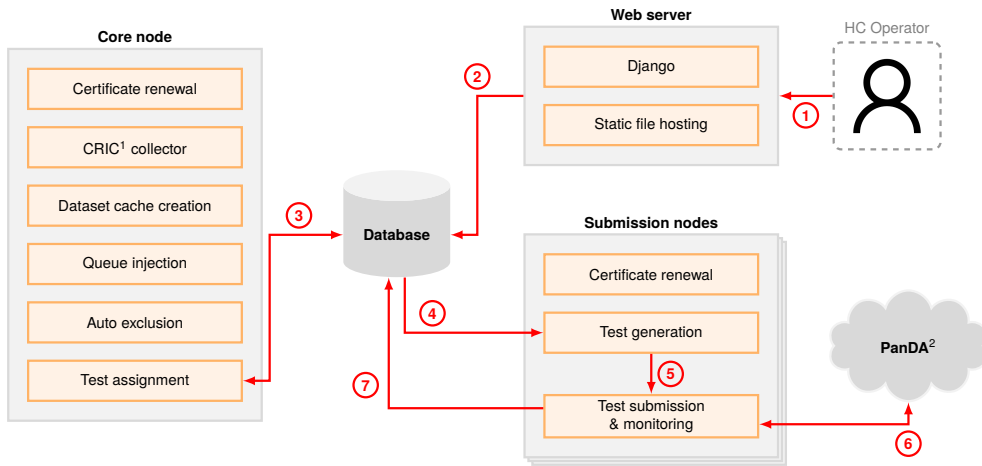
4.+5. generate config files for PanDA jobs

¹Computing Resource Information Catalog

Benjamin Rottler

²Production and Distributed Analysis – workload management system

CHEP 2023 – ATLAS HammerCloud Containerization



1.+2. create test

3. assign test to submission node

4.+5. generate config files for PanDA jobs

6.+7. submit and monitor PanDA jobs

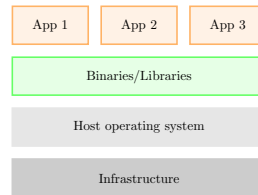
¹Computing Resource Information Catalog

Benjamin Rottler

²Production and Distributed Analysis – workload management system

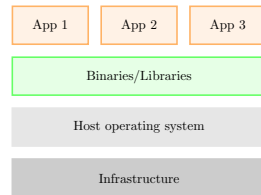
Current setup

- ▶ Python 2.7 / Django 1.1
- ▶ no containerization – binaries/libraries
 - depend on host operating system
 - are shared between apps/services
- ▶ HammerCloud software distributed via RPMs
- ▶ dedicated machines needed for development



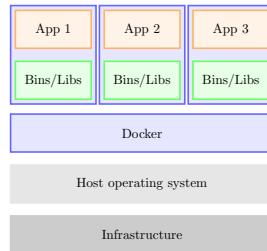
Current setup

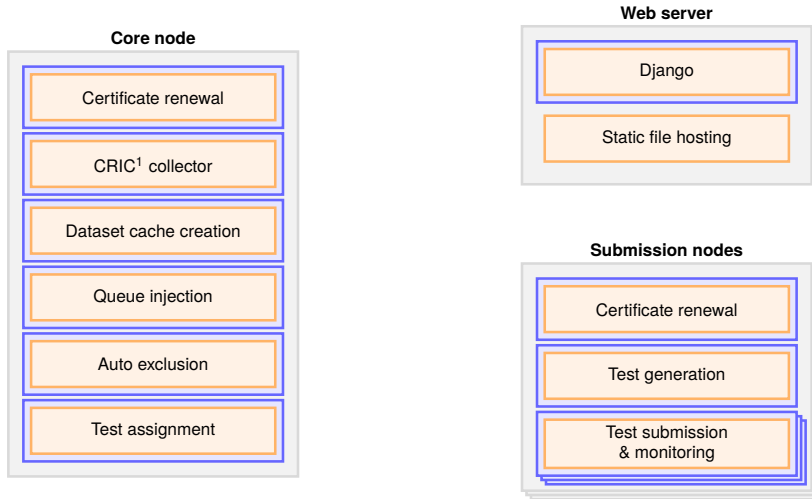
- ▶ Python 2.7 / Django 1.1
- ▶ no containerization – binaries/libraries
 - depend on host operating system
 - are shared between apps/services
- ▶ HammerCloud software distributed via RPMs
- ▶ dedicated machines needed for development



Containerized setup

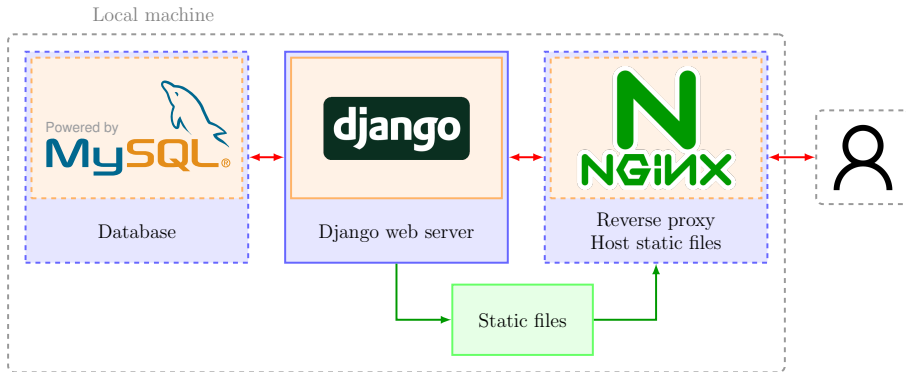
- ▶ migrate codebase to Python 3.8 / Django 4.0 (almost done)
- ▶ containerization via Docker
 - services can run in their own environment
- ▶ more agile development process
 - local development on laptop possible
 - application of CI/CD workflow
- ▶ long term goal: streamline commissioning of HPC sites
 - provide containerized HammerCloud setup
 - HPC sites can test their setup with a local HammerCloud instance



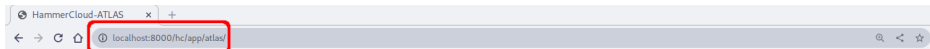


- docker philosophy: use separate container for each service

¹Computing Resource Information Catalog



- ▶ local container orchestration with `docker compose`
- ▶ include config and static files via bind mounts
→ persistent storage on host system
- ▶ allows one to operate the entire web server on a laptop



HammerCloud | ATLAS

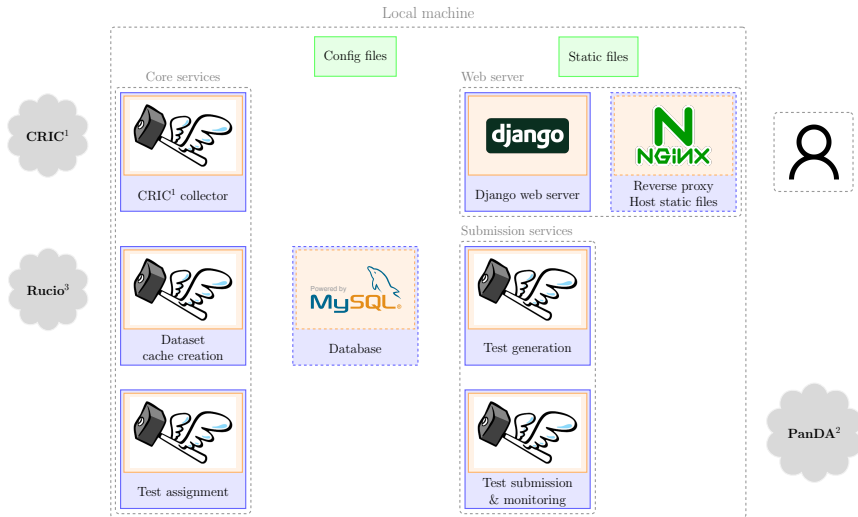
Home	Tests	Robot	Overviews	PanDA Dashb.	More HC...	Help	Administration
------	-------	-------	-----------	--------------	------------	------	----------------

Welcome to HammerCloud-ATLAS.

Running and Scheduled AFT/PFT Tests

State	Id	Host	Template	Start (Europe/Zurich)	End (Europe/Zurich)	Sites	subm jobs	run jobs	comp jobs	fail jobs	fail %	tot jobs
running	20252624	hammercloud-ai-74	1184: PFT mc15 Sim_tf 21.0.15 FC	21/Nov, 7:32	22/Nov, 5:50	AGLT2_TEST, BEIJING, GoeGrid, 162 more...	102	90	5890	1273	17	7355
running	20252627	hammercloud-ai-77	952: AFT Eventloop 21.2.1 Analy	21/Nov, 8:16	22/Nov, 7:23	ANALY_ARNES_DIRECT, ANALY_SIGNET_DIRECT, ANALY_CERN_T0_ART, 108 more...	76	22	8460	1018	11	9576
running	20252638	hammercloud-ai-79	1192: PFT mc21 Sim_tf 22.0.73	21/Nov, 15:14	22/Nov, 14:13	CERN, UNI-FREIBURG, AGLT2, 162 more...	52	86	1963	333	14	2434
running	20252640	hammercloud-ai-78	1013: AFT AthDerivation 21.2.33.0	21/Nov, 16:44	22/Nov, 17:38	ANALY_ARNES_DIRECT, ANALY_SIGNET_DIRECT, ANALY_CERN_T0_ART, 108 more...	46	53	2358	186	7	2643
running	20252645	hammercloud-ai-74	1169: GPU Container GPU Available + Vector Multiplication	22/Nov, 0:08	23/Nov, 0:58	ANALY_MANC_GPU, ANALY_BNL_GPU_ARC, ANALY_SLAC_GPU, 4 more...	1	3	6	7	41	17
running	20252646	hammercloud-ai-77	883: AFT PlottingJobOptions_ExampleCode 21.0.8	22/Nov, 0:46	23/Nov, 2:09	ANALY_ARNES_DIRECT, ANALY_SIGNET_DIRECT, ANALY_CERN_T0_ART, 108 more...	65	27	241	14	4	347

► HammerCloud website running locally on laptop

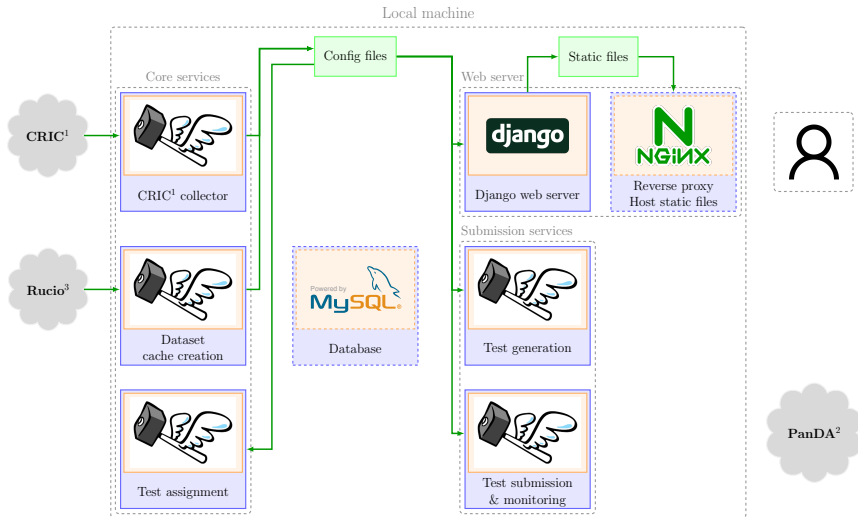


► minimal setup to submit HammerCloud jobs from a local machine (e.g. personal laptop) to PanDA

¹Computing Resource Information Catalog

²Production and Distributed Analysis – workload management system

³Data management tool

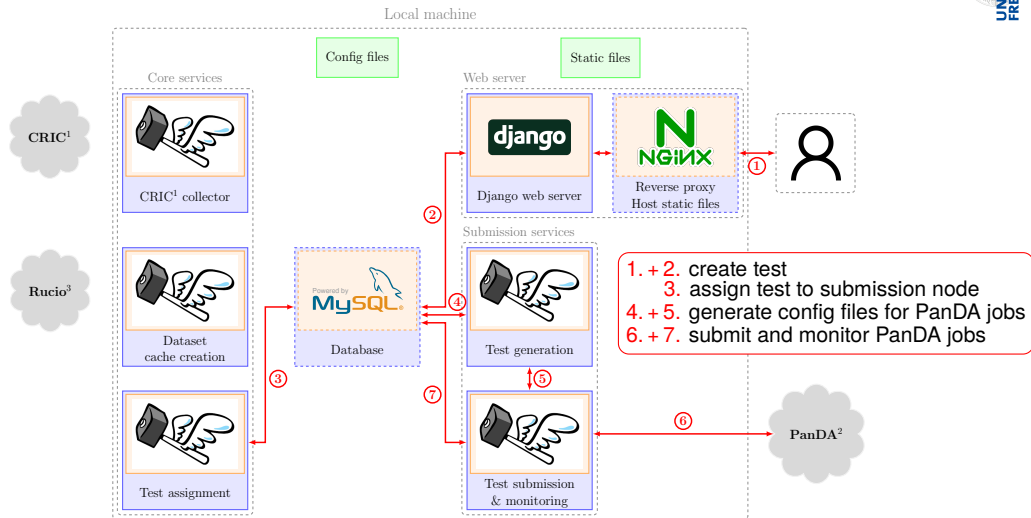


- minimal setup to submit HammerCloud jobs from a local machine (e.g. personal laptop) to PanDA

¹Computing Resource Information Catalog

²Production and Distributed Analysis – workload management system

³Data management tool



► minimal setup to submit HammerCloud jobs from a local machine (e.g. personal laptop) to PanDA

¹Computing Resource Information Catalog

²Production and Distributed Analysis – workload management system

³Data management tool

Containerization of HammerCloud

- ▶ more than 80 % of the code base has been migrated successfully to Python 3.8 / Django 4.0
- ▶ developed a suitable container architecture for HammerCloud infrastructure
 - ▶ this allows local development
 - ▶ more agile code deployment and CI
- ▶ web server completely containerized



Outlook

- ▶ migrate remaining HammerCloud components to Python 3.8 / Django 4.0
- ▶ containerize remaining services
- ▶ put containerized setup into production
- ▶ develop CI/CD workflow