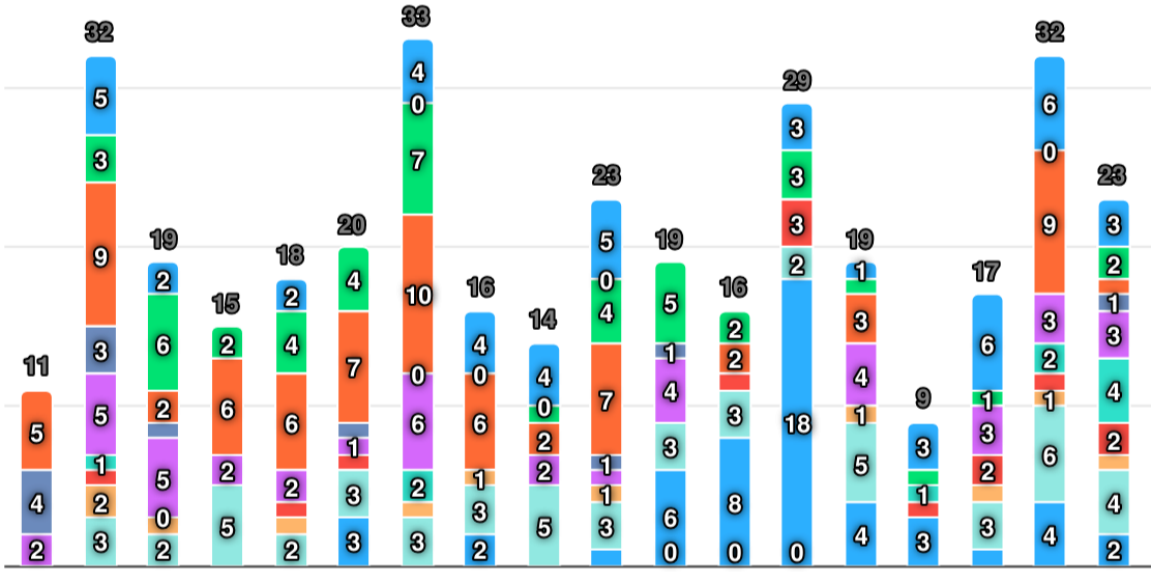
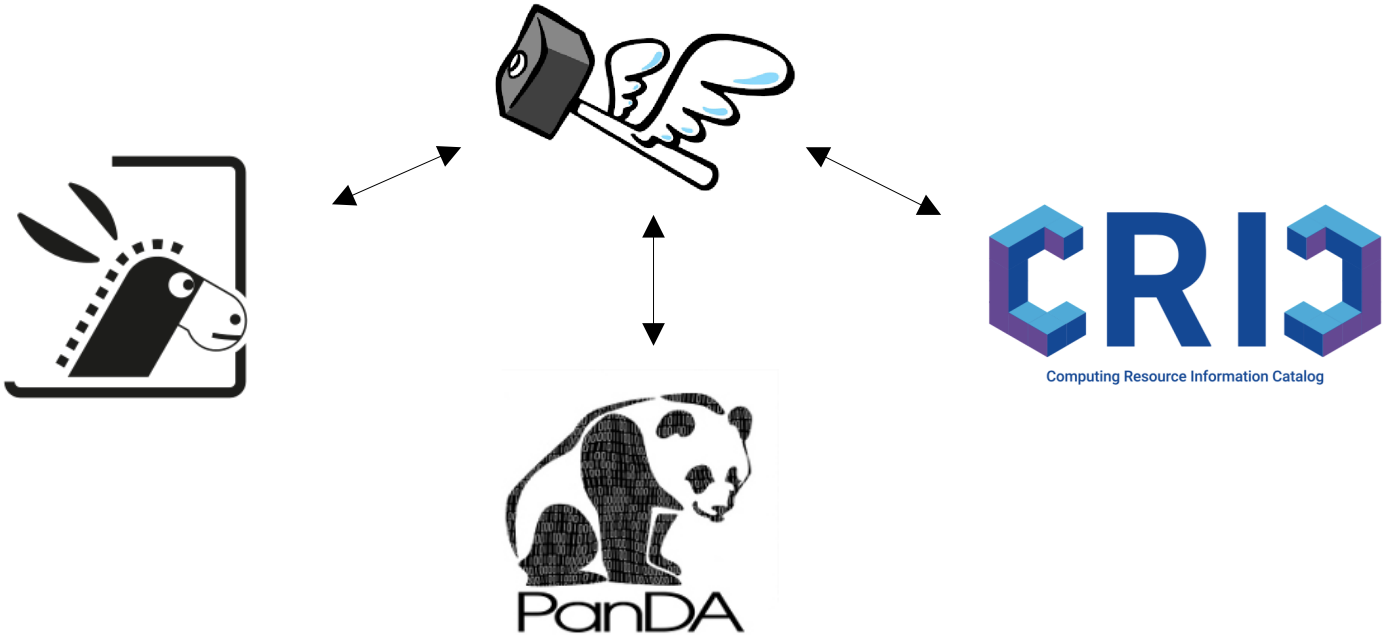
Enhancements and resource optimisations for ATLAS use of HammerCloud

Alexander Lory on behalf of the **ATLAS collaboration**

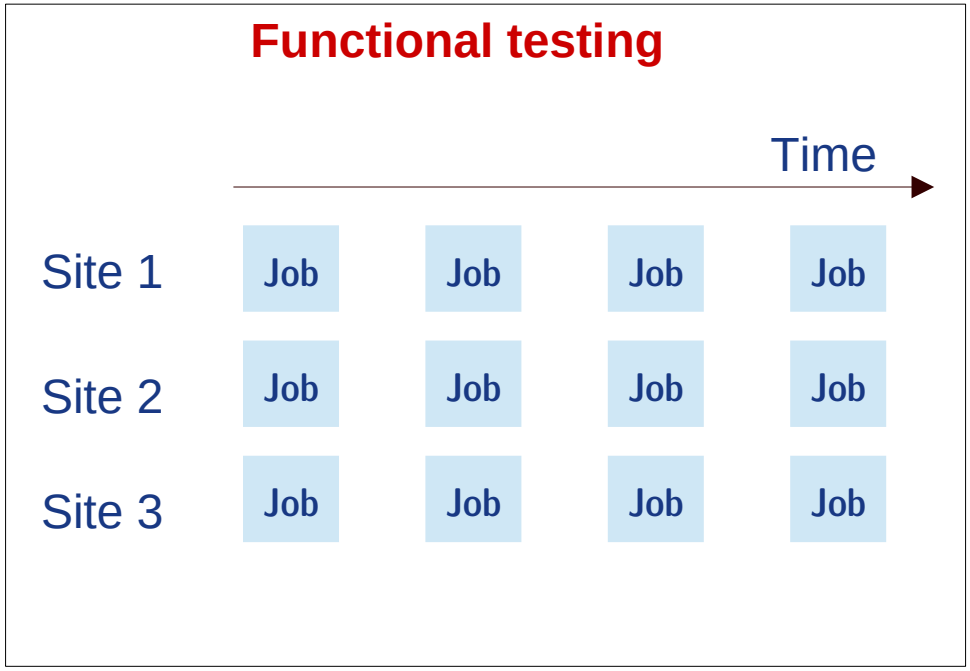
CHEP 2024 - Kraków

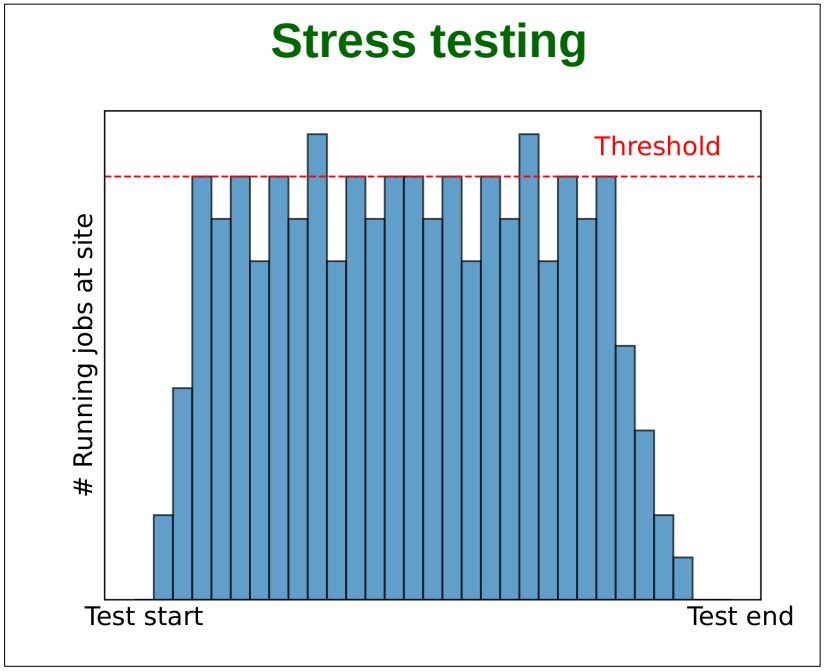


- LHC experiments utilize a large quantity of heterogeneous resources
  - ATLAS: **33 countries, 154 sites, 239 queues**
- Ensuring proper function and commissioning new sites requires **testing resources** with **full-chain workloads**
- **HammerCloud: A WLCG testing framework**



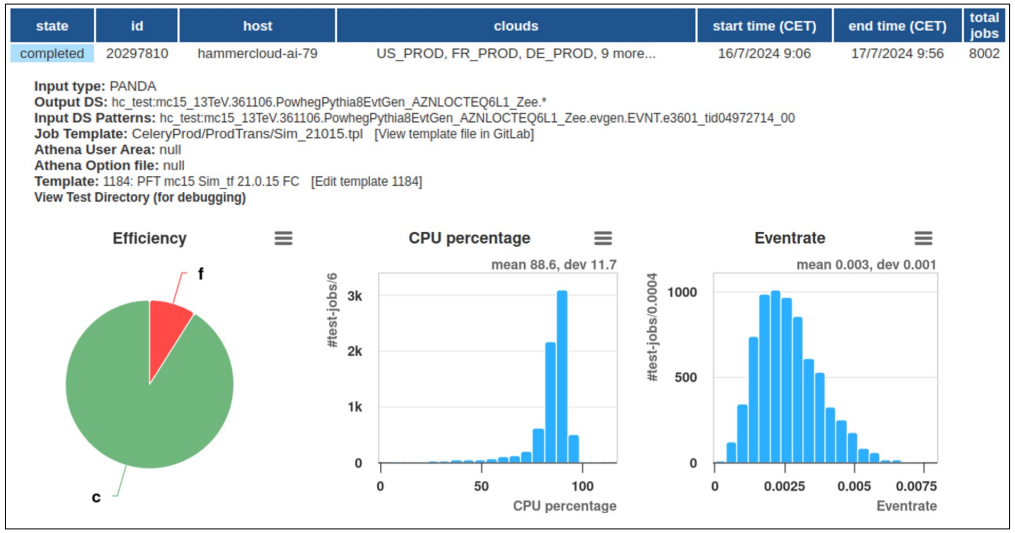
- **HammerCloud**:
  - **Automated** submission of **standardized** jobs
  - in **regular,** adjustable **intervals**
  - with adjustable number of **parallel** running jobs
- Two typical **modes**:
  - **Functional testing**: constant **stream** of **short jobs** on many (all) grid sites
  - **Stress testing**: Large amount of **parallel** jobs over given time-frame

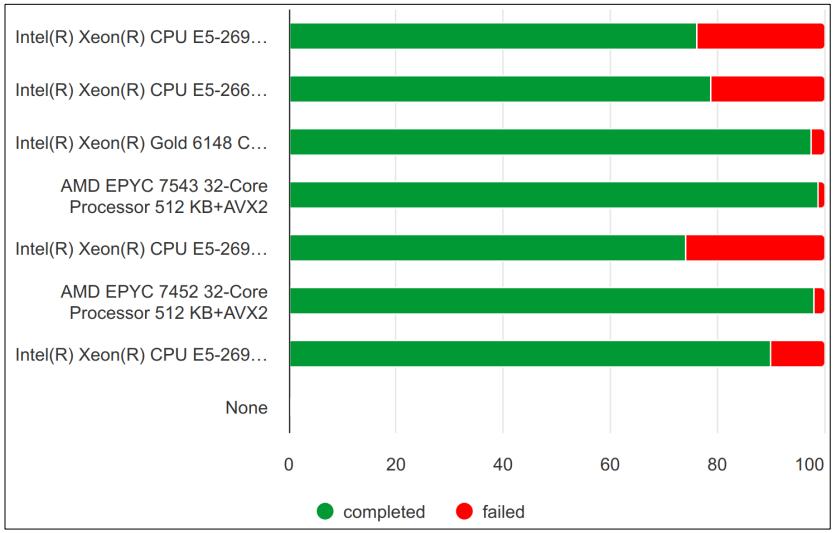


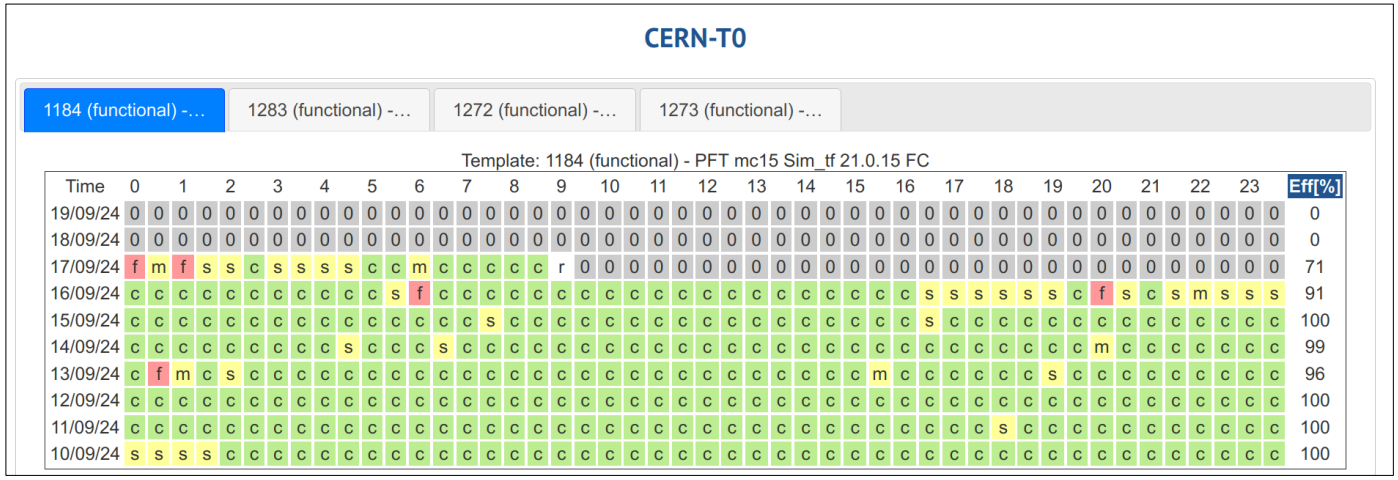


Also provides **monitoring tools** for HammerCloud and site admins

hammercloud.cern.ch

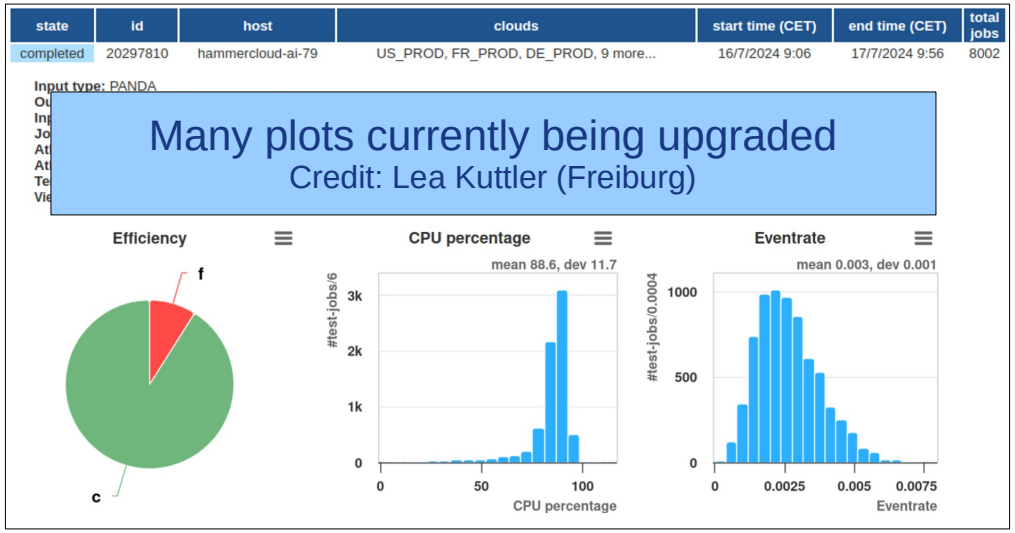


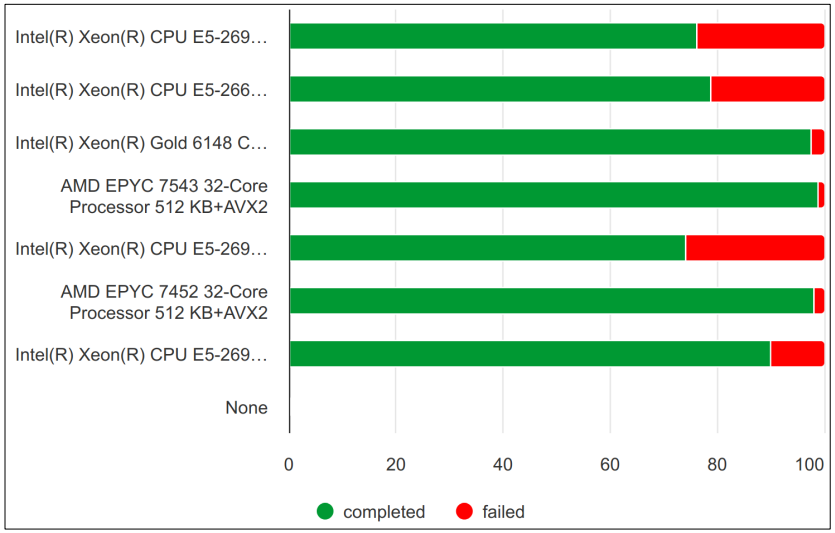


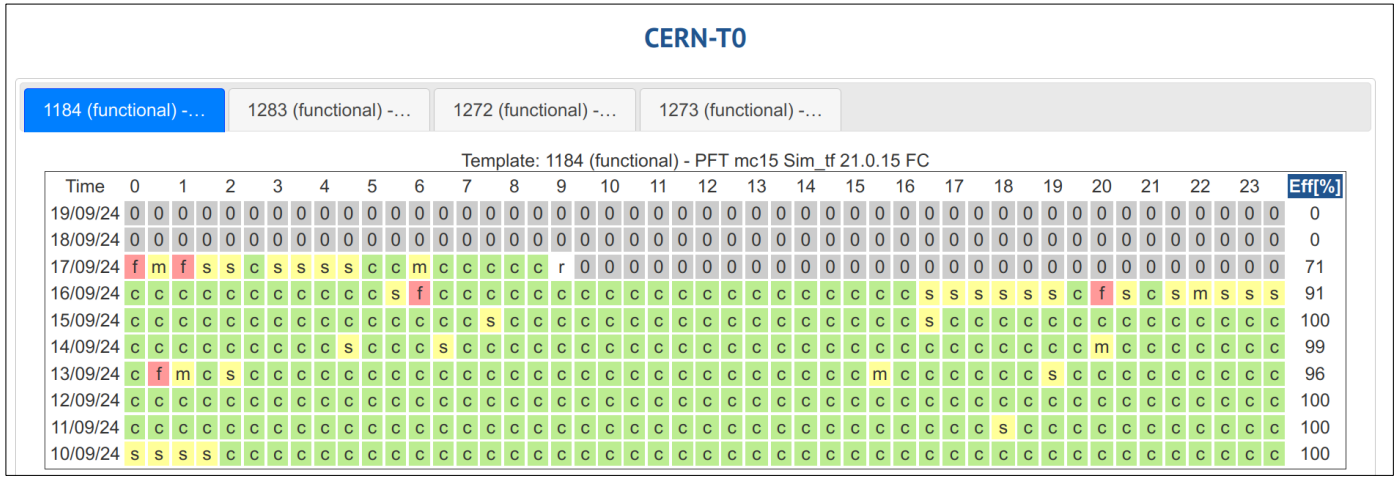

CERN-T0
1184 (functional) -…
1283 (functional) -…
1272 (functional) -…
1273 (functional) -…
Template: 1184 (functional) - PFT mc15 Sim_tf 21.0.15 FC
Time 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 Eff[%]
19/09/24 0 0 0

Also provides **monitoring tools** for HammerCloud and site admins

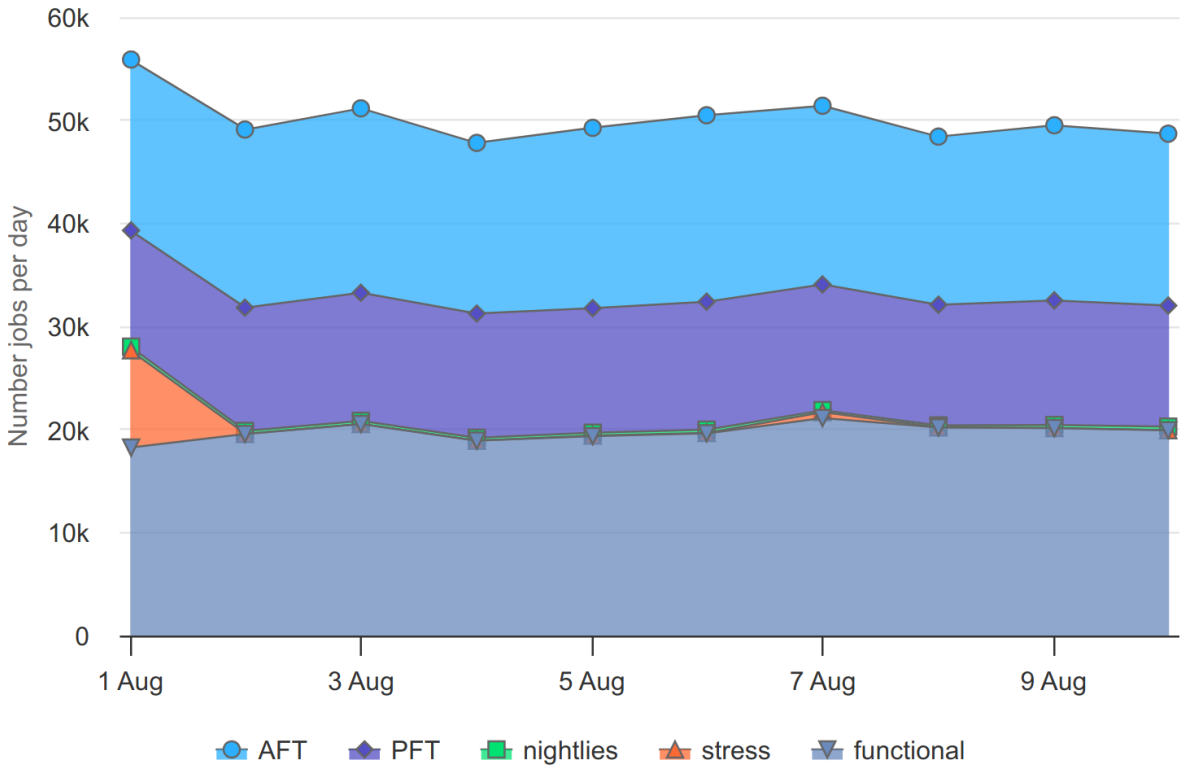
hammercloud.cern.ch



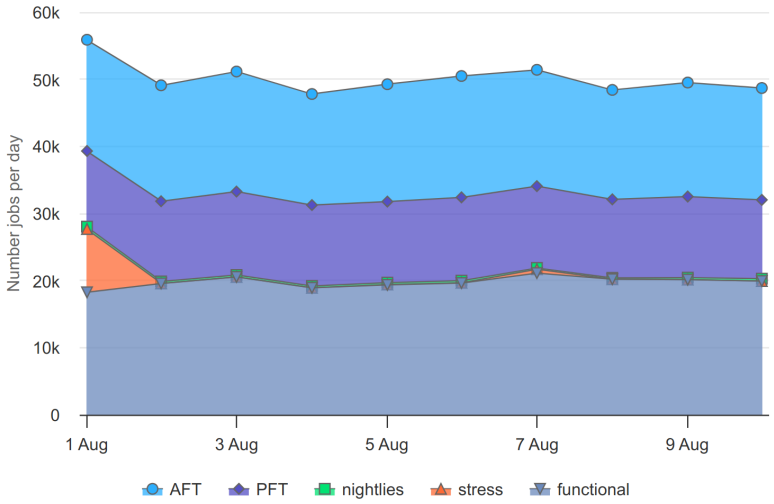



CERN-T0
1184 (functional) -…
1283 (functional) -…
1272 (functional) -…
1273 (functional) -…
Template: 1184 (functional) - PFT mc15 Sim_tf 21.0.15 FC
Time 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 Eff[%]
19/09/24 0 0 0

- At ATLAS:
  - 50k jobs daily, 750 job slots
- Categories:
  - Analysis Functional Tests (AFTs), Production Functional Tests (PFTs)
  - On demand stress tests
  - Other functional tests



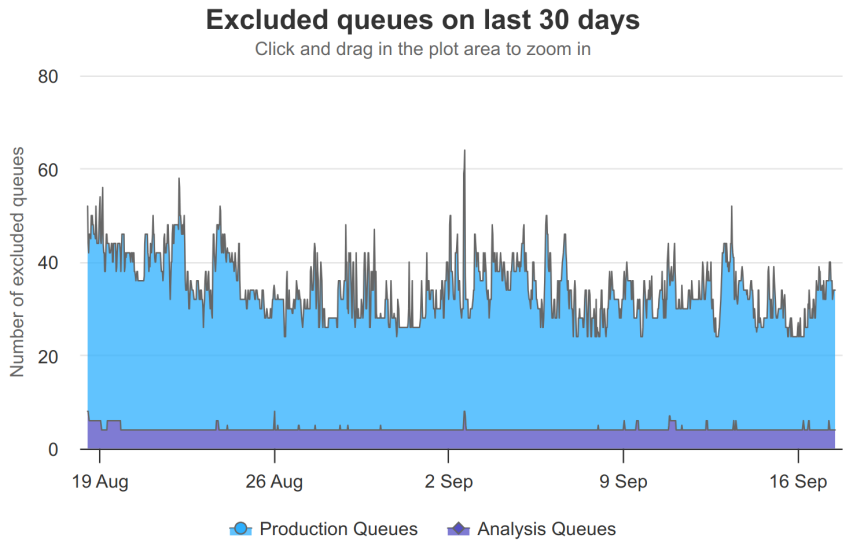
- The ***golden*** functional tests
  - High frequency, short duration, active 24/7
  - 7 tests, covering majority of grid workflows
    - AFTs: user analysis
    - PFTs: simulation
    - ARM: simulation
    - GPU: vector multiplication
  - Test results used for automatic exclusion / recovery of resources



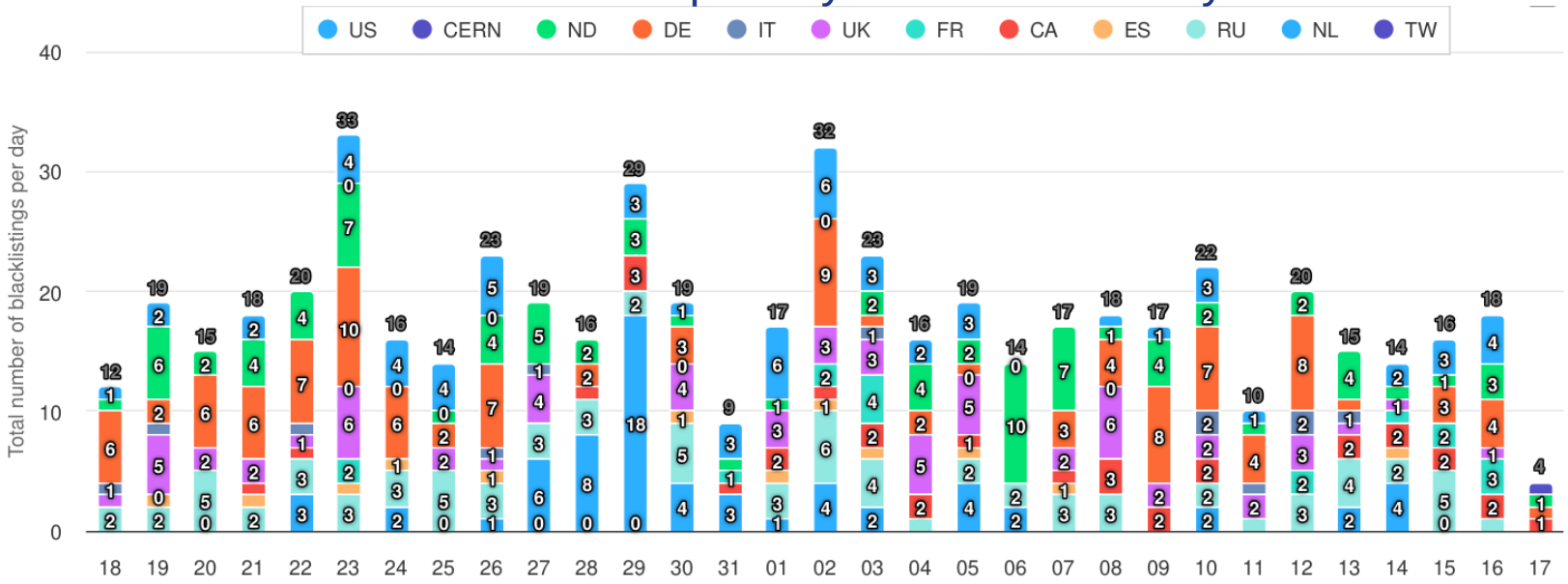
Running and Scheduled AFT/PFT Tests

State	Id	Host	Template	Start (Europe/ Zurich)	End (Europe/ Zurich)	Sites	subm jobs	run jobs	comp jobs	fail jobs	fail %	tot jobs
running	20301408	hammercloud-ai-74	1214: GPU Container GPU Available + Vector Multiplication (job based submission)	16/Sep, 13:44	17/Sep, 15:20	ANALY_BNL_GPU_ARC, ANALY_INFN-T1_GPU, ANALY_MANC_GPU, 6 more...	3	3	415	263	38	684
running	20301412	hammercloud-ai-73	1252: ARM mc21 Sim_tf 23.0.31 mcore	16/Sep, 15:04	17/Sep, 16:05	UKI-SCOTGRID-GLASGOW_ARM, INFN-CNAF_ARM, CERN-ARM, 2 more...	3	2	108	23	17	136
running	20301413	hammercloud-ai-72	1272: AFT EventLoop 22.2.113 centos7 directIO	16/Sep, 15:20	17/Sep, 16:10	ANALY_TOKYO, TOKYO, AGLT2, 115 more...	81	51	6033	409	6	6574
running	20301418	hammercloud-ai-75	1273: AFT EventLoop 25.2.7 el9 directIO	16/Sep, 17:22	17/Sep, 15:08	ANALY_TOKYO, TOKYO, AGLT2, 115 more...	86	48	5688	377	6	6199
running	20301421	hammercloud-ai-78	1283: PFT mc21 Sim_tf 22.0.73 mcore lomem clone	16/Sep, 18:54	17/Sep, 19:50	CERN, UNI-FREIBURG, AGLT2, 167 more...	89	78	1894	519	20	2580
running	20301432	hammercloud-ai-72	1184: PFT mc15 Sim_tf 21.0.15 FC	16/Sep, 23:36	18/Sep, 1:44	AGLT2_TEST, BEIJING, GoeGrid, 167 more...	103	86	3117	550	14	3856

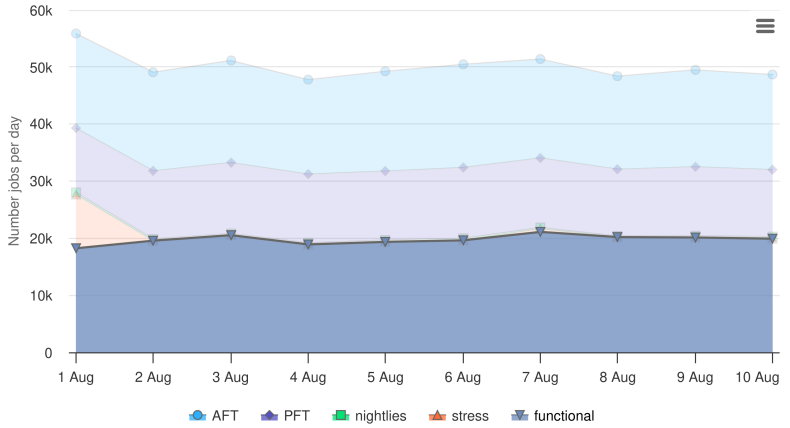
- Consider amount of **recent succesful/failed job**
- Set of **rules** triggers the **automatic exclusion / recovery** of sites (queues) from the pool of resources available to users
  - ~ 10 – 40 queues excluded at given time
  - ~ 10 – 30 daily exclusions / recoveries



Total exclusions per day for the last 30 days

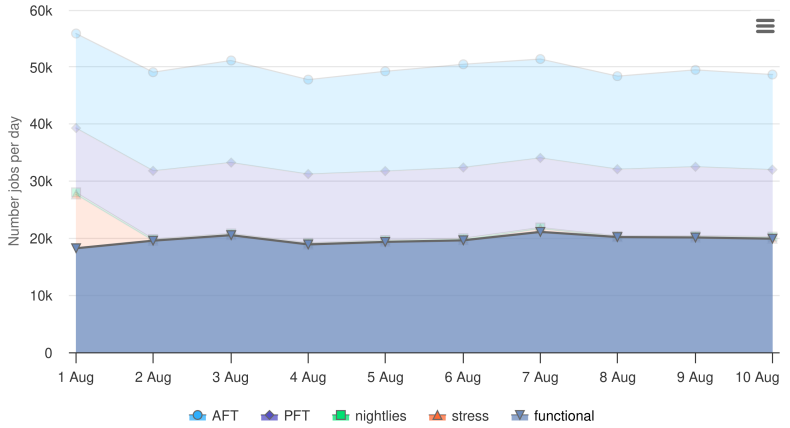


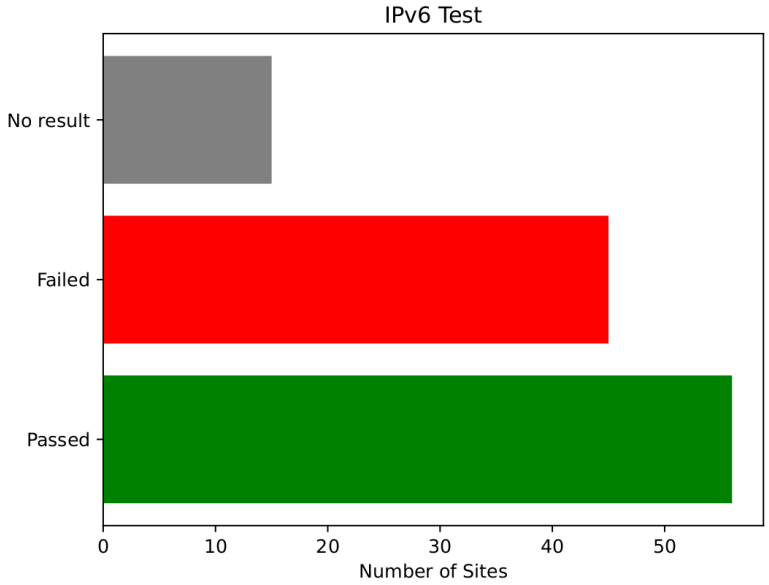
- Other functional tests:
  - Testing new software versions
    - Duplicate standard tests with different software



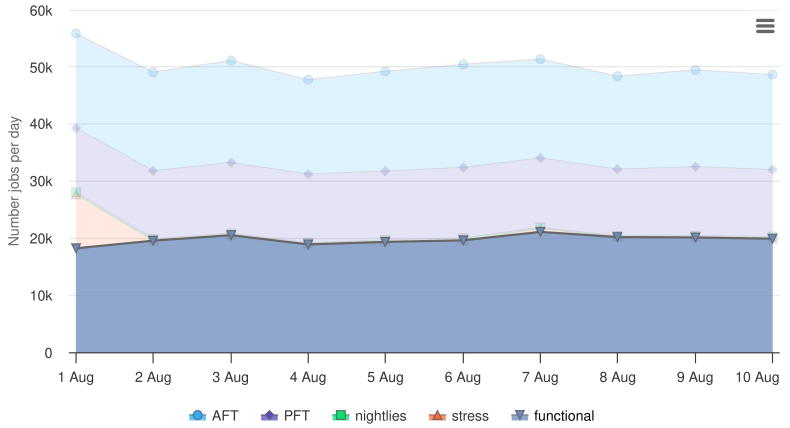
Jobs of template "1272 (functional) - AFT EventLoop 22.2.113 centos7 directIO"	Jobs of template "1274 (functional) - ALRB clone AFT EventLoop 22.2.113 centos7 directIO (each 2 hours)"
C: 89 | F: 2 | S: 1 | Total: 92 C: 96% | F: 2% |	vs. C: 11 | F: 1 | Total: 12 C: 91% | F: 8% |
c 6337344155 modificationHost: compute-5-39.local	c 6337303153 modificationHost: compute-21-27.local
c 6337353621 modificationHost: compute-11-32.local	c 6337312661 modificationHost: compute-21-27.local
c 6337355940 modificationHost: compute-2-26.local	c 6337470514 modificationHost: compute-23-28.local
c 6337357714 modificationHost: compute-12-33.local	c 6337650683 modificationHost: compute-24-10.local
c 6337360631 modificationHost: compute-18-37.local	c 6337839746 modificationHost: compute-6-3.local
c 6337366792 modificationHost: compute-12-33.local	c 6338010899 modificationHost: compute-10-14.local
c 6337382840 modificationHost: compute-12-35.local	c 6338145206 modificationHost: compute-6-32.local
c 6337393119 modificationHost: compute-21-18.local	c 6338266316 modificationHost: compute-6-27.local
c 6337411391 modificationHost: compute-2-29.local	c 6338381804 modificationHost: compute-12-16.local
c 6337433196 modificationHost: compute-24-15.local	c 6338510578 modificationHost: compute-19-15.local
c 6337441121 modificationHost: compute-11-32.local	c 6338594880 modificationHost: compute-26-26.local
c 6337449192 modificationHost: compute-13-28.local	f 6338669976 modificationHost: compute-12-33.local pilot:::1378 Info: /cvmfs mounted; do 'setupATLAS -d -c ...' to skip default mounts. Info: $HOME mounted; do 'setupATLAS -d -c ...' to skip default mounts. ---------- Apptainer: 1.2.2 Host: Linux, CentOS Linux 7 (Core), x86_64, 3.10.0-1160.88.1.el7.x86_64 From: /cvmfs/atlas.cern.ch/repo/containers/sw/apptainer/x86_64-el7/1.2.2/bin/apptainer ContainerType: atlas-default apptainer exec -e -B /condor-ce/gk01/htcondor-spool/1073/0/cluster296

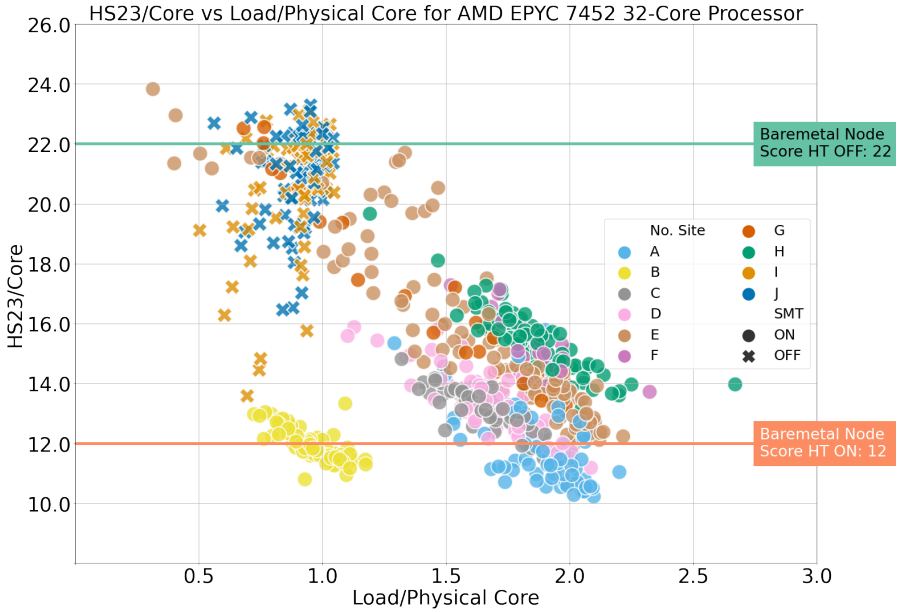
- Other functional tests:
  - Testing new software versions
    - Duplicate standard tests with different software
  - Monitoring IPv6 deployment on CEs
    - One IPv6 test job / day





- Other functional tests:
  - Testing new software versions
    - Duplicate standard tests with different software
  - Monitoring IPv6 deployment on CEs
  - Benchmarking sites
    - Measurement of HEPscore every 3h on all sites
    - Used e.g. to measure and increase efficiency of computing in production environment





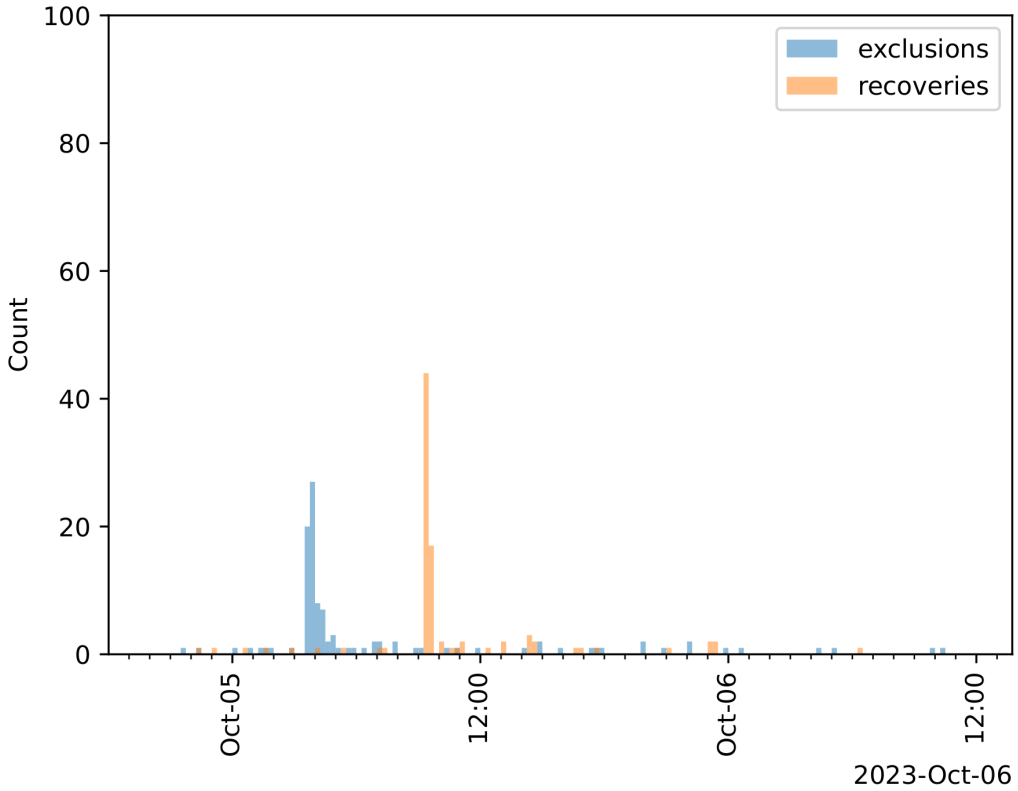
Analysis by Natalia Diana Szczepanek in the next talk (slides)

Massive recovery

- Sometimes a central issue causes a **massive exclusion** of resources
- This central issue does not reflect the state of the sites
- Recovery of sites sometimes not as fast as desired, due to lacking test results

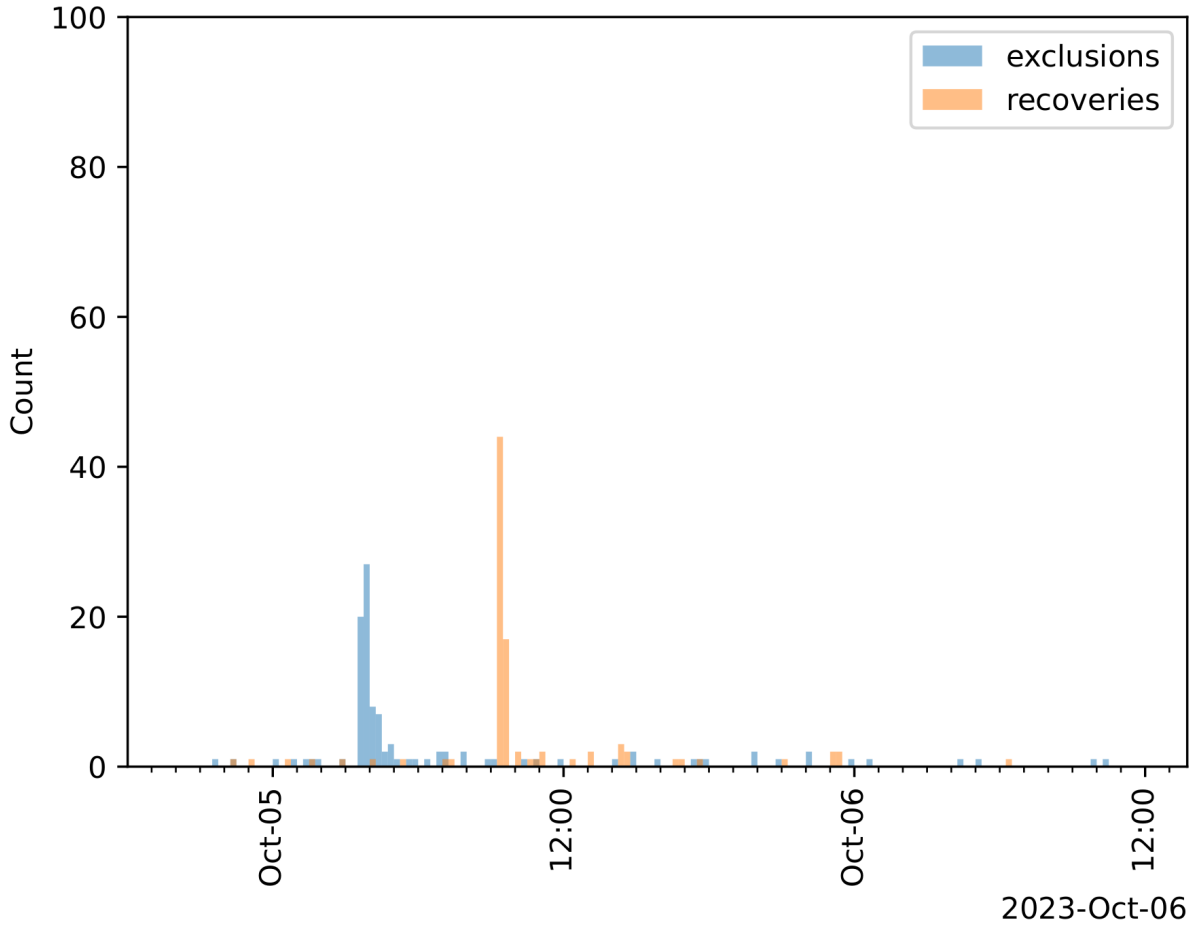
  → A feature introduced this year **speeds up the recovery** of lagging sites

Massive automatic recovery

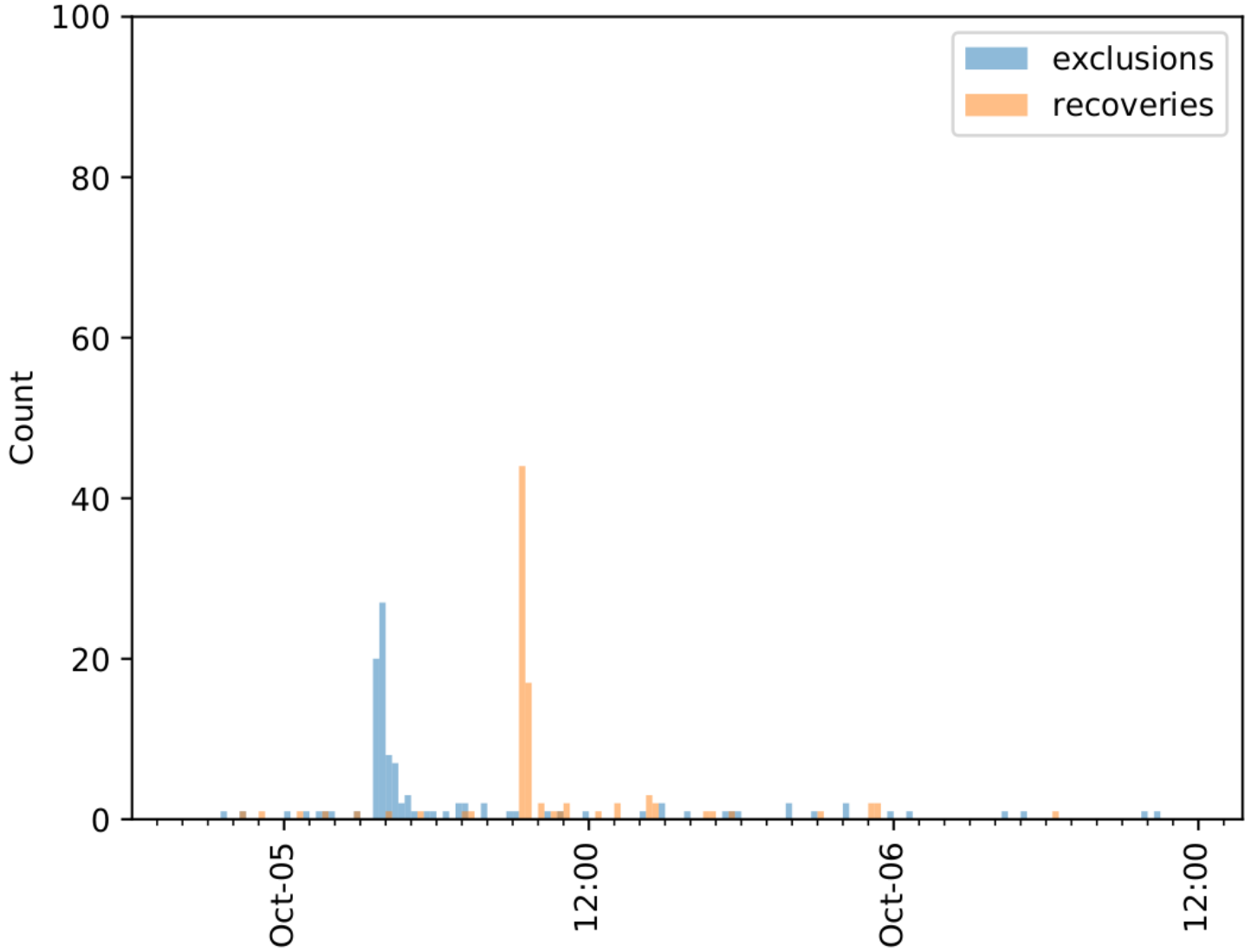


Massive recovery

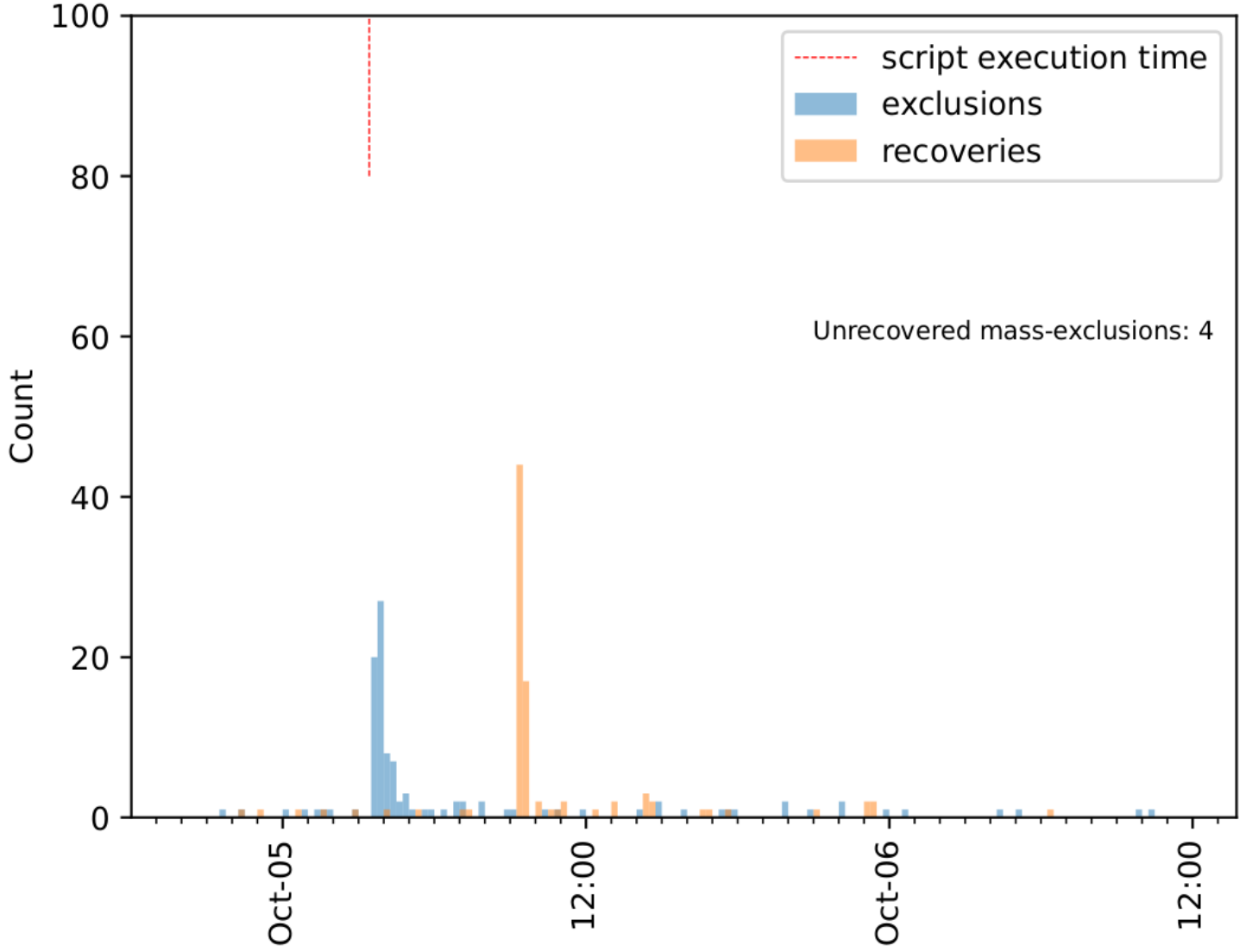
- Basic idea:
  - Detect massive exclusion event
  - When first sites start to recover, recover all remaining sites



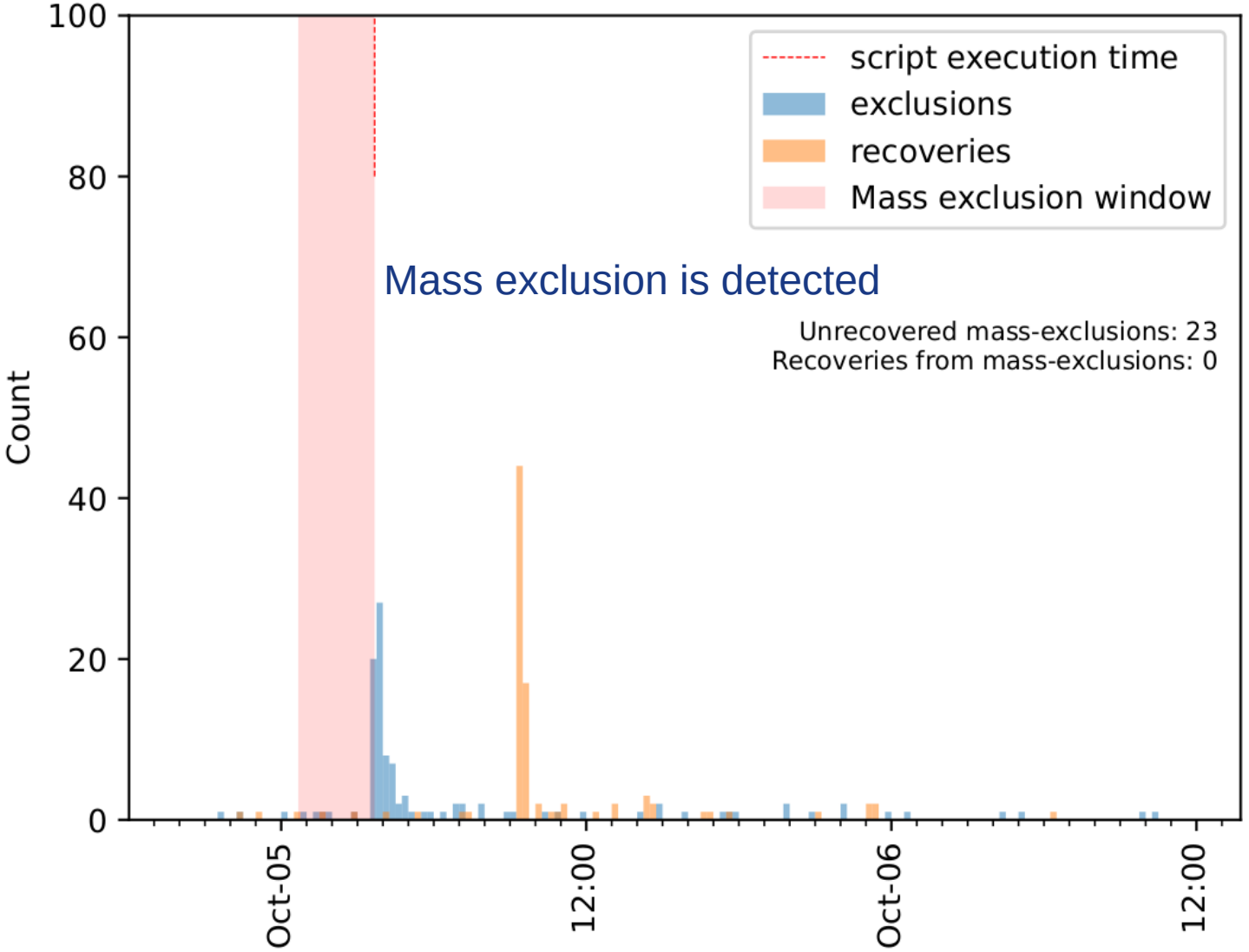
Massive recovery



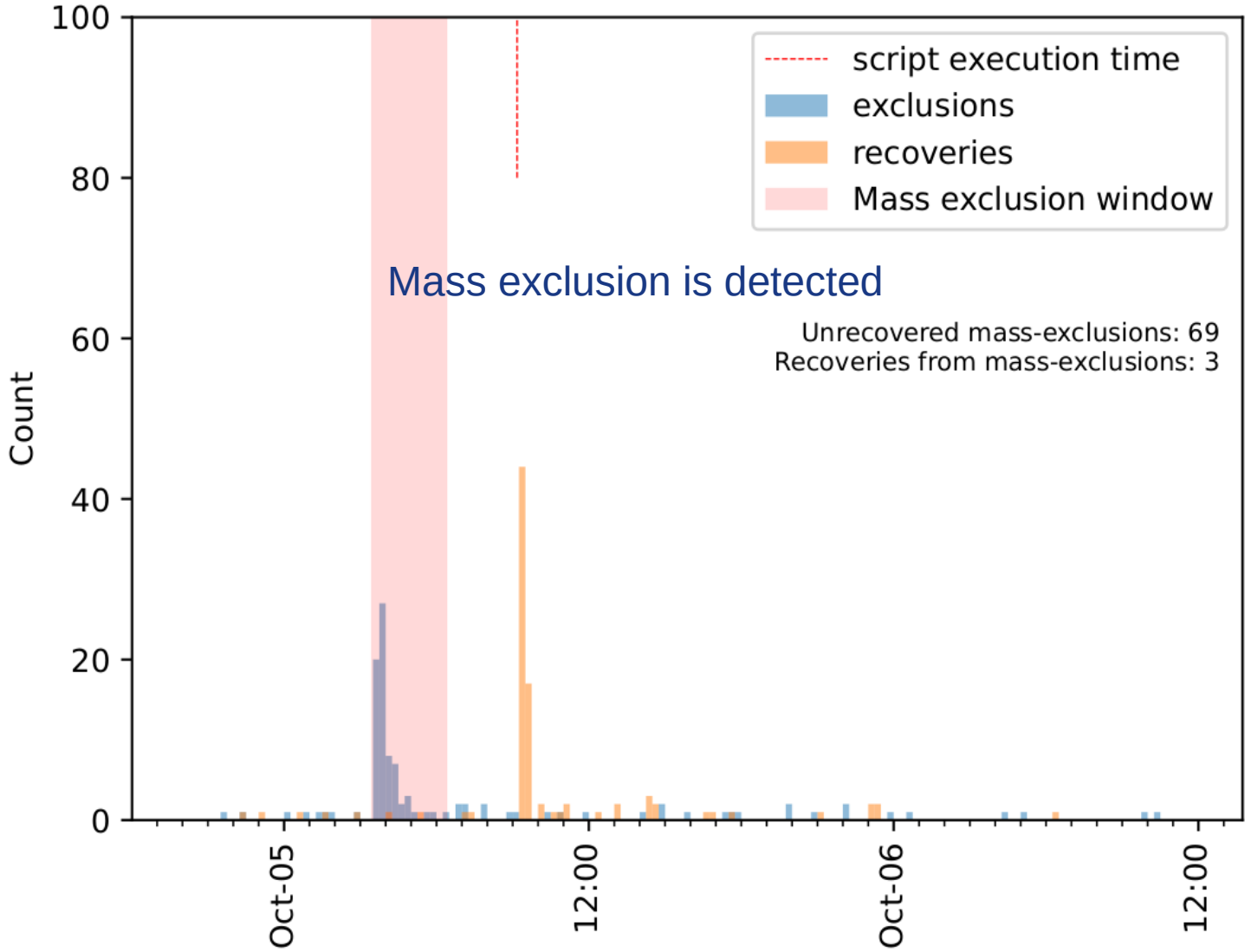
Massive recovery



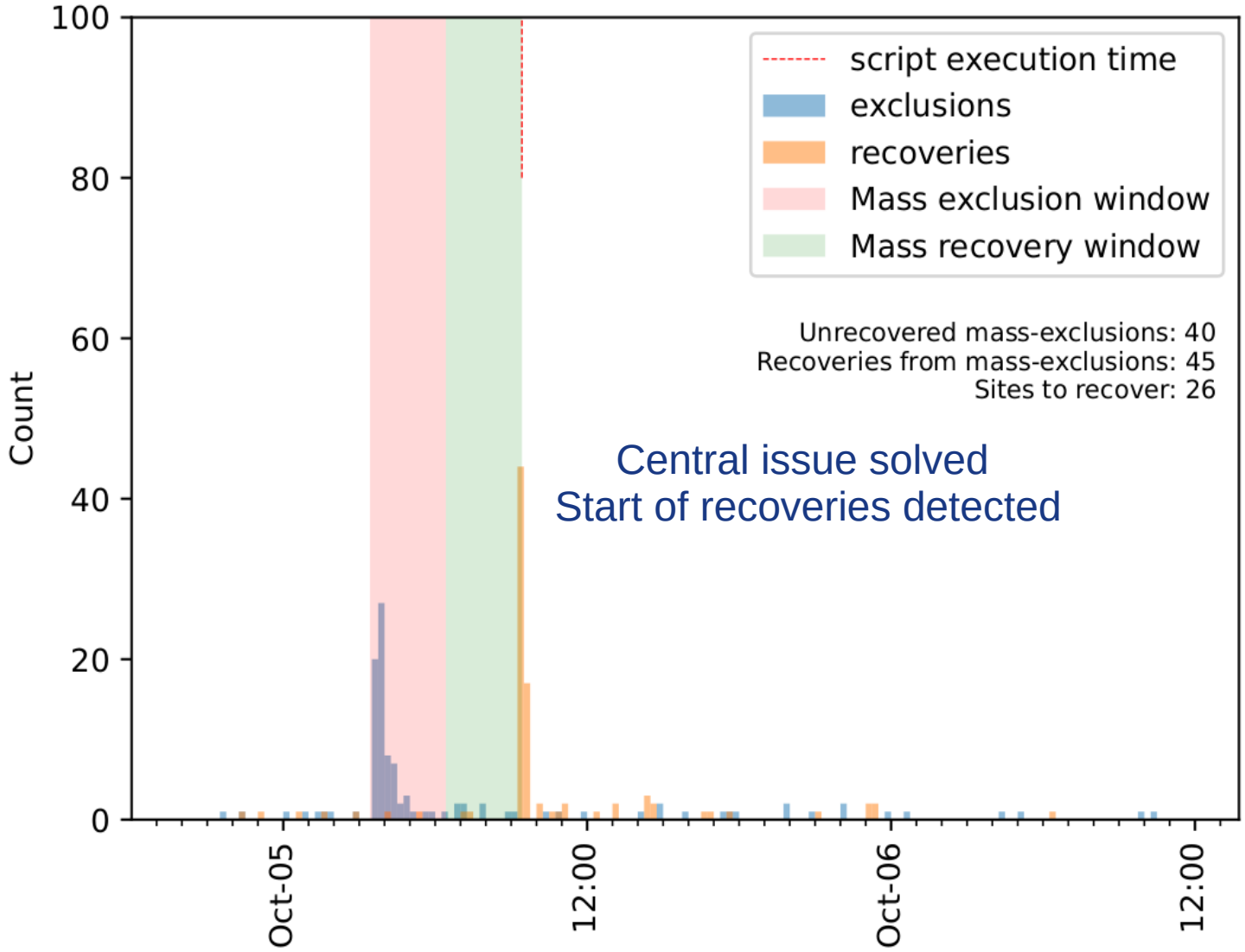
Massive recovery



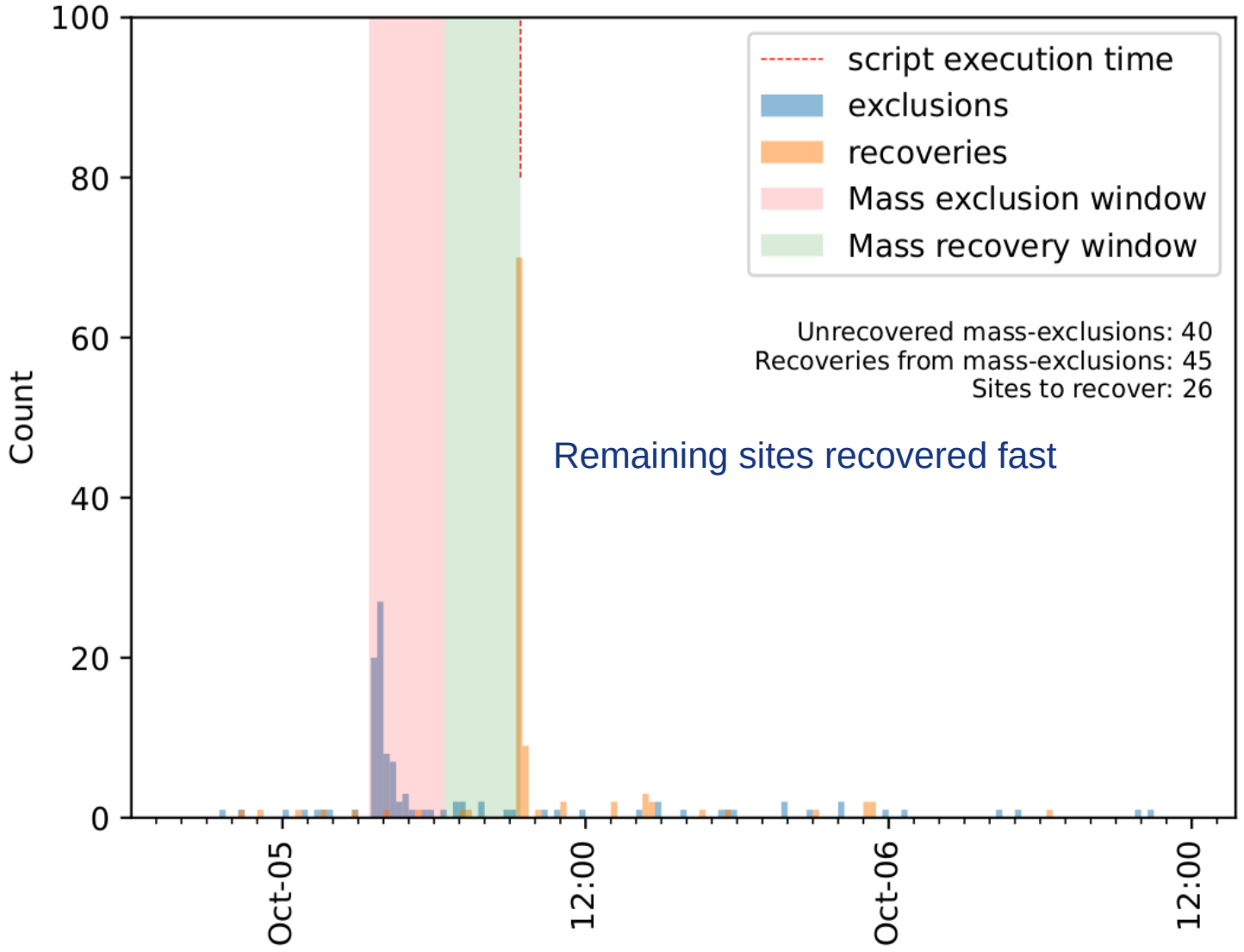
Massive recovery



Massive recovery



Massive recovery

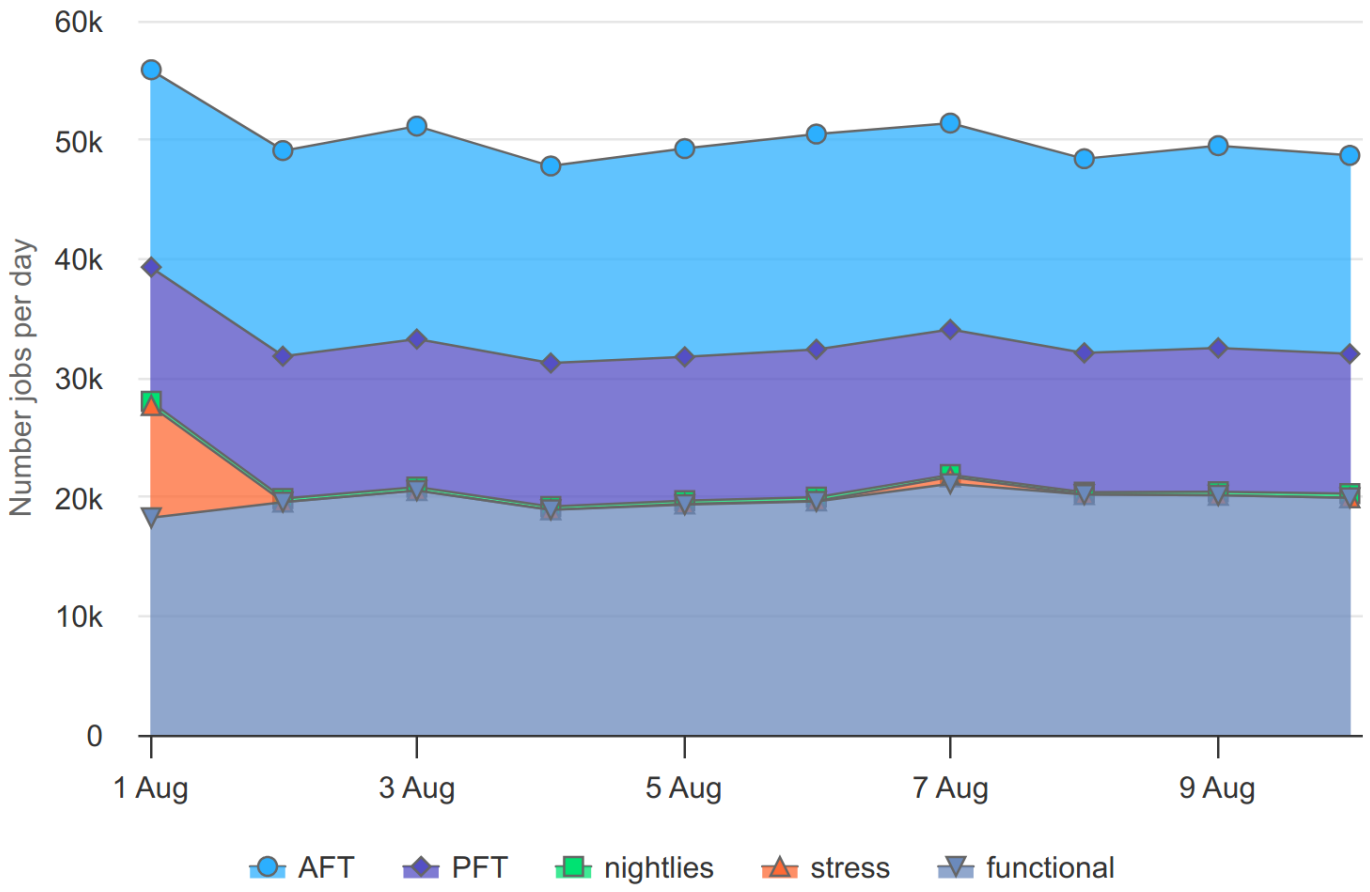


Massive recovery

The mass-recovery feature kicked in two times since its introduction (Feb. 2024)

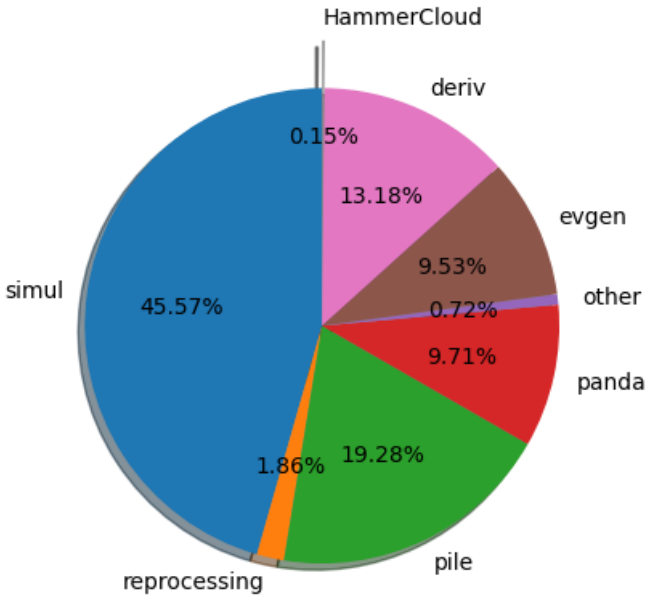
Speeding up recovery of 19 and 55 queues respectively

Resources utilisation



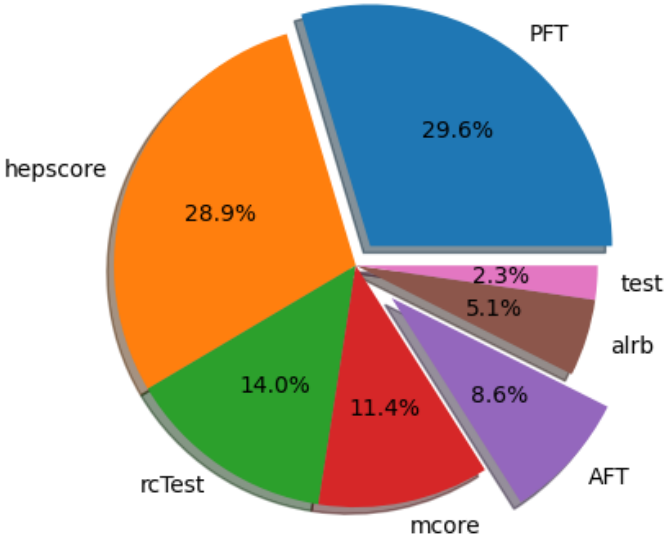
50k jobs/ day, waste of resources?

Share of resources used by HammerCloud



0.15% to the hs23 hours of the grid in 2024

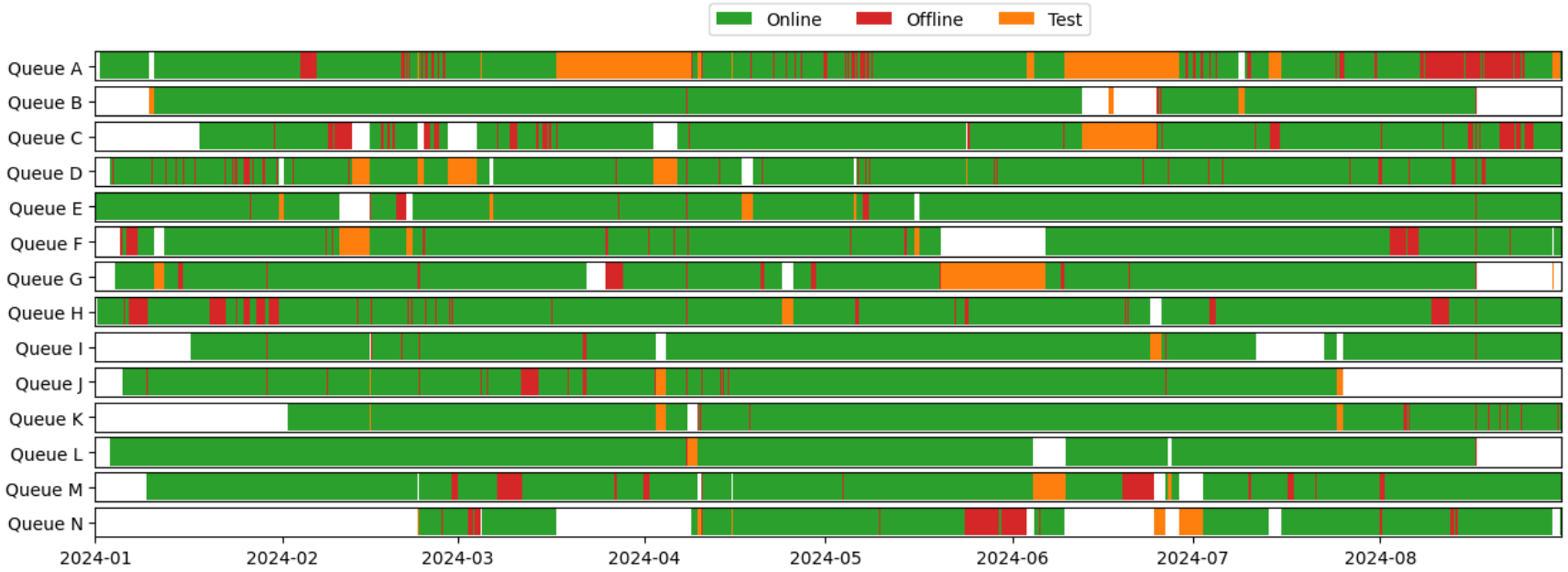
Shares of HammerCloud test types



38% for auto-exclusion

→ **0.058%** of total ATLAS hs23 hours 2024

- What is the impact of queue exclusion?
  - Typical availability profile:



  - ~ 4.6% of the total runtime in 2024 queues were auto-excluded
  - Saving effect depends on site
    - If nodes get **idle** when excluded: **O(50%) energy saved**
    - If shared/full site → **CPU used for other VO → 100% energy saved**

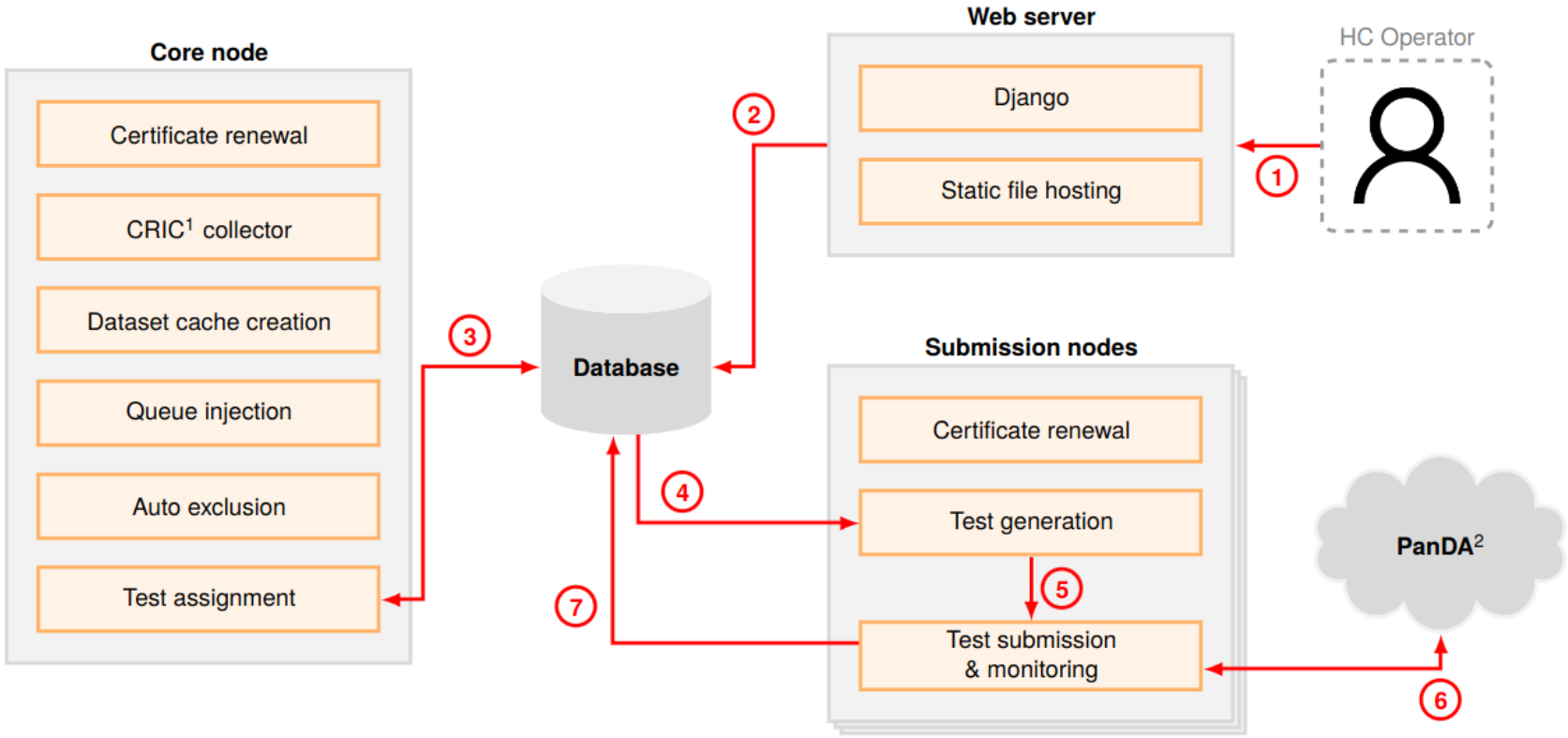
    → Excluding resources prevents **2 – 4%** of wasted energy
      - Compared to this, **0.058%** of energy used for auto-exclusion is small

Through **testing**, **benchmarking**, and **auto-excluding** sites

HammerCloud ensures a
good grid user experience and
efficient resource utilisation

Thank you!

HammerCloud design



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

[1] **C**omputing **R**esource **I**nformation **C**atalog

[2] **P**roduction **and** **D**istributed **A**nalysis – workload management system