

A new HTCondor monitoring for CNAF Tier-1

HTCondor Workshop Autumn 2021



The project

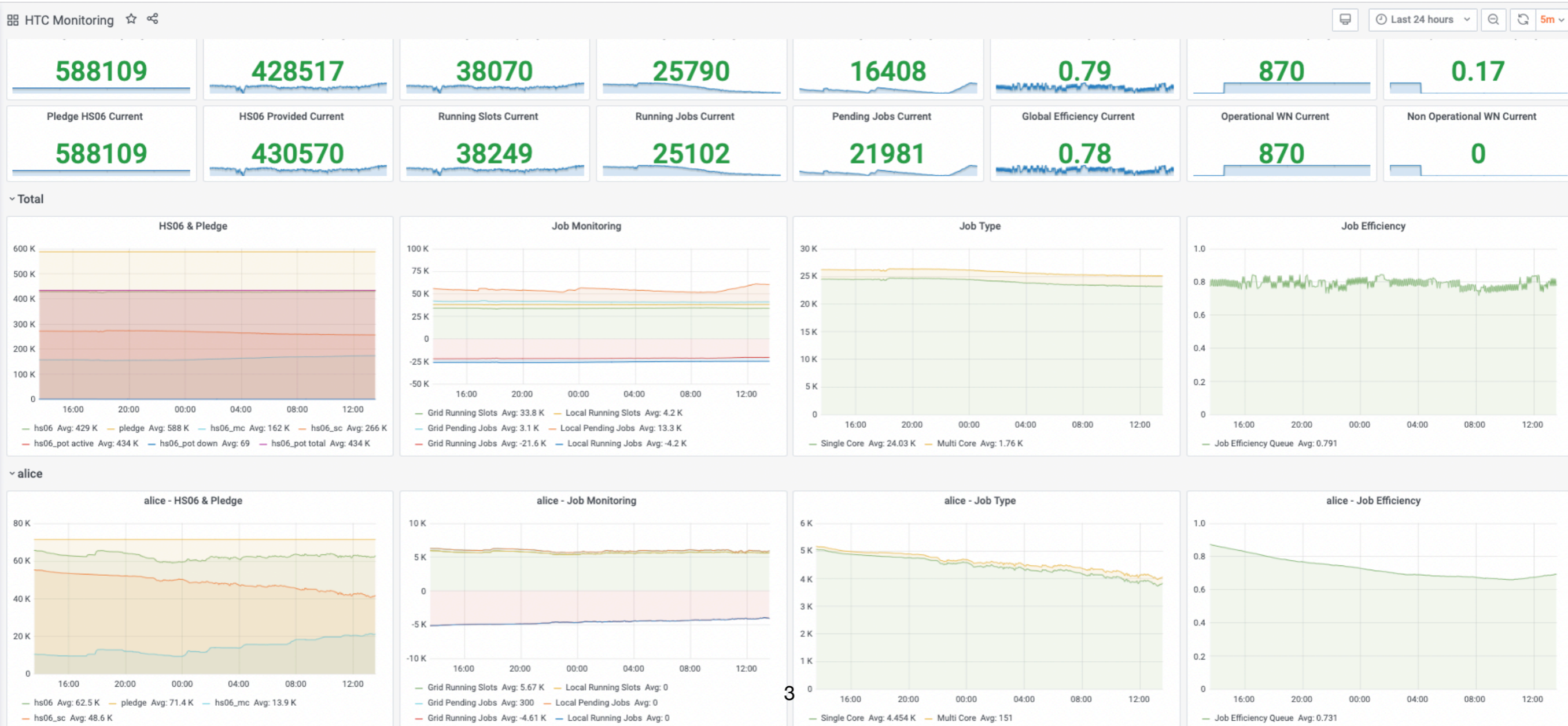
outline

- **htc tier-1 farm** —> see previous talk by Stefano Dal Pra
- **Introduction**
- **Data storage:**
 - PostgreSQL database
- **Data collection:**
 - condor_q
 - condor_status
- **Data presentation:**
 - monitoring dashboard
- **What's next**

Introduction


Where we start and why another monitoring system

- Monitoring:  *influxdb*
 Grafana
- Only aggregate views \rightarrow need for different granularities



Data storage

PostgreSQL

- PostgreSQL database 
 - **Job data tables:**
 - hj —> last 24h jobs
 - hj_old —> backup
 - timestamp index
 - **Machine data table:**
 - hm —> all data
 - timestamp index

JOBs Data collection

condor_q -> running jobs



```
cmd_runjobs = ""condor_q -global -all -cons 'Member(JobStatus,{2,3,4}) && (JobUniverse != 7) && (RegExp("wn-2|cn-6|ba-3",  
(RemoteHost ? : LastRemoteHost))) && ((time() - JobStartDate) > 3)' -af:j \  
AcctGroup \  
'IfThenElse(ScheddHostName == "sn-01","local","grid")' \  
'split(splitSlotName(RemoteHost ? : LastRemoteHost)[1],".")[0]' \  
'time() - JobStartDate <= 180' \  
'int(Member(JobStatus,{3,4}) && (time() - CompletionDate) <= 3 * 60)' \  
JobStatus \  
'CpusProvisioned ? : min({RequestCpus ? : 1,8})' \  
'((int(MATCH_t1_wn_hs06) + 0.0)/MATCH_TotalSlotCpus)' \  
'(int(MATCH_t1_wn_hs06) + 0.0)' \  
'(CumulativeRemoteSysCpu ? : 0) + (CumulativeRemoteUserCpu ? : 0)' \  
'(time() - JobStartDate)' \  
Owner \  
'ResidentSetSize_RAW/1e6 ? : 0.0' \  
'ImageSize_RAW/1e6 ? : 0.0' \  
jobstartdate \  
'ScheddHostName ? : "unknown"' \  
'(DiskUsage_RAW ? : 0)/1e6'""
```

Custom ClassAds

MACHINEs Data collection

condor_status



```
cmd_condor_status = ""condor_status -comp -cons '(RegExp("wn-2|cn-6|ba-3",Machine))' \
-af \
'split(Machine,"@.")[0]' \
Cpus \
Disk \
t1_GPFS_CHECK \
t1_CPU1_T \
t1_CPU2_T \
t1_Syst_T \
t1_Peripheral_T \
t1_MemAvailable \
t1_MemTotal \
t1_SwapFree \
t1_SwapTotal \
TotalCondorLoadAvg \
TotalLoadAvg \
MonitorSelfAge
```

HTCondor cron: /usr/bin/ipmitool sdr list full

Custom ClassAds

HTCondor cron: /proc/meminfo

Dashboard

Plotly Dash plotly

- **Requirements:**
 - Fully customizable dashboard and plots
 - **Dash-Plotly** (<https://plotly.com>) libraries + **Pandas**
 - Tools for “job selection”
 - Plots of the full farm overview as a function of time
- **Implementation:**
 - Cloud machine (2 VCpus + 2GB RAM) with service exposed in an internal CNAF net
 - Load share between the database and the dashboard machine

—> **see live demo**



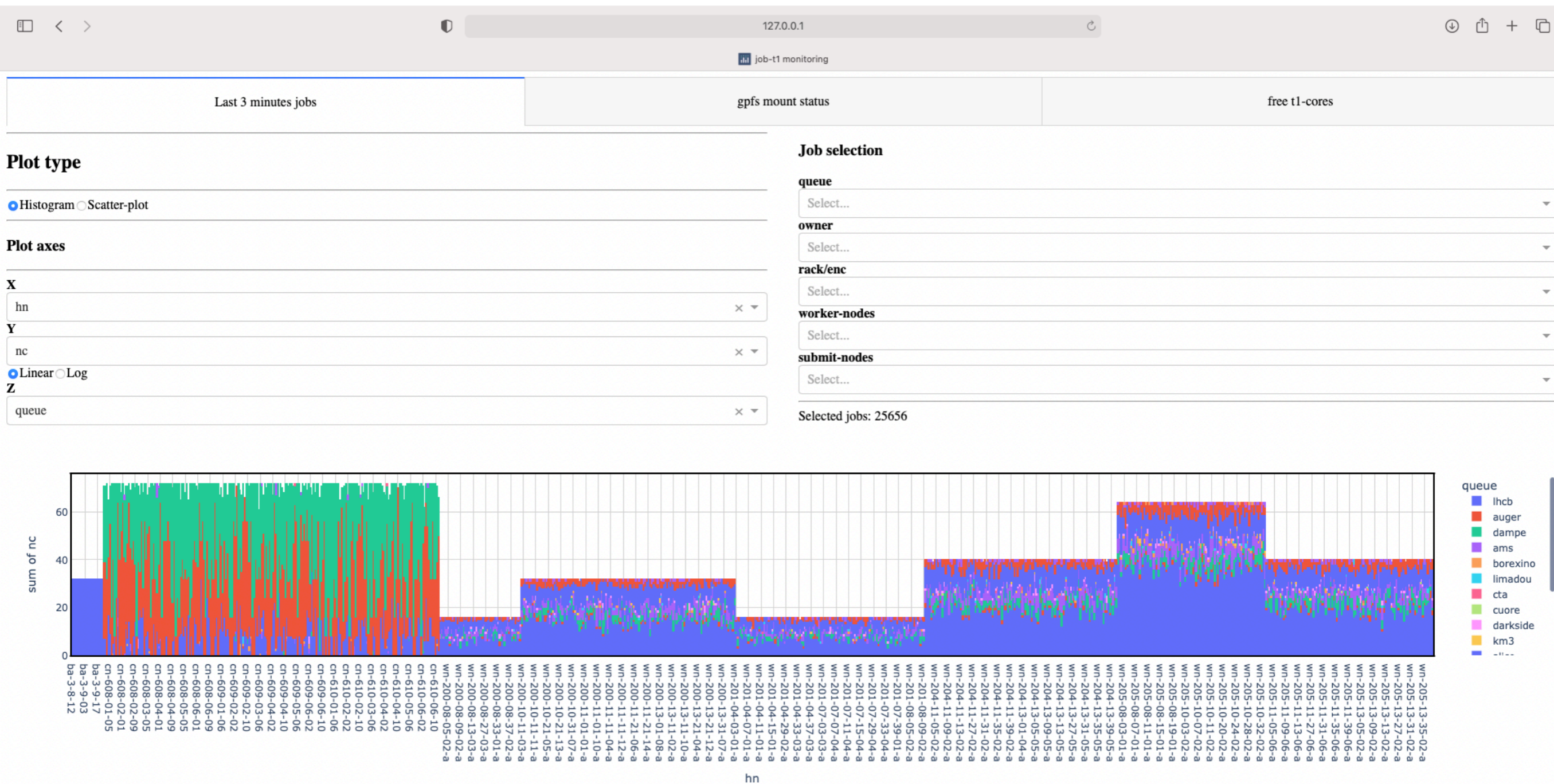
What's next

- **Data collection:**
 - **Add** more **ClassAds** of interest
 - Add **pending jobs** information
 - **Data aggregation** for long-time studies
- **Dashboard:**
 - Selection of *most useful/informative* plots
 - Add *time-dependent* job plots
 - Cosmetic improvements
- **Off-line data analyses:**
 - JOB-clustering
 - Predicting jobs behavior



Backup

Dashboard layout: 1



Dashboard layout: 2

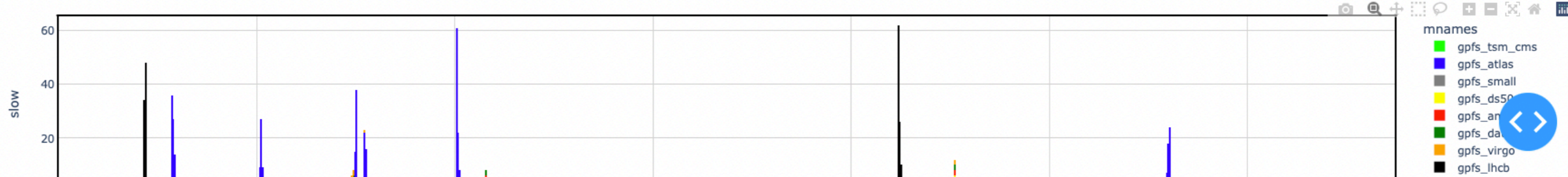
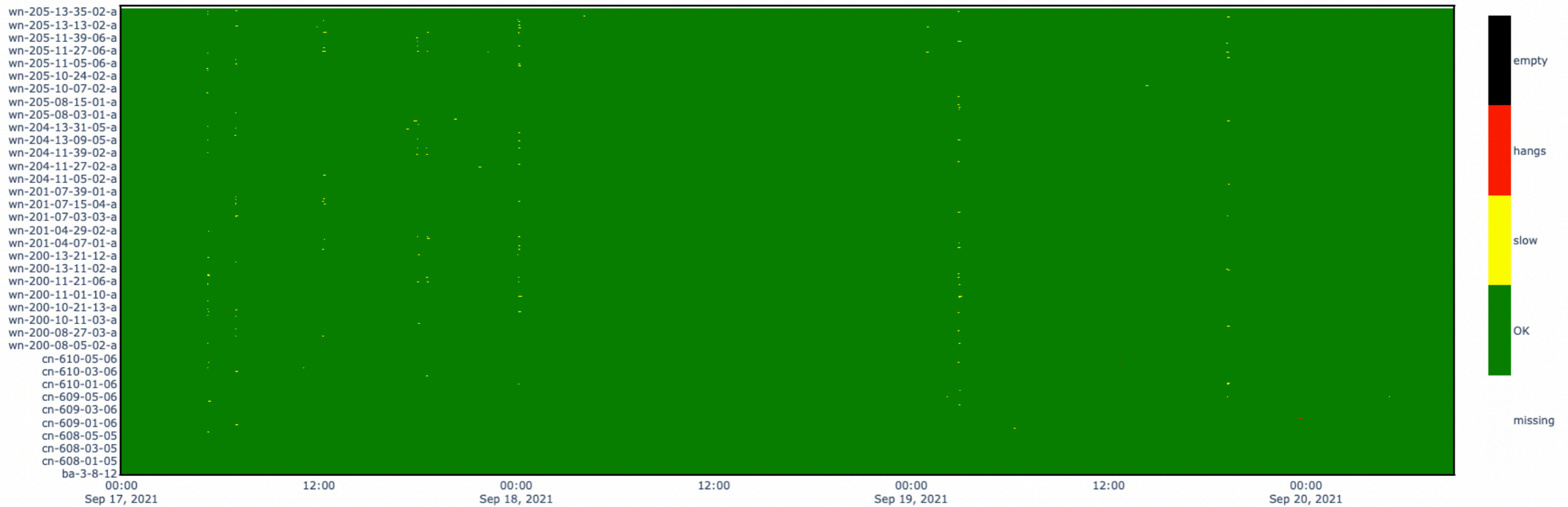
127.0.0.1 job-t1 monitoring

Last 3 minutes jobs gpfs mount status free t1-cores

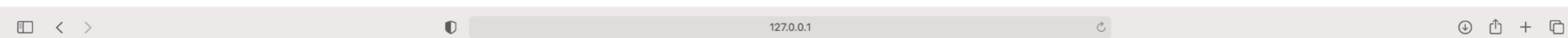
Start Date → End Date Update plots

Select data time interval.

- Default refresh: last 12 hours
- This could take a while...around 1 minute for each selected day



Dashboard layout: 3



job-t1 monitoring

Last 3 minutes jobs

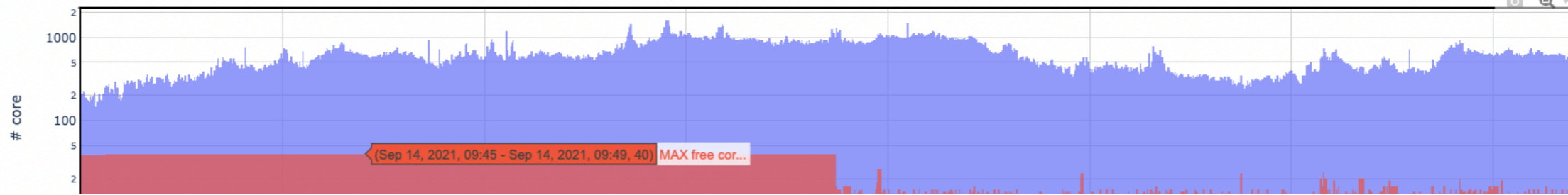
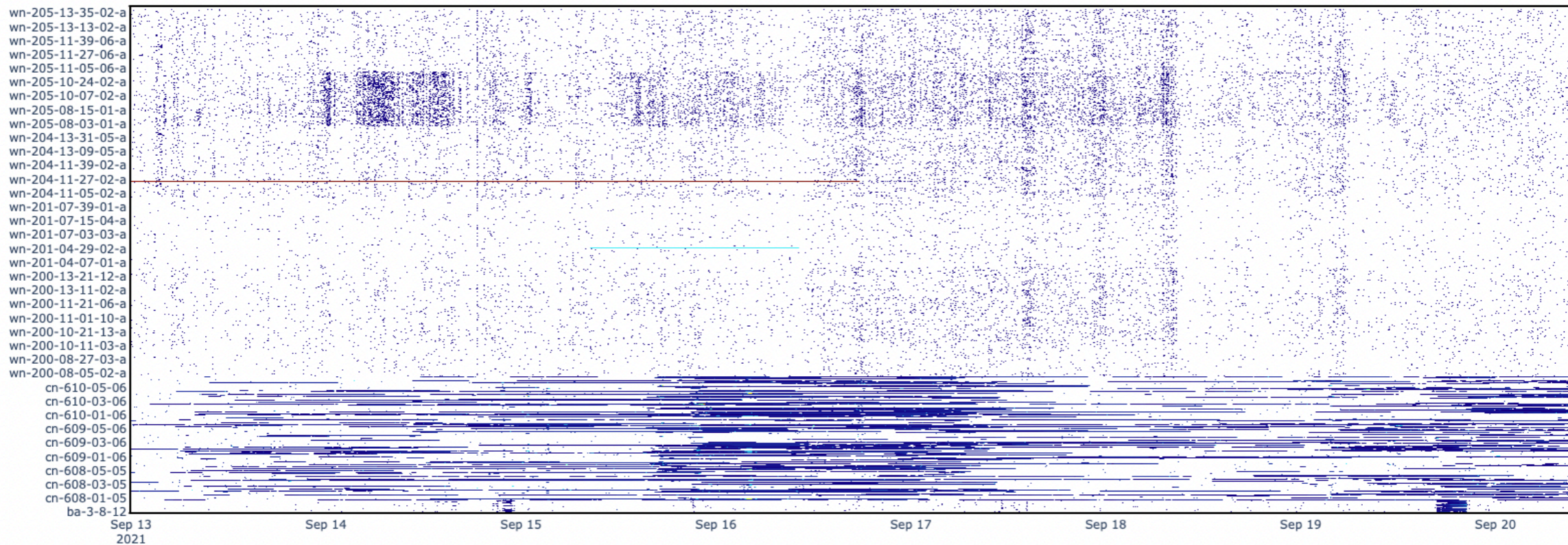
gpfs mount status

free t1-cores

Select data time interval.

Start Date → End Date Update plots

- Default refresh: last 12 hours
- This could take a while...around 1 minute for each selected day



Dashboard layout: 1

job scatter-plot

