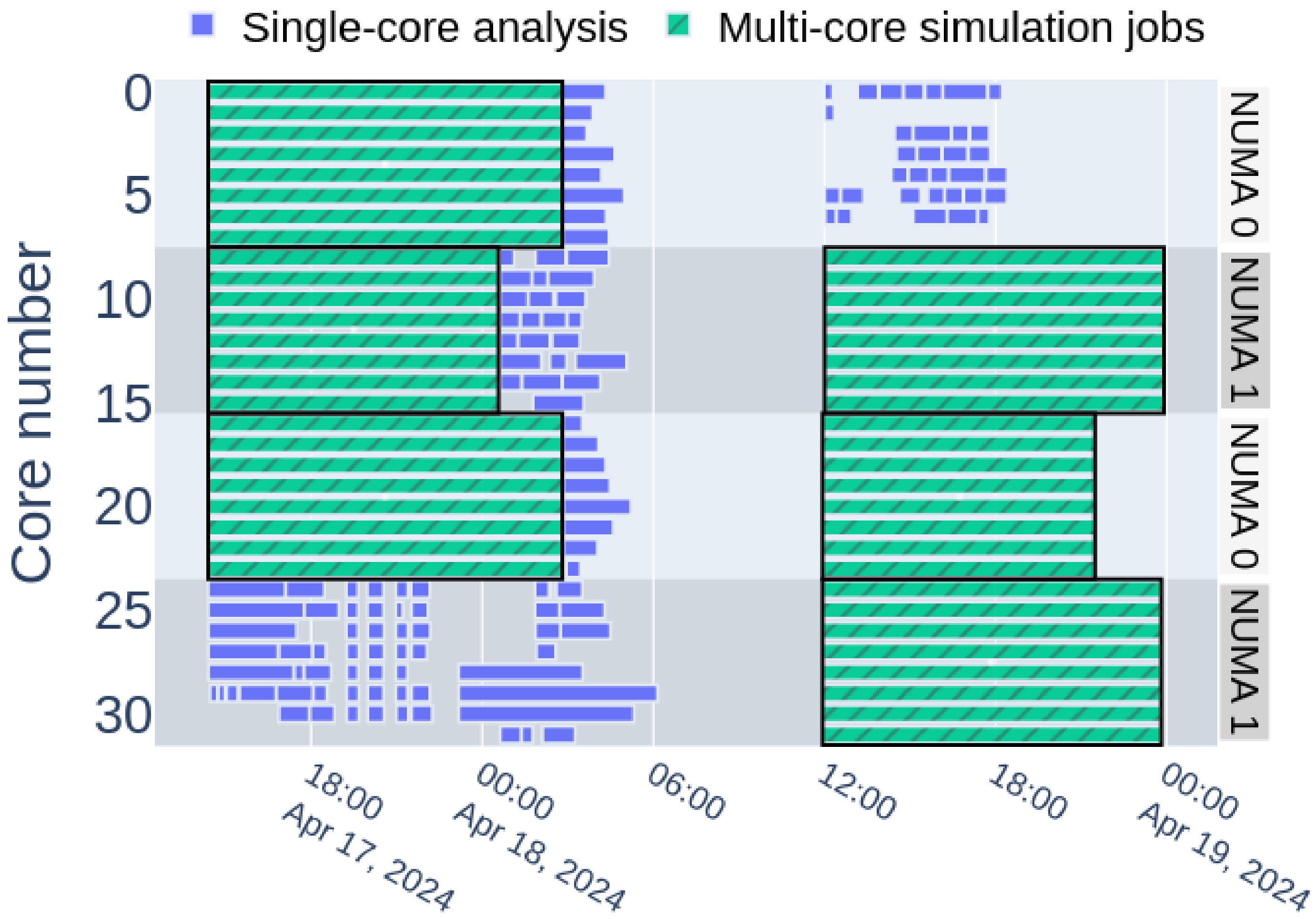
Job TTL Optimization

Elena Mihăilescu

ALICE Day @ UPB, 2024

FACULTY OF
AUTOMATIC CONTROL
AND COMPUTERS

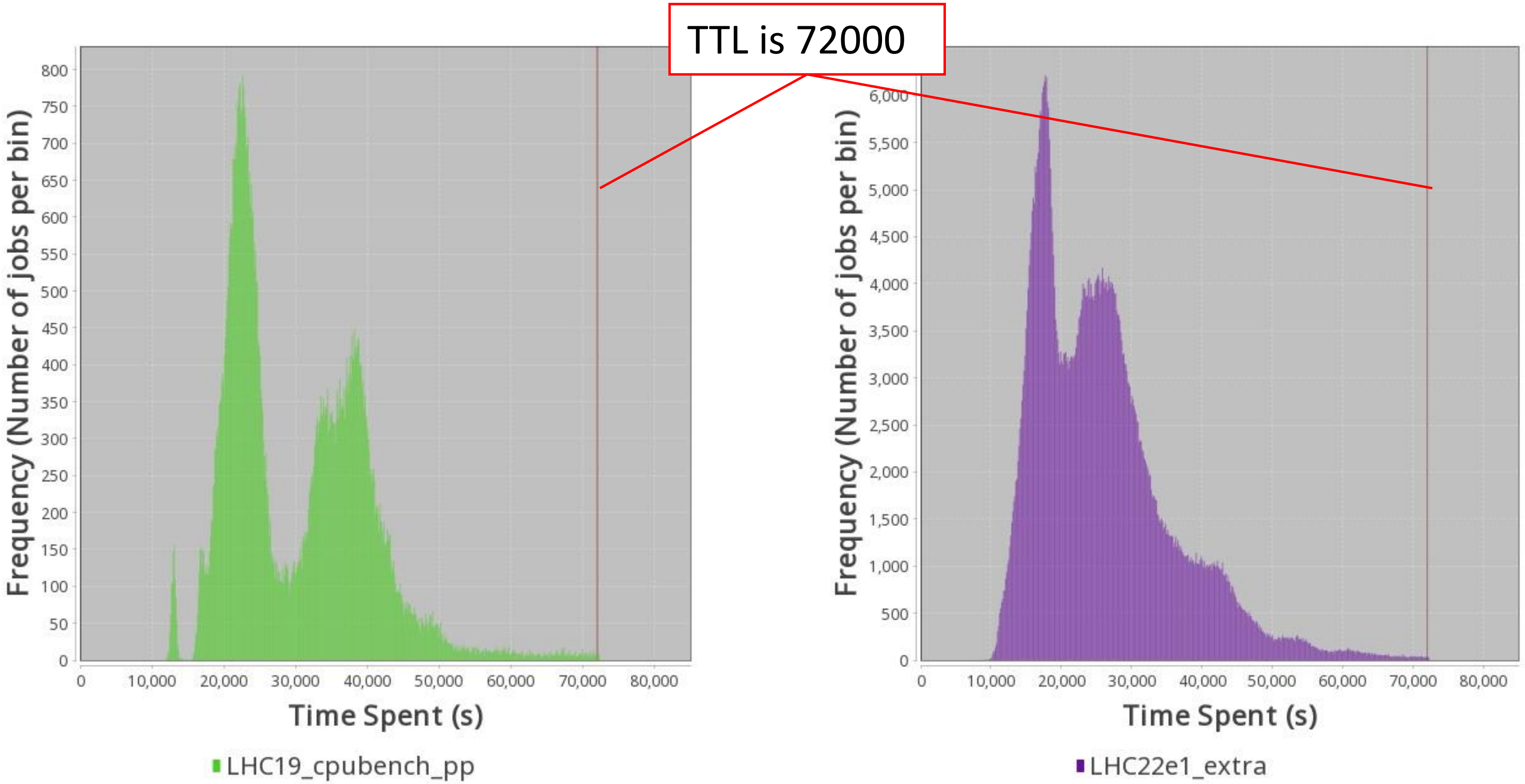
Job distribution in a Grid slot



Marta Bertran Ferrer, *Whole-node scheduling in the ALICE Grid*

What we have seen

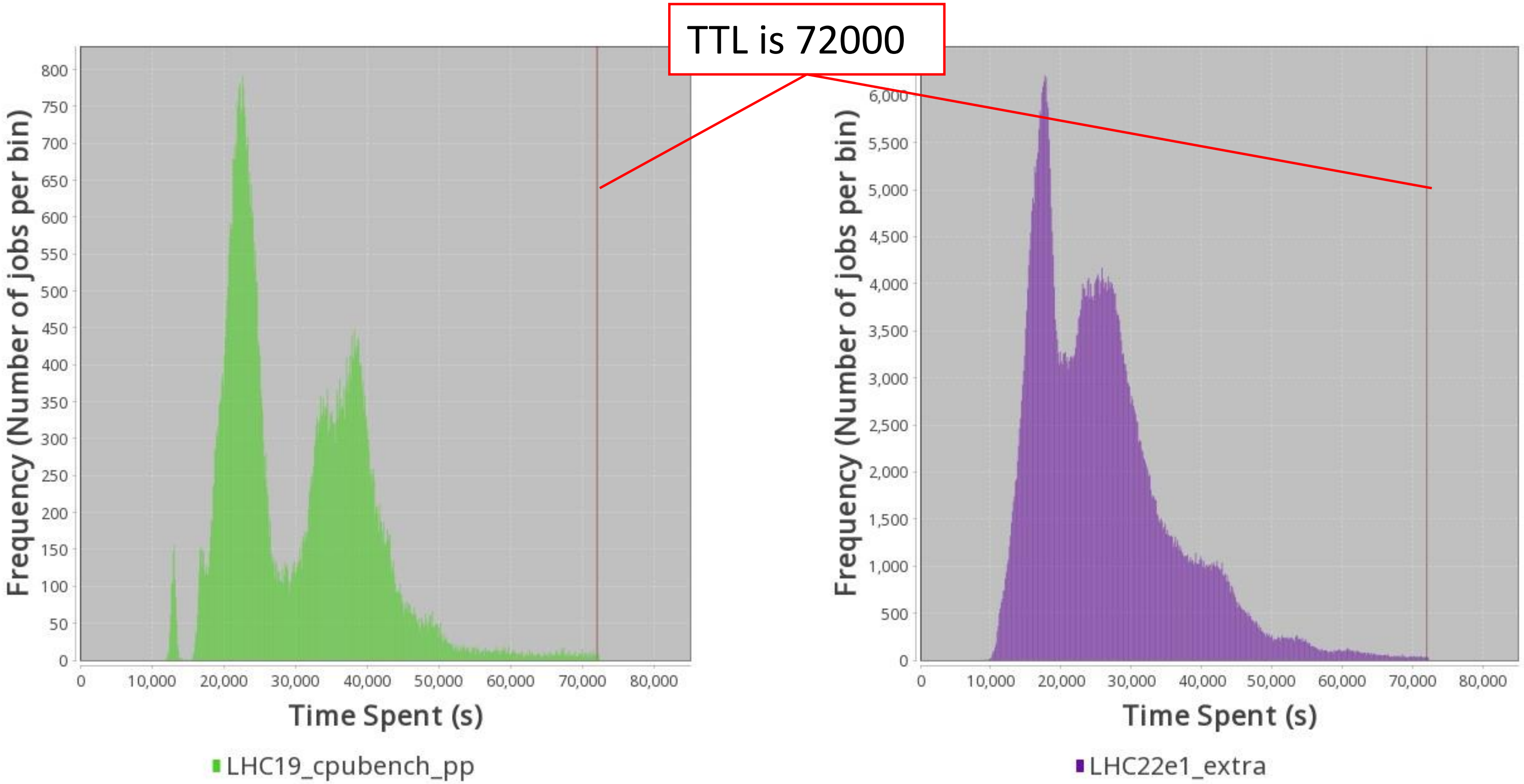
- In certain conditions, jobs finish much faster than the time they asked for



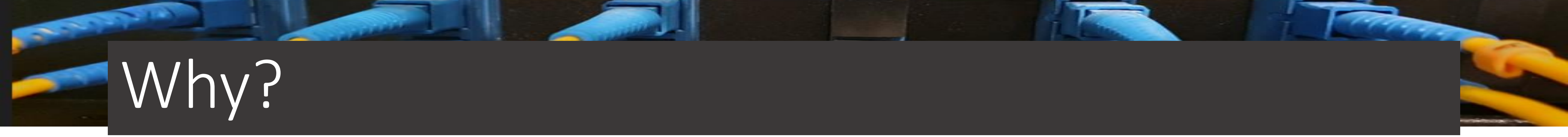
TTL is 72000

Issues

- Jobs will wait for newly created slots to be scheduled on.



TTL is 72000

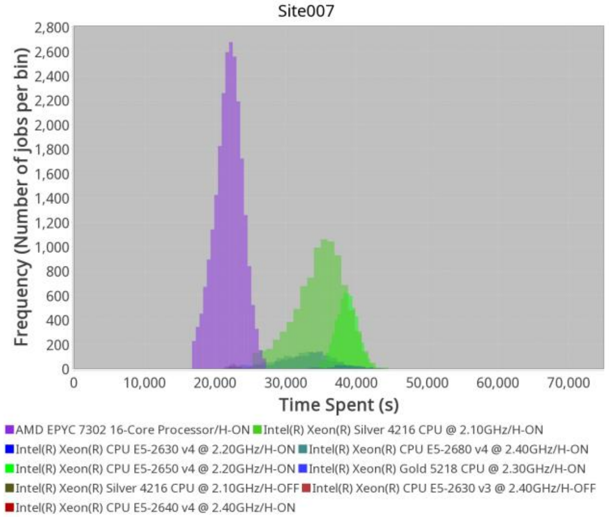


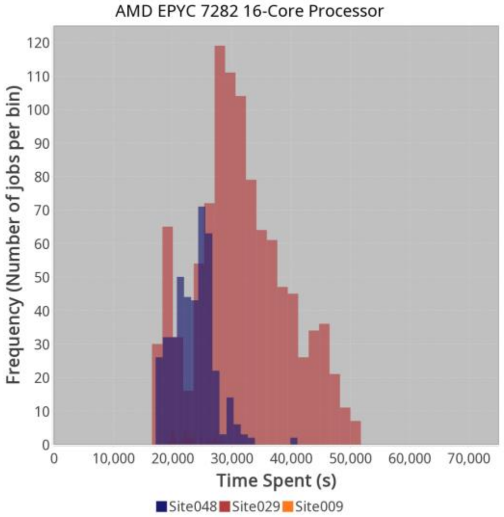
Why?

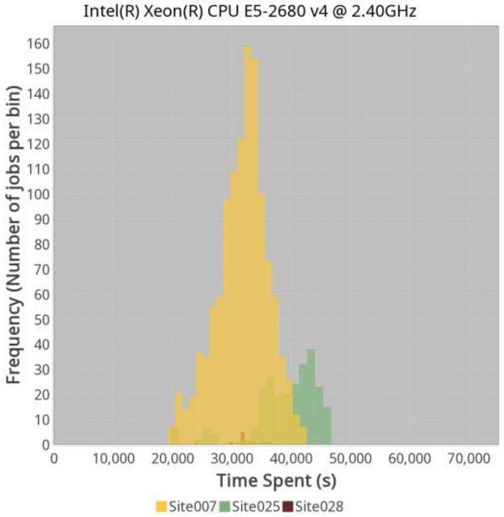
- Heterogenous resources and configurations
  - 55 sites
  - 11000 hosts
  - 15 operating systems
  - 131 CPU Model Names
  - 143 (CPU Model Name, hyperthreading) combinations
  - Different CPU flags
  - 1680 unique combinations of (CPU Model Name, microcode, CPU flags, hyperthreading, OS, RAM, SWAP)
  - + virtualization, security patches, batch queue solution, containerization

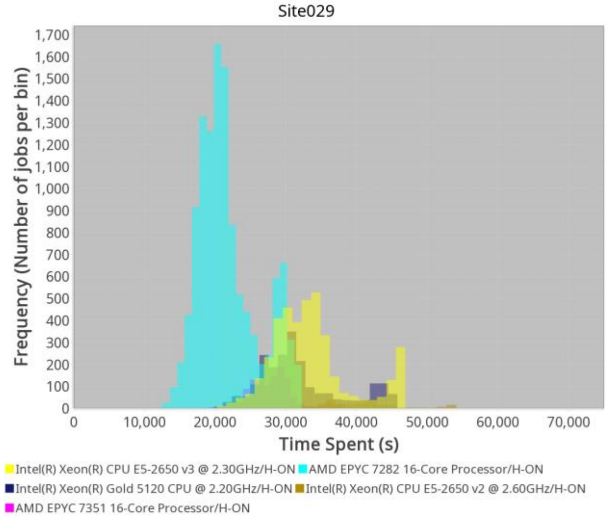
Find the common factors

- How to
  - Split datasets based on different keys
  - Check if the data distribution is normal (Gaussian model)



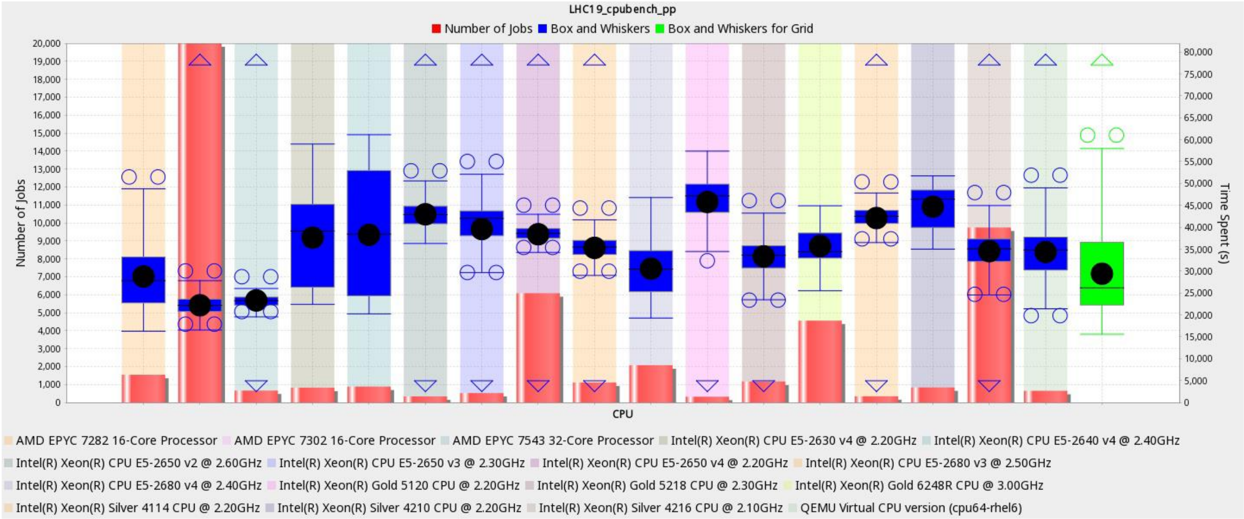






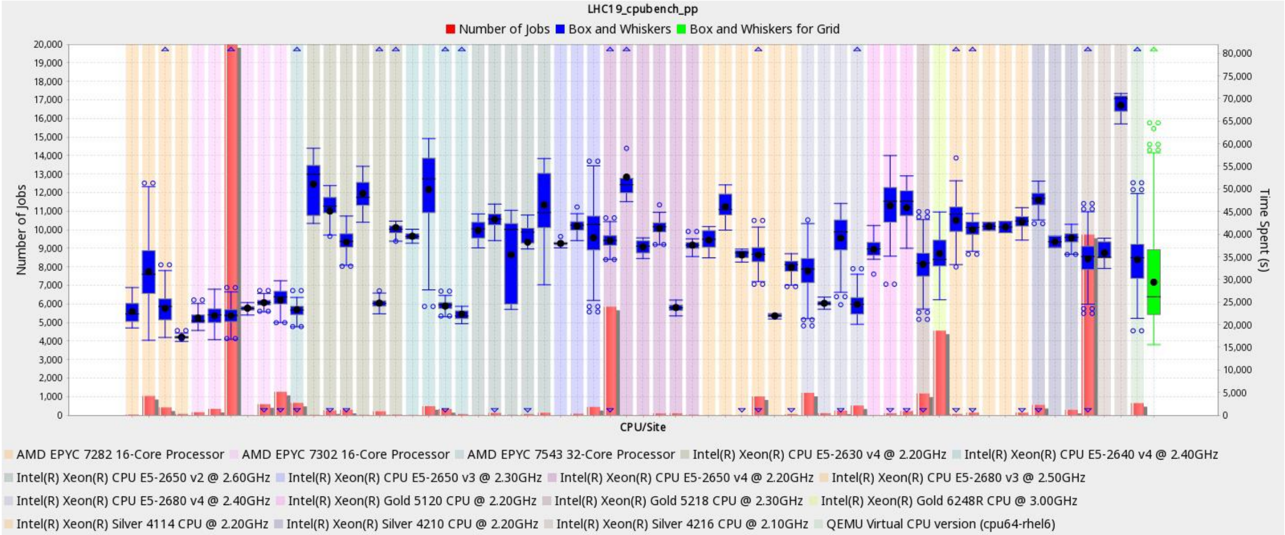
Find the common factors

- How to
  - Split datasets based on different keys
  - Check if the data distribution is normal (Gaussian model)
  - Check if the values are not spread in a wide interval



Find the common factors

- How to
  - Split datasets based on different keys
  - Check if the data distribution is normal (Gaussian model)
  - Check if the values are not spread in a wide interval



Job TTL Optimization

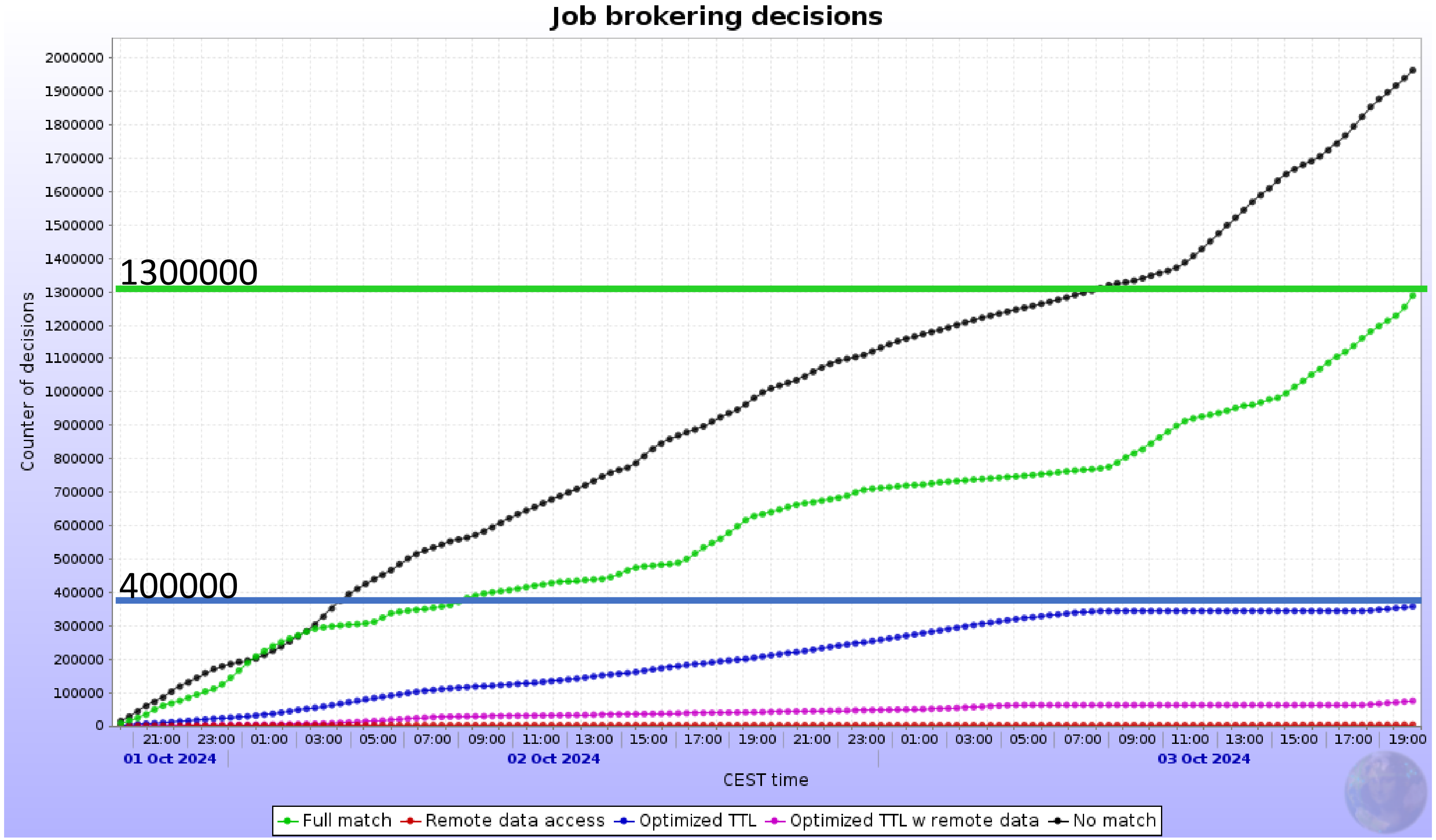
- Predict TTL based on the history of the job duration for the same:
  - Production Id
  - CPU Model Name
  - Site

- Keep history metadata in database – TTLPredictionHistoryTable

Job TTL Optimization in action

- Step 1: Collect history
  - Each job that finishes its execution will update the job history table
  - Compute the predicted TTL in advance
- Step 2: Modify the Job Scheduling process
  - Try default matching option – if this doesn’t work then
  - Try to alter the TTL based on the job history table
- Step 3: Alter Job’s JDL file and send it to the Grid site

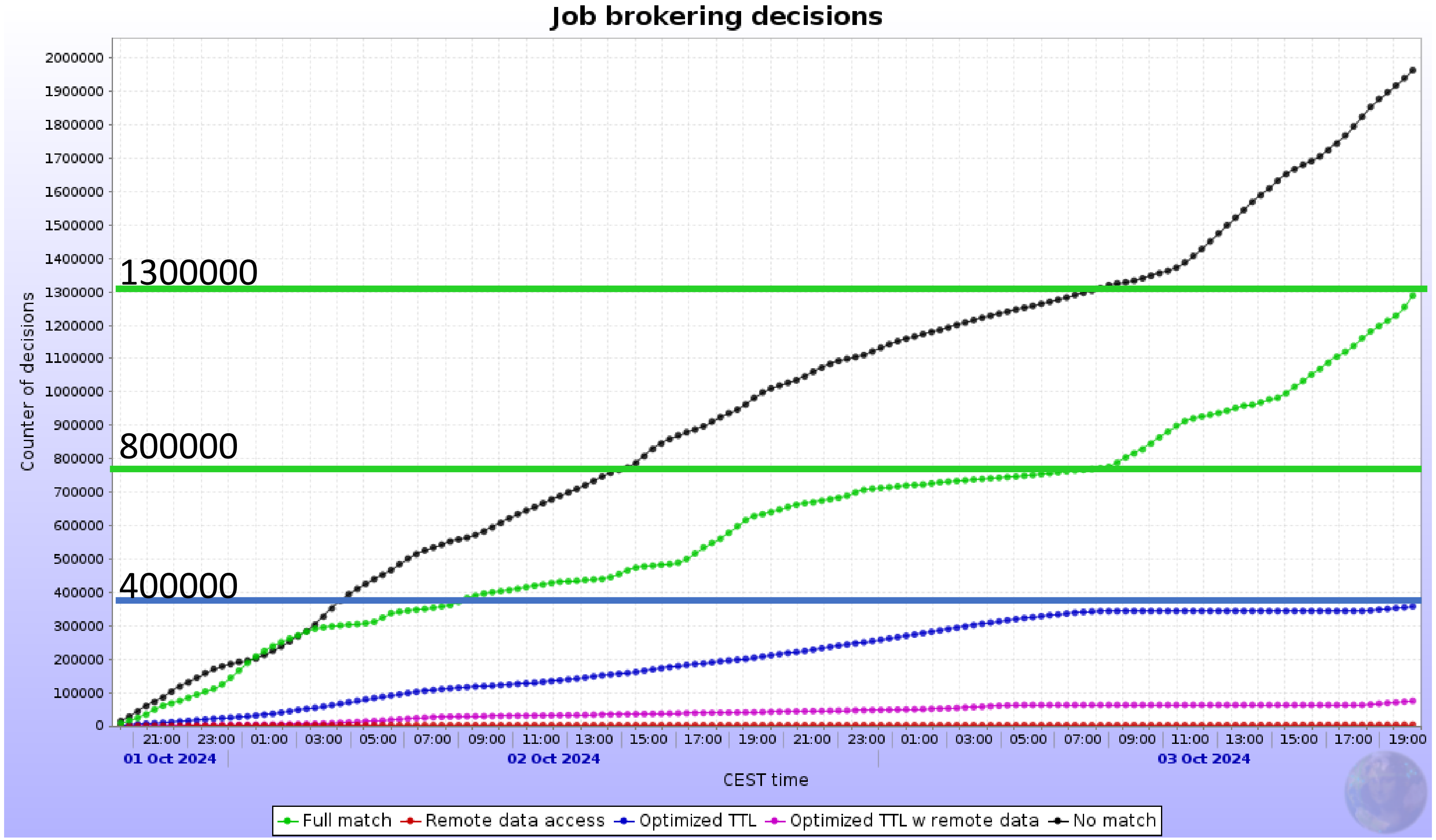
Results



1300000

400000

Results


Job brokering decisions
Counter of decisions
2000000
1900000
1800000
1700000
1600000
1500000
1400000
1300000
1200000
1100000
1000000
900000
800000
700000
600000
500000
400000
300000
200000
100000
0
1300000
800000
400000
21:00
23:00
01:00
03:00
05:00
07:00
09:00
11:00
13:00
15:00
17:00
19:00
21:00
23:00
01:00
03:00
05:00
07:00
09:00
11:00
13:00
15:00
17:00
19:00
01 Oct 2024
02 Oct 2024
03 Oct 2024
CEST time
Full match
Remote data access
Optimized TTL
Optimized TTL w remote data
No match