

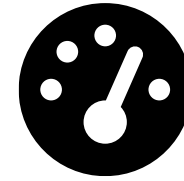
Verifying Performance with the RWTH Job Monitoring

Christian Wassermann¹

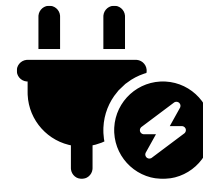
¹ IT Center, RWTH Aachen University, Germany

CLAIX-2025 Porting and Tuning Workshop in Aachen, Germany

Working With New Cluster Hardware



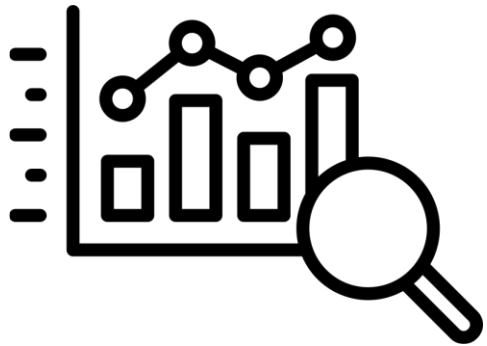
Higher execution performance



Improved energy efficiency



Requires code porting



- Measurements needed to assess performance characteristics
- Some applications or tools might be non-trivial to configure ... and sometimes results can be non-trivial to analyze
- Solution: always-on background monitoring accessible in the browser

Agenda

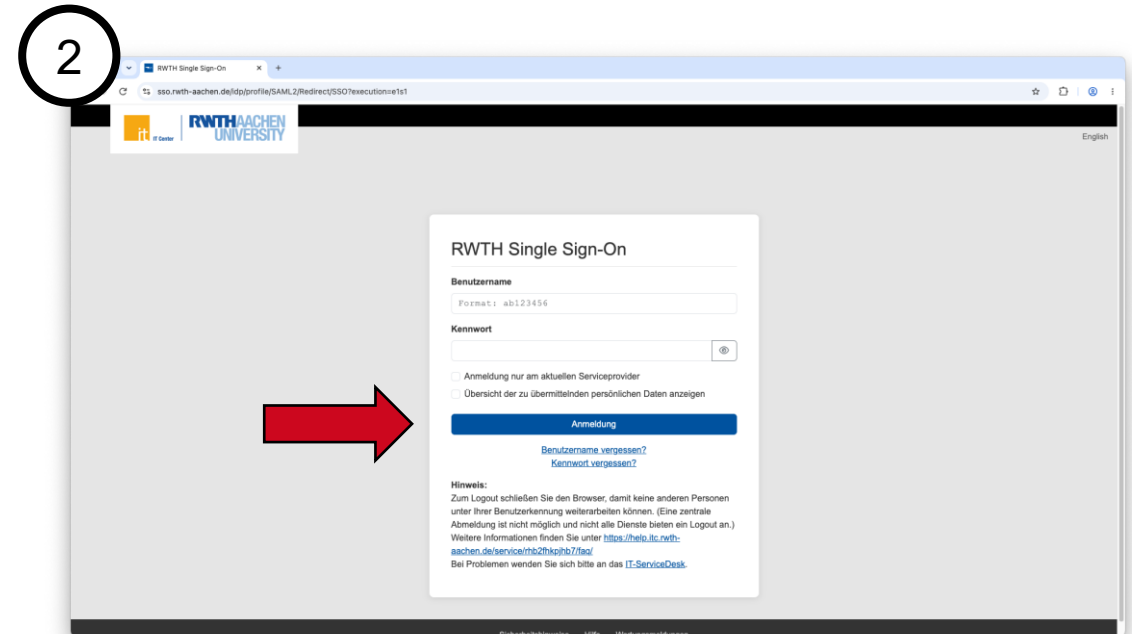
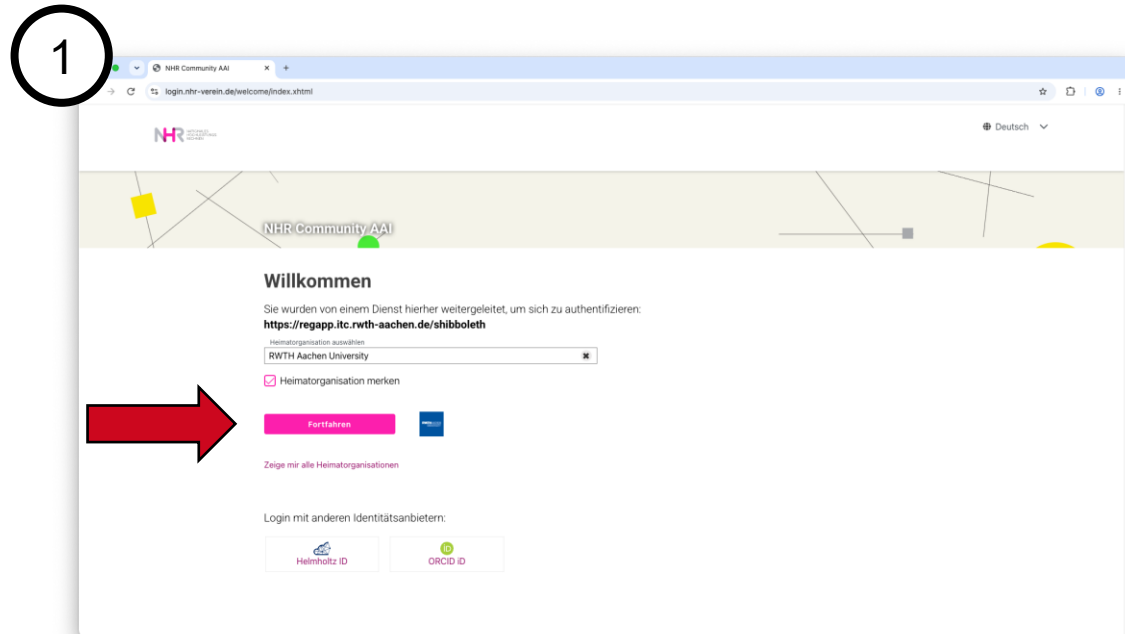
- Accessing the Monitoring System
 - The Webportal
 - Joblist Overview
 - Job Details
 - Caveats and Gotchas
- Detecting Misconfiguration
 - Wrong Core Counts
 - Single Core Binding
 - Underutilization of Resources
- Try It Yourself

Agenda

- **Accessing the Monitoring System**
 - **The Webportal**
 - **Joblist Overview**
 - **Job Details**
 - **Caveats and Gotchas**
- Detecting Misconfiguration
 - Wrong Core Counts
 - Single Core Binding
 - Underutilization of Resources
- Try It Yourself

Accessing the Monitoring System – The Webportal

- Online documentation @ [help.itc](https://help.itc.rwth-aachen.de)
- Visit Grafana @ perfmon.hpc.itc.rwth-aachen.de



! Single Sign-On credentials (not HPC account)

Accessing the Monitoring System – Joblist Overview

- History of finished and running Jobs

3

Filter by cluster partition

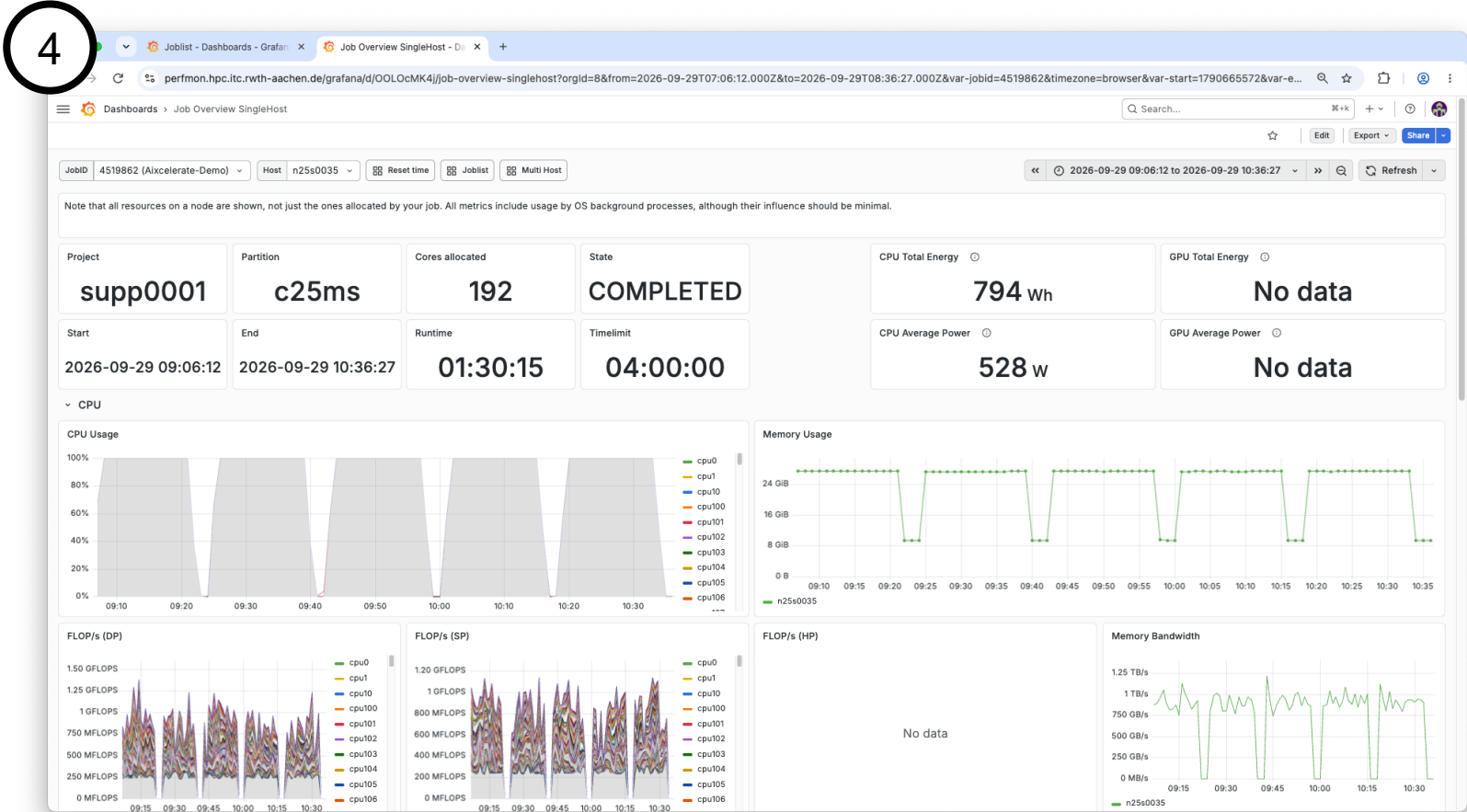
Sort by runtime

Increase time range

Access job details

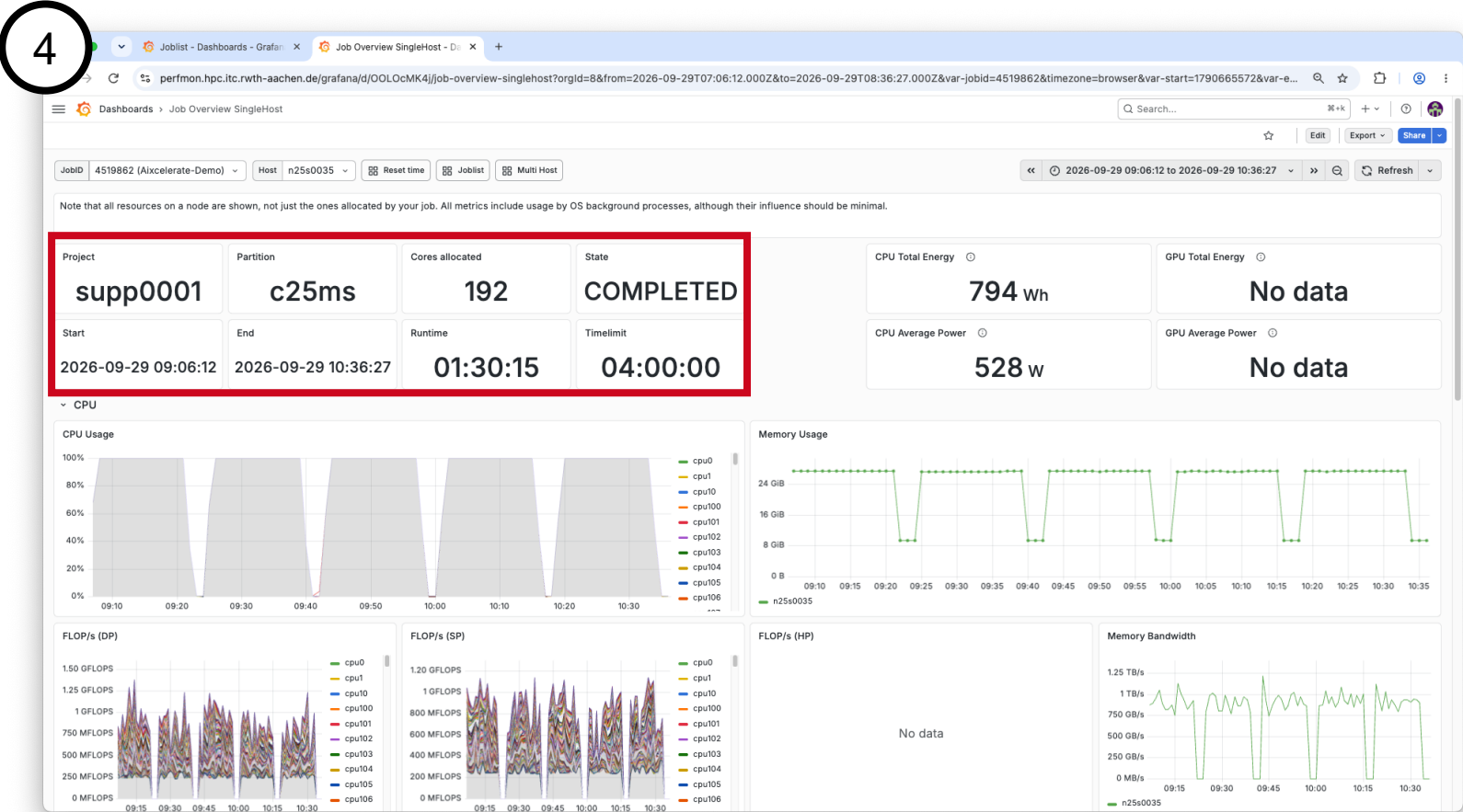
JobID	JobName	Partition	Start	End	Runtime	JobState	NumNodes	Nodelist
4521014	Aixcelerate-Demo	c25ms	2026-09-29 10:01:14	2026-09-29 11:27:11	1.43 hours	COMPLETED	1	n25s0054
4519862	Aixcelerate-Demo	c25ms	2026-09-29 09:06:12	2026-09-29 10:36:27	1.50 hours	COMPLETED	1	n25s0035
4519863	Aixcelerate-Demo	c25ms	2026-09-29 09:06:12	2026-09-29 11:51:02	2.75 hours	COMPLETED	1	n25s0036
4519864	Aixcelerate-Demo	c25ms	2026-09-29 09:06:12	2026-09-29 13:08:28	4.04 hours	TIMEOUT	1	n25s0038
4519836	Aixcelerate-Demo	c25ms	2026-09-29 09:02:00	2026-09-29 10:32:15	1.50 hours	COMPLETED	1	n25s0033
4519837	Aixcelerate-Demo	c25ms	2026-09-29 09:02:00	2026-09-29 13:04:25	4.04 hours	TIMEOUT	1	n25s0034
4506521	Aixcelerate-Demo	c25ms	2026-09-28 17:46:04	2026-09-28 18:48:10	1.03 hours	TIMEOUT	1	n25s0150
4505573	Aixcelerate-Demo	c25ms	2026-09-28 17:27:17	2026-09-28 17:42:22	15.1 mins	COMPLETED	1	n25s0087
4505335	Aixcelerate-Demo	c25ms	2026-09-28 17:20:47	2026-09-28 17:46:06	25.3 mins	CANCELLED	1	n25s0104
4505147	Aixcelerate-Demo	c25ms	2026-09-28 17:12:32	2026-09-28 17:20:49	8.28 mins	CANCELLED	1	n25s0087
4505039	Aixcelerate-Demo	c25ms	2026-09-28 17:08:13	2026-09-28 17:12:08	3.92 mins	CANCELLED	1	n25s0097
4206185	test	c25ms	2026-09-17 21:48:59	2026-09-17 21:49:15	16 s	COMPLETED	1	n25s0067
4206160	test	c25ms	2026-09-17 21:45:26	2026-09-17 21:45:28	2 s	COMPLETED	1	n25s0067
4206161	test	c25ms	2026-09-17 21:45:26	2026-09-17 21:45:28	2 s	COMPLETED	1	n25s0091
4206129	test	c25ms	2026-09-17 21:41:53	2026-09-17 21:42:08	15 s	COMPLETED	1	n25s0067
4206130	test	c25ms	2026-09-17 21:41:53	2026-09-17 21:41:57	4 s	COMPLETED	1	n25s0091
4206131	test	c25ms	2026-09-17 21:41:53	2026-09-17 21:41:58	5 s	COMPLETED	1	n25s0098
4206132	test	c25ms	2026-09-17 21:41:53	2026-09-17 21:41:57	4 s	COMPLETED	1	n25s0116
4206121	test	c25ms	2026-09-17 21:39:22	2026-09-17 21:39:37	15 s	COMPLETED	1	n25s0067
4206117	test	c25ms	2026-09-17 21:38:51	2026-09-17 21:38:54	3 s	COMPLETED	1	n25s0067

Accessing the Monitoring System – Job Details

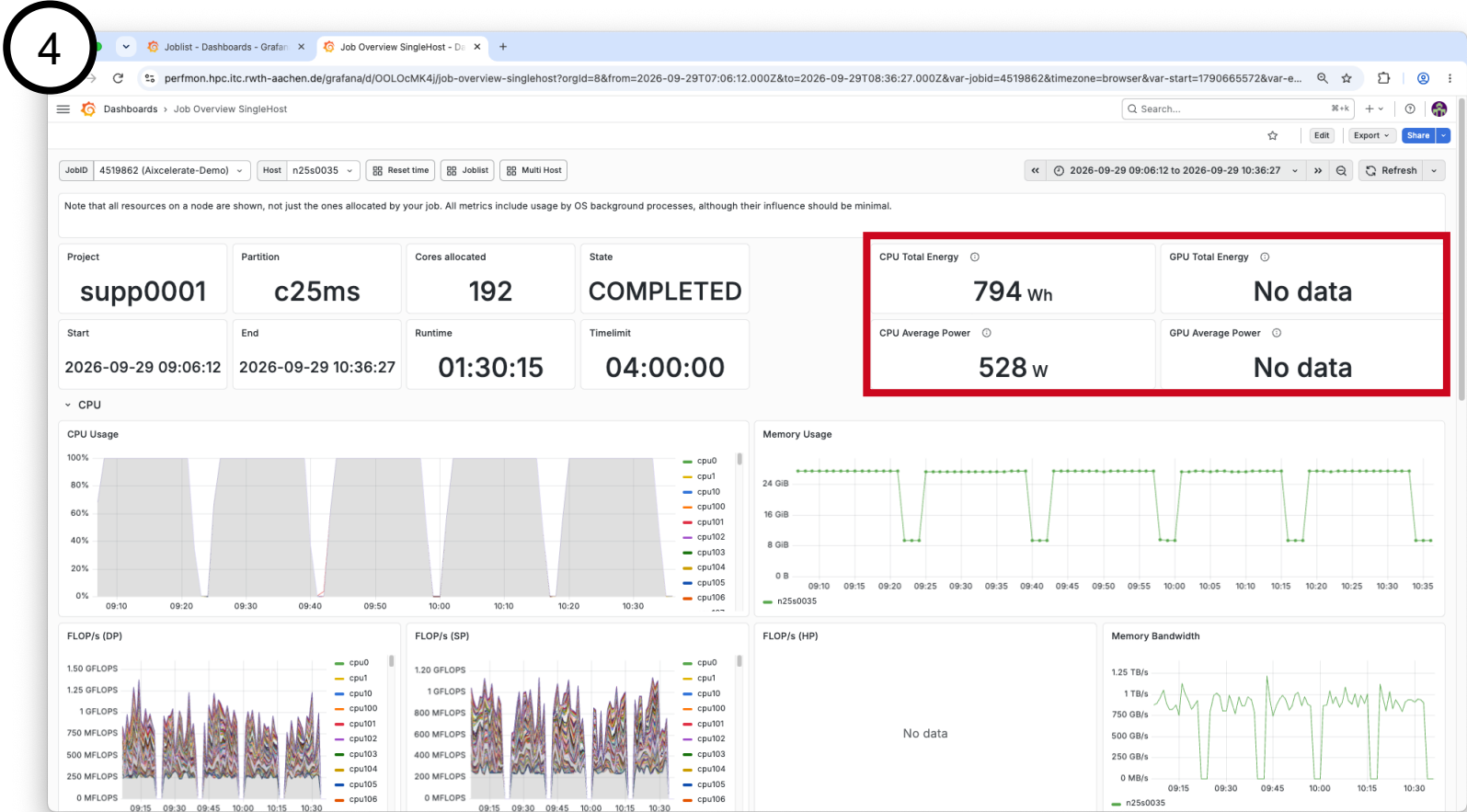


Accessing the Monitoring System – Job Details

Basic job
meta data



Accessing the Monitoring System – Job Details

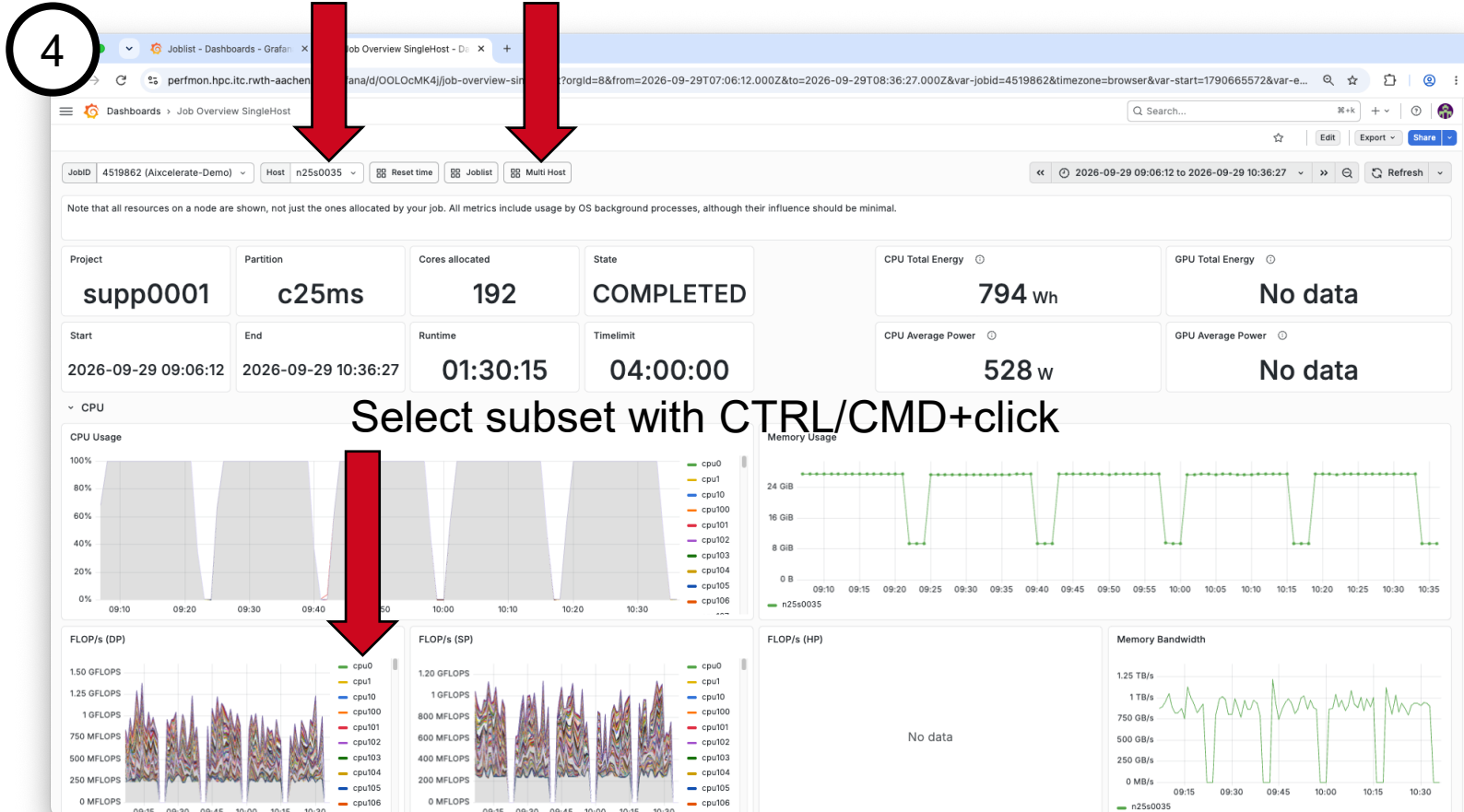


Energy data
(„No data“ for GPU
due to CPU node)

Accessing the Monitoring System – Job Details

Select different host

Switch single-/multi-host view



Accessing the Monitoring System – Caveats and Gotchas

- Sampling frequency: 1 sample / minute
 - Almost no influence on application runtime
 - ⚠ Short application stages can be missed
- Metric-specific sampling quality:
 - Sampling length of 1 second for CPU FLOP/s and CPU memory bandwidth
 - ⚠ This data is missing with `#SBATCH --hwctr=OPTION` with `OPTION ∈ { vtune, vtuneperf, likwid, papi }`
 - Continuous measurement for power, GPU, fabric, Lustre, NFS
 - Collection of current value for CPU, memory
- Shared node usage
 - ⚠ Data may be influenced by behavior of other jobs
 - View resources of job: `$ scontrol show job -d "${SLURM_JOB_ID}" > "slurm-${SLURM_JOB_ID}.job"`
 - Force exclusive node allocation with `#SBATCH --exclusive`
- Power and energy measurements
 - For CPUs measured with RAPL interface → for details see [here](#) and [here](#)
 - For GPUs measured with NVML's `nvmDeviceGetTotalEnergyConsumption`
 - ⚠ Data quality and availability limited by tools of hardware vendors

Agenda

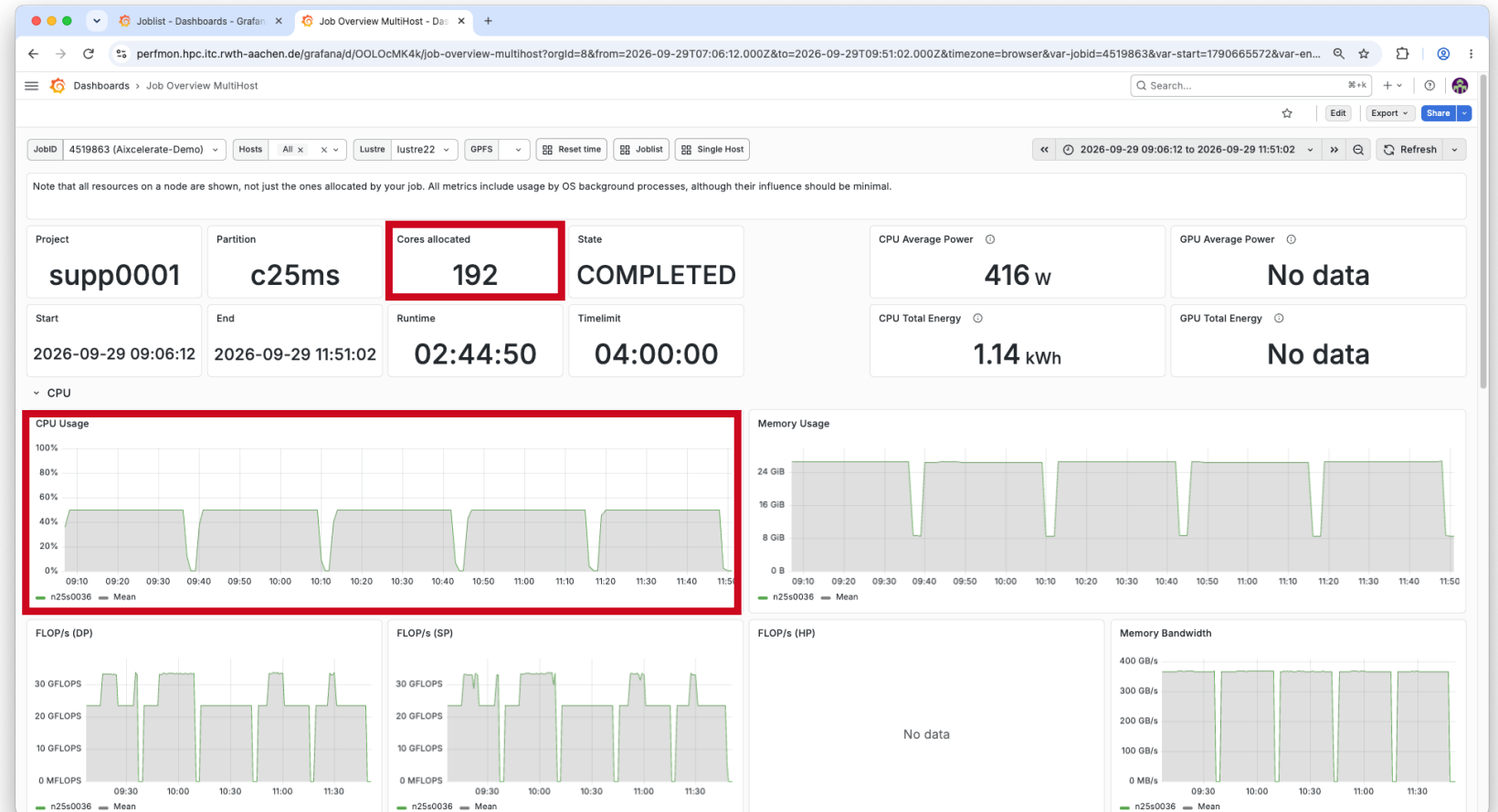
- Accessing the Monitoring System
 - The Webportal
 - Joblist Overview
 - Job Details
 - Caveats and Gotchas
- **Detecting Misconfiguration**
 - **Wrong Core Counts**
 - **Single Core Binding**
 - **Underutilization of Resources**
- Try It Yourself

Detecting Misconfiguration – Wrong Core Counts

✓ 192 cores allocated

⚠ only 50% CPU usage

- Lessons learned:
 - CLAIX-2023: 96 cores / node
 - CLAIX-2025:
 - c25ms: 192 cores / node
 - c25ml: 192 cores / node
 - c25g: 128 cores / node
- Adjust hard-coded values for
 - `OMP_NUM_THREADS`
 - `mpiexec -n`
 - `srun -n`

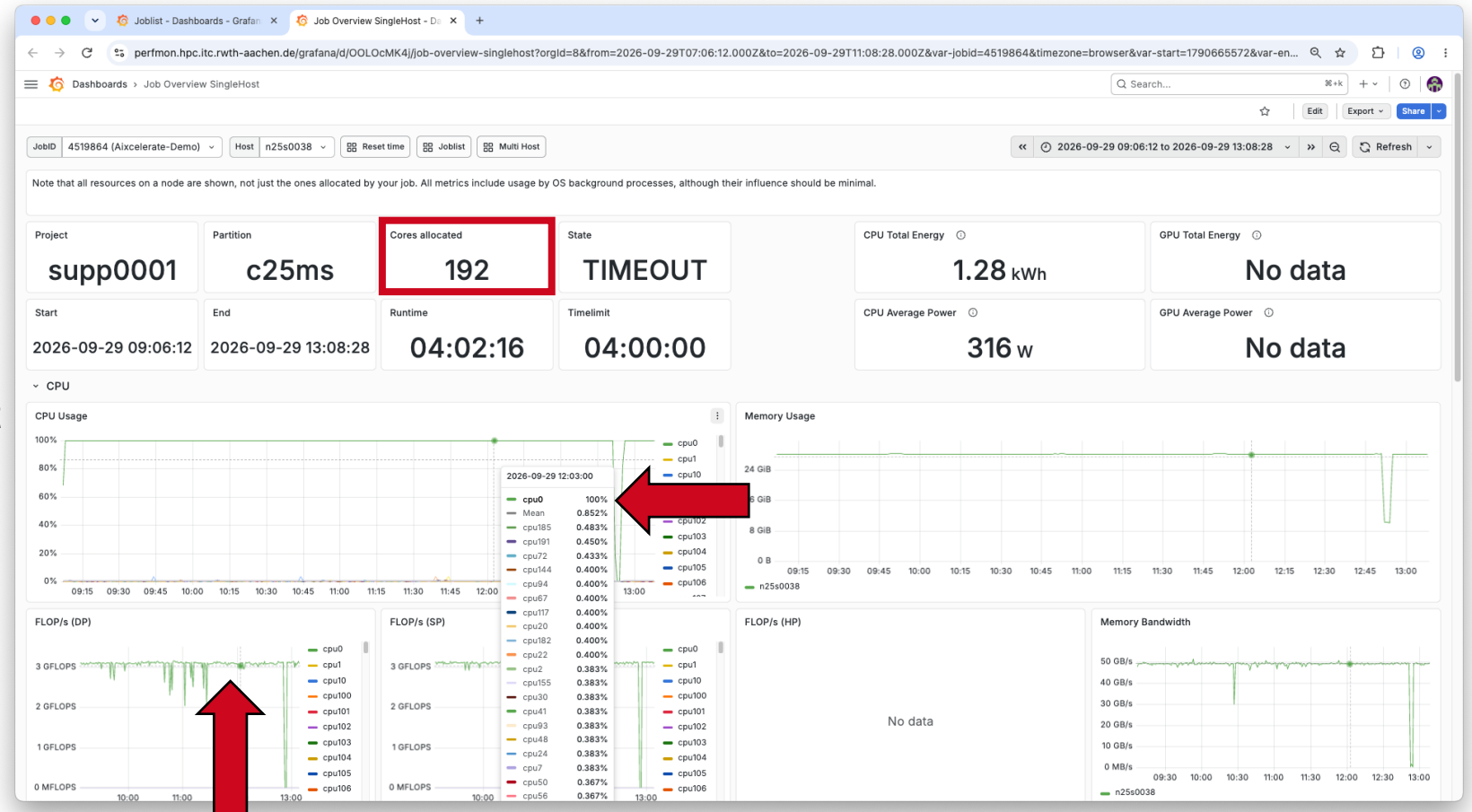


Detecting Misconfiguration – Single Core Binding

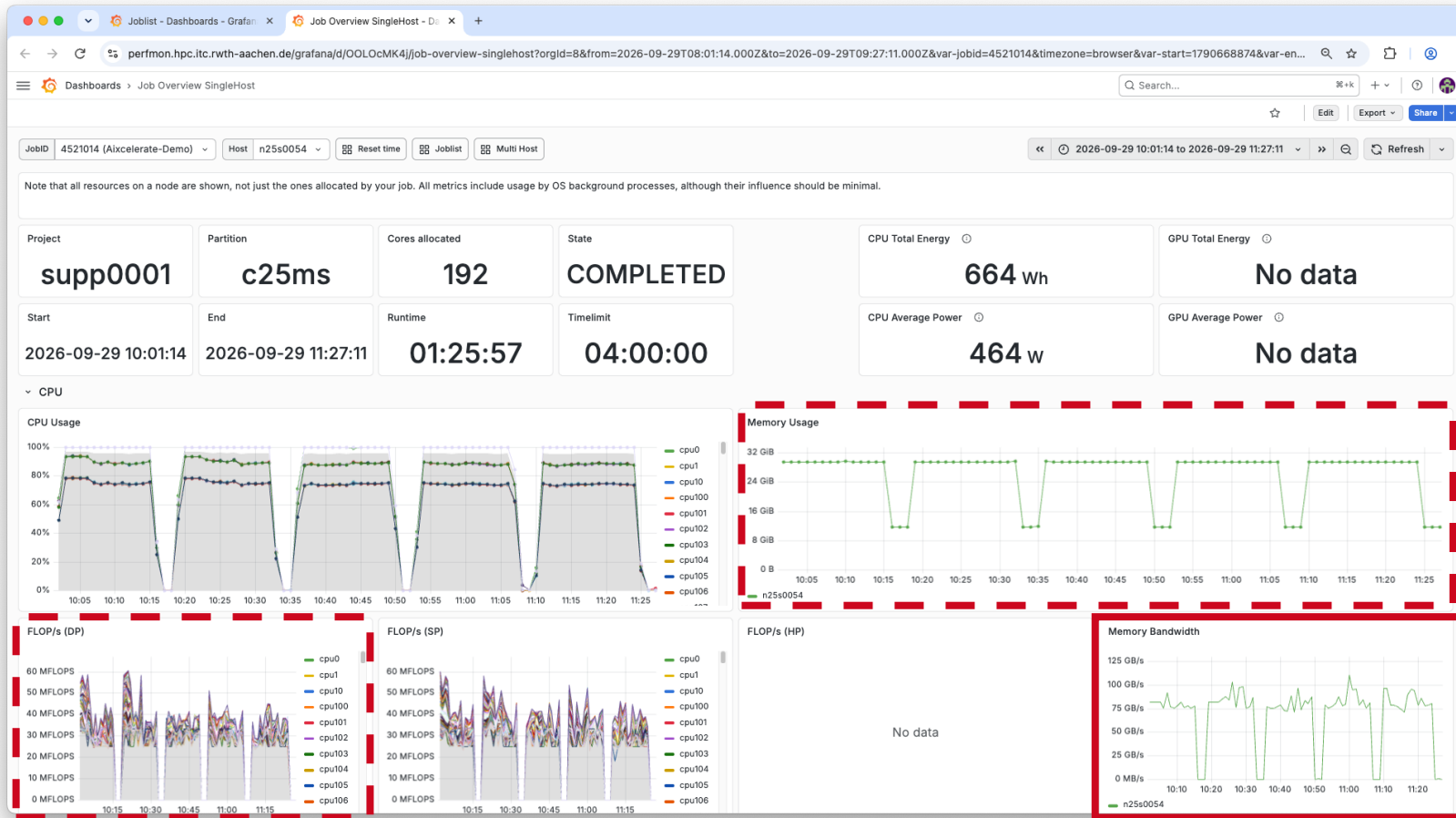
✓ 192 cores allocated

⚠ only a single core used

- Lessons learned:
 - Check thread / process placement
 - `OMP_DISPLAY_AFFINITY=true`
 - `mpiexec --report-bindings`
 - `mpiexec -print-rank-map`
 - `srun --cpu-bind verbose`
- Consult documentation for
 - `OMP_PLACES`
 - `mpiexec`
 - `srun`



Detecting Misconfiguration – Underutilization of Resources



- Reference run for comparison:

- ⚠ 75 GB/s vs. 900 GB/s



- Previous experience with job:
 - “Usually 90% of peak bandwidth”
 - “Before > 200 MFLOP/s occurred”

- Talk to us about:
 - Expectations for CLAIX-2025
 - Deviations from previous results

Agenda

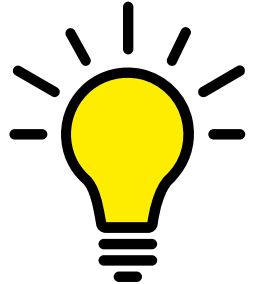
- Accessing the Monitoring System
 - The Webportal
 - Joblist Overview
 - Job Details
 - Caveats and Gotchas
- Detecting Misconfiguration
 - Wrong Core Counts
 - Single Core Binding
 - Underutilization of Resources
- **Try It Yourself**

Try It Yourself

- Monitoring data is valuable for verifying high-level performance trends
- Always-on collection enables the analysis of historic jobs (even if not initially intended)
- Now it is your turn:
 - go to perfmon.hpc.itc.rwth-aachen.de,
 - scan the QR code, or
 - click the logo



**Missing something?
New ideas are welcome!**



References

Acknowledgments:

Publications:

Links:

Images:

- CLAIX-2025 from https://blog.rwth-aachen.de/itc/files/2026/07/CLAIX_25_Blog-scaled.jpg (own illustration)
- Gauge High by Font Awesome from <https://fontawesome.com/icons/gauge-high> (CC BY 4.0 License)
- Plug Circle Bolt by Font Awesome from <https://fontawesome.com/icons/plug-circle-bolt> (CC BY 4.0 License)
- File Code by Font Awesome from <https://fontawesome.com/icons/file-code> (CC BY 4.0 License)
- Measure by littlestar23 from <https://thenounproject.com/icon/measure-6414403> (CC BY 3.0 License)
- Grafana from https://github.com/grafana/grafana/blob/5f4490098b3555e5e9ce6d0a59c07d7463eaeadc/public/img/grafana_icon.svg (AGPL-3.0 License)
- Meeting by Hermine Blanquart from <https://thenounproject.com/icon/meeting-4220250> (CC BY 3.0 License)
- Ideas by Ian Ransley from <https://thenounproject.com/icon/ideas-1879490> (CC BY 3.0 License)