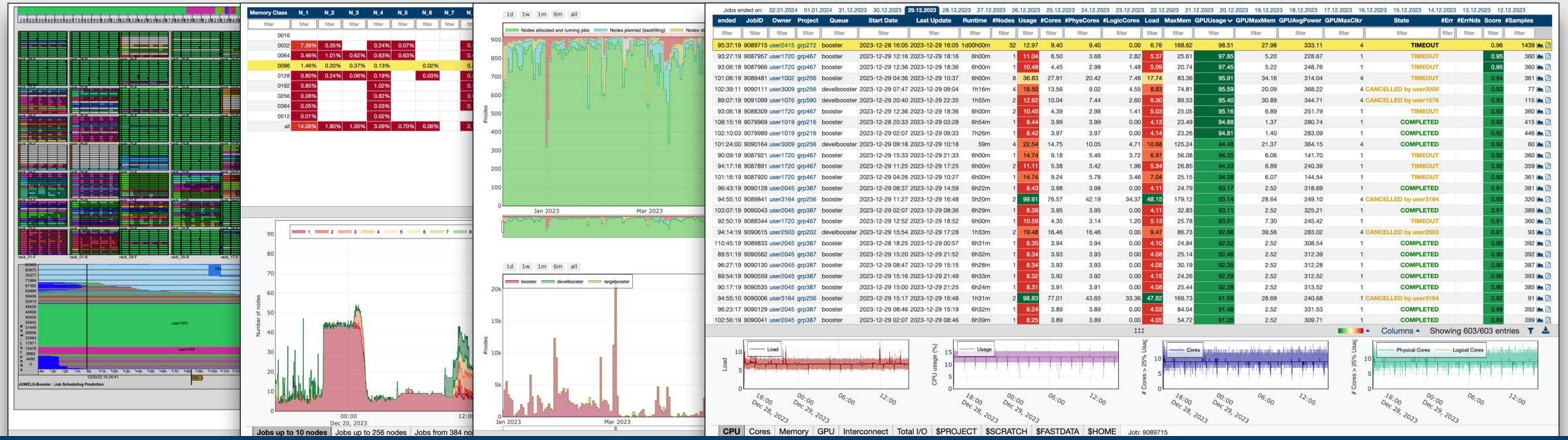


Supporting HPC Users with LLview

6th ISC HPC International Workshop on “Monitoring & Operational Data Analytics”

13.06.2025 I **FILIPE GUIMARÃES**, ARAVIND SANKARAN, WOLFGANG FRINGS | HAMBURG

Member of the Helmholtz Association

Presentation DOI: 10.5281/zenodo.15681176

JÜLICH Forschungszentrum

What is operational data analytics? Why do ~~we~~ users need it?

Disclaimer: from the point of view of the support team

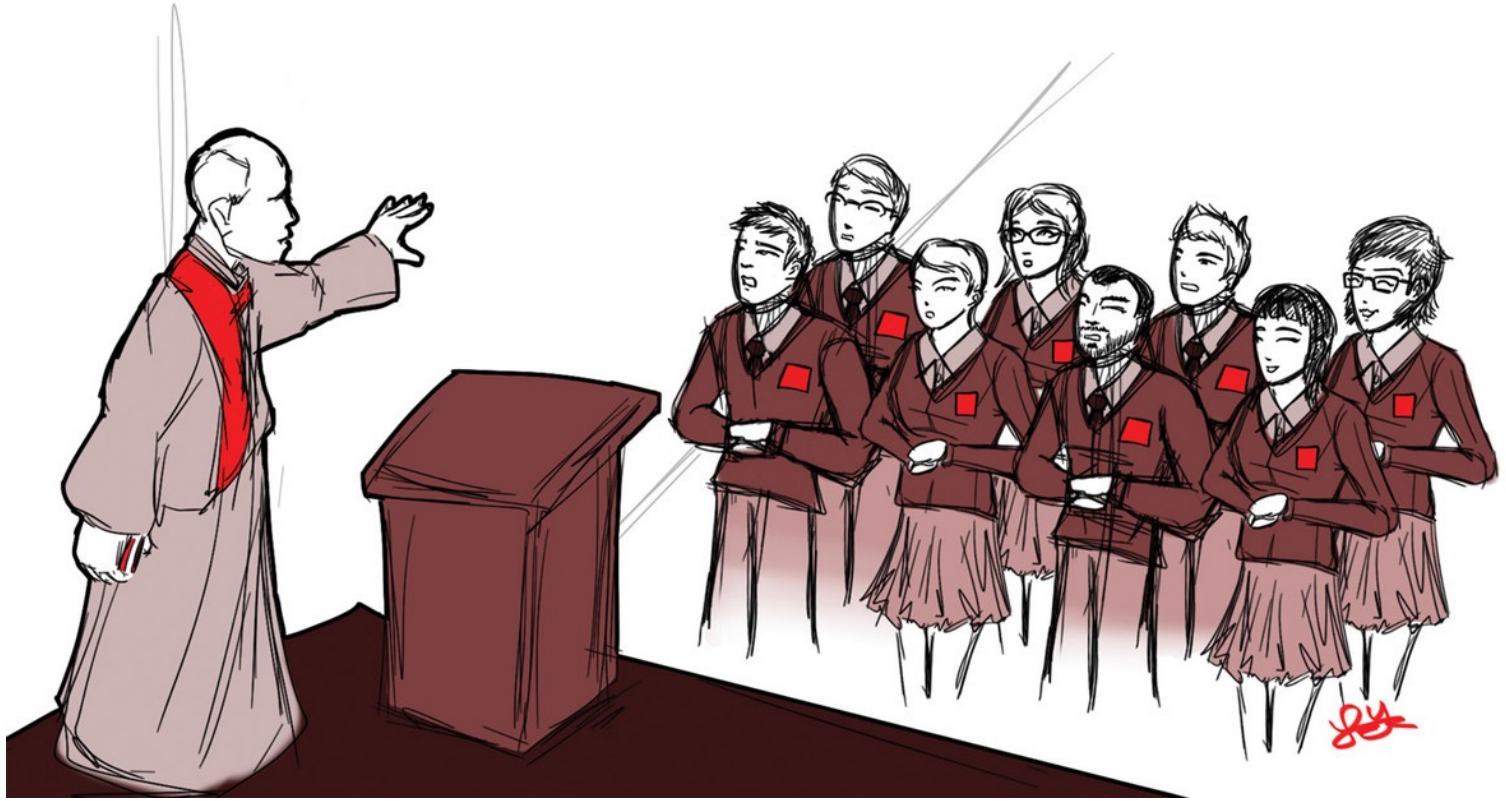
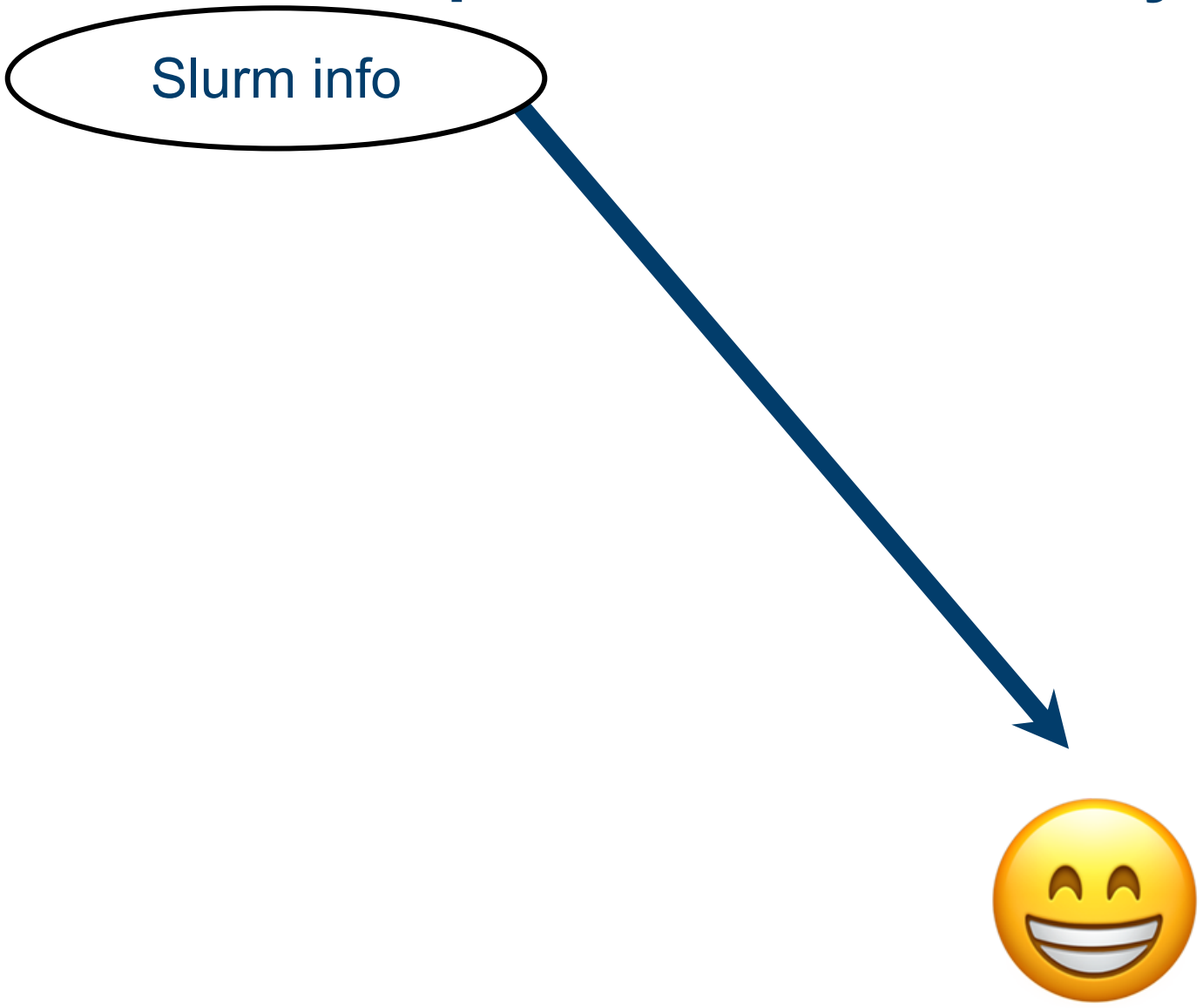
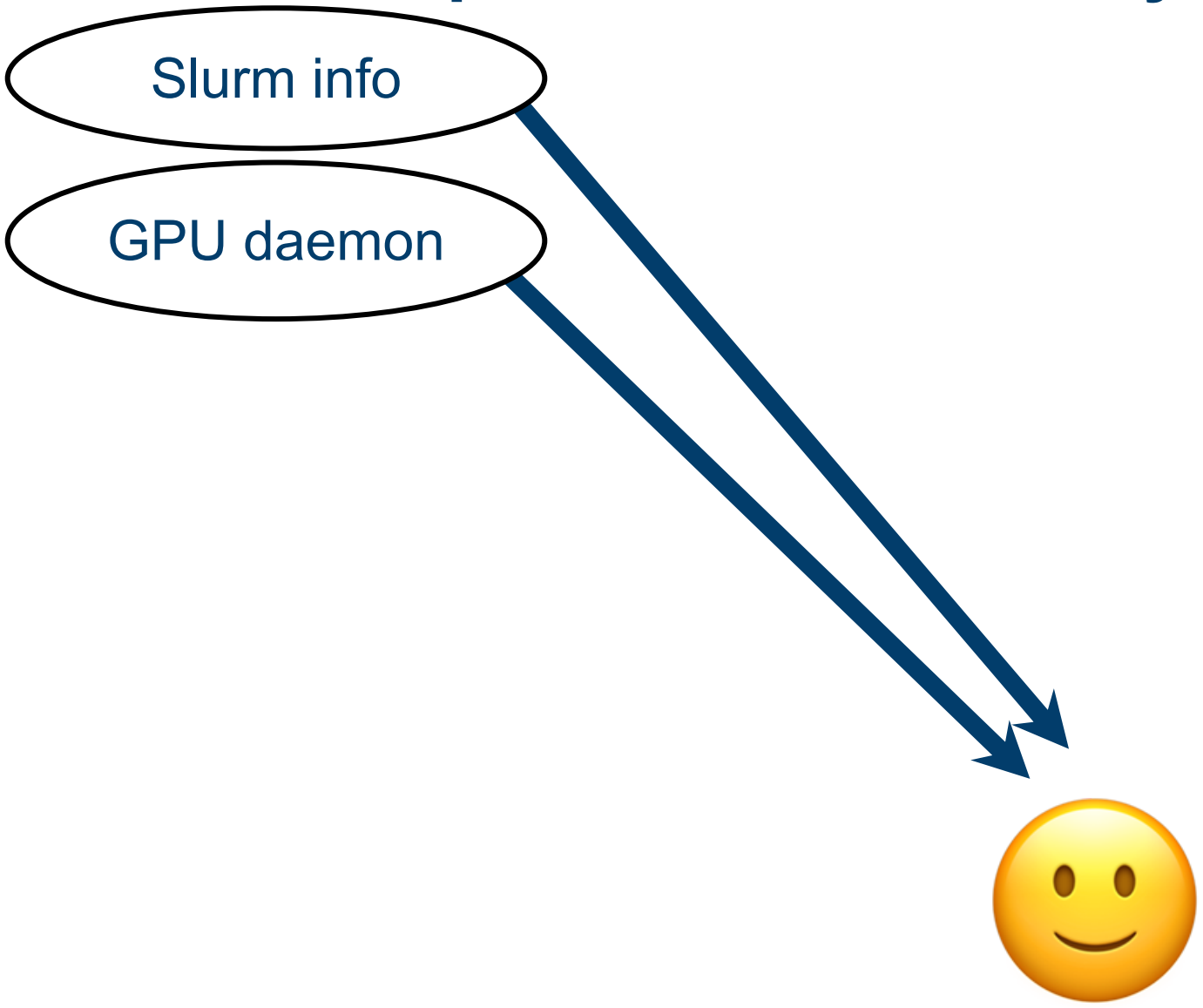


Illustration from: https://www.linkedin.com/pulse/preaching-choir-worship-church-services-steve-cranston-ma/

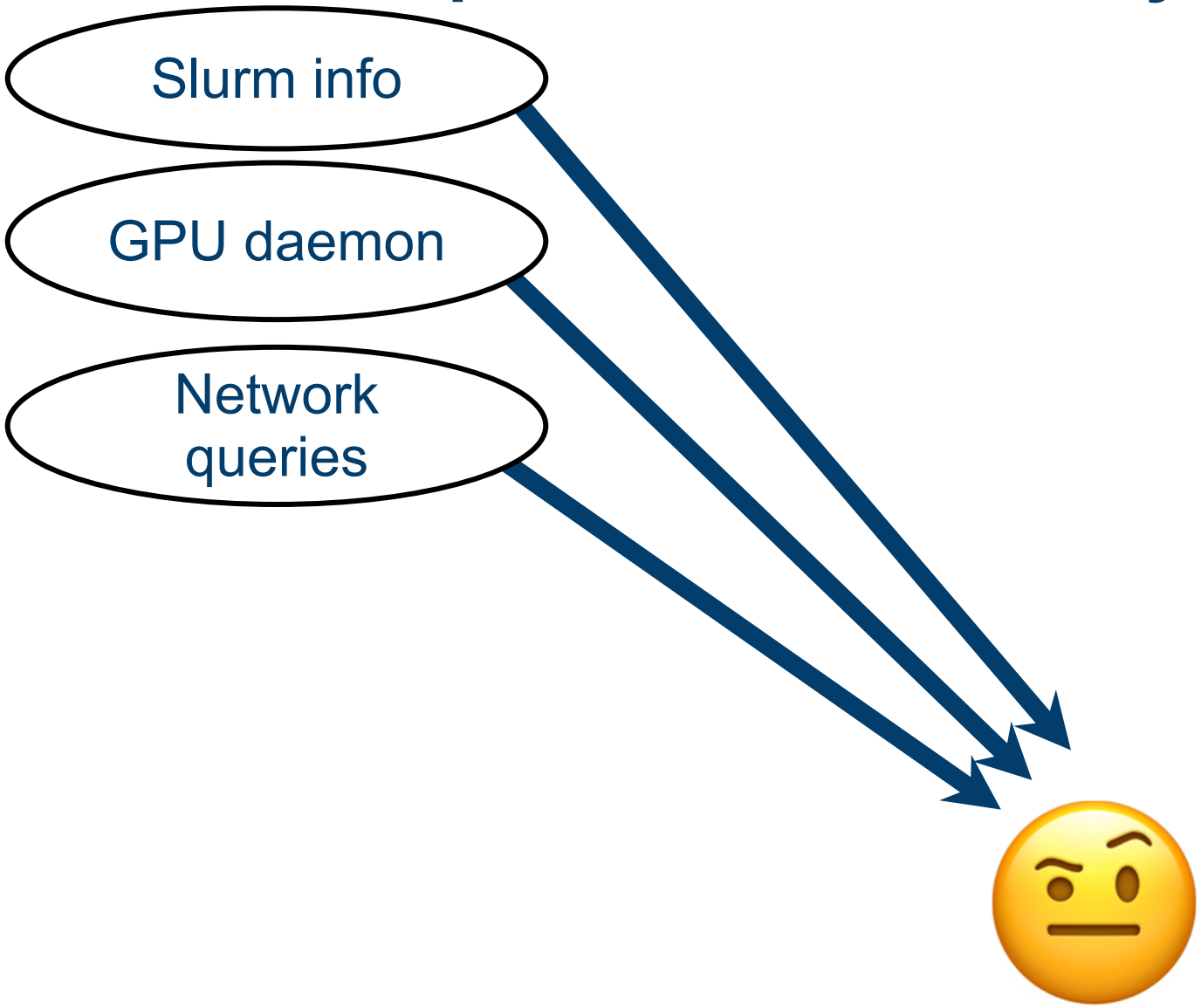
What is operational data analytics? Why do ~~we~~ users need it?



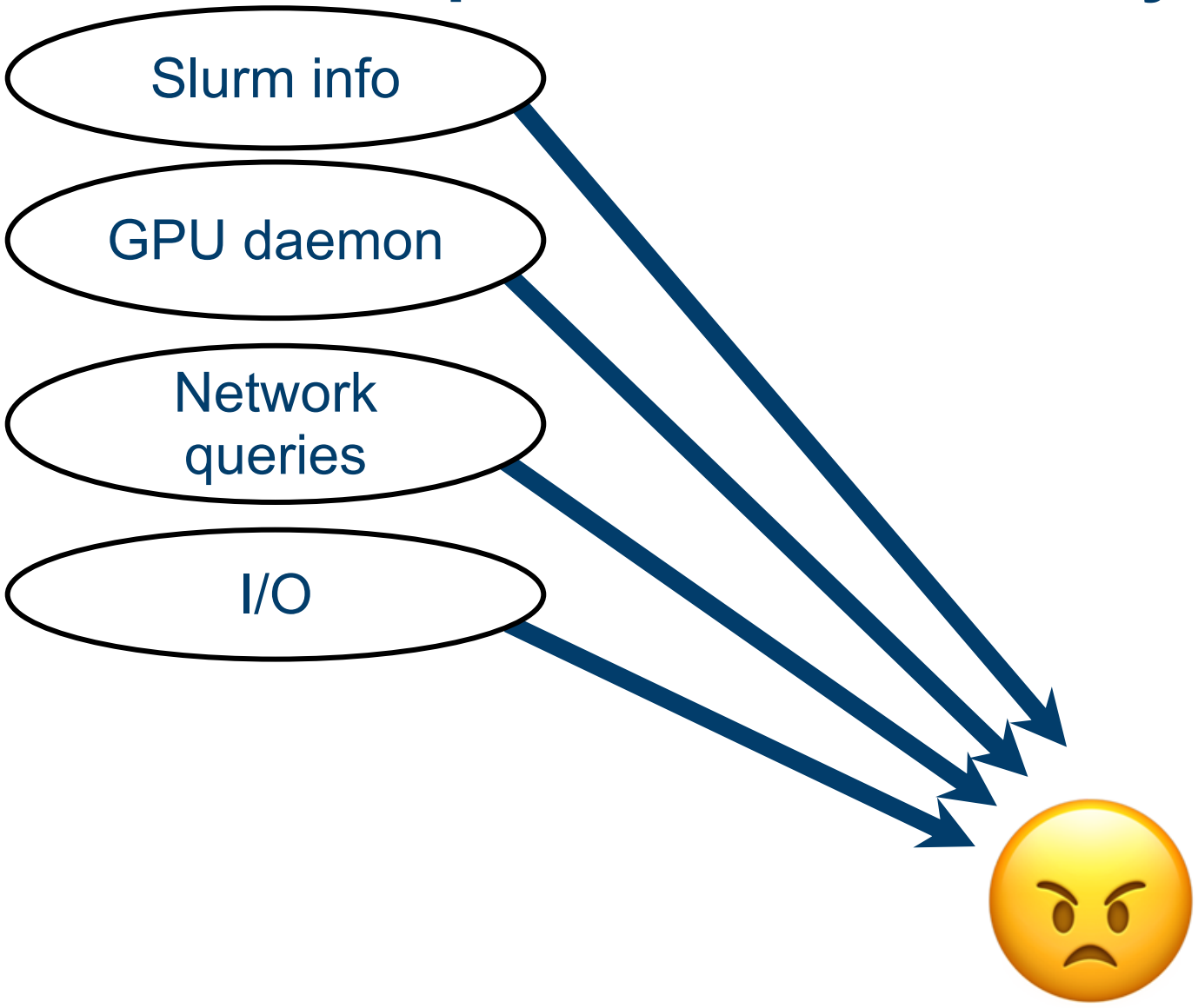
What is operational data analytics? Why do ~~we~~ users need it?



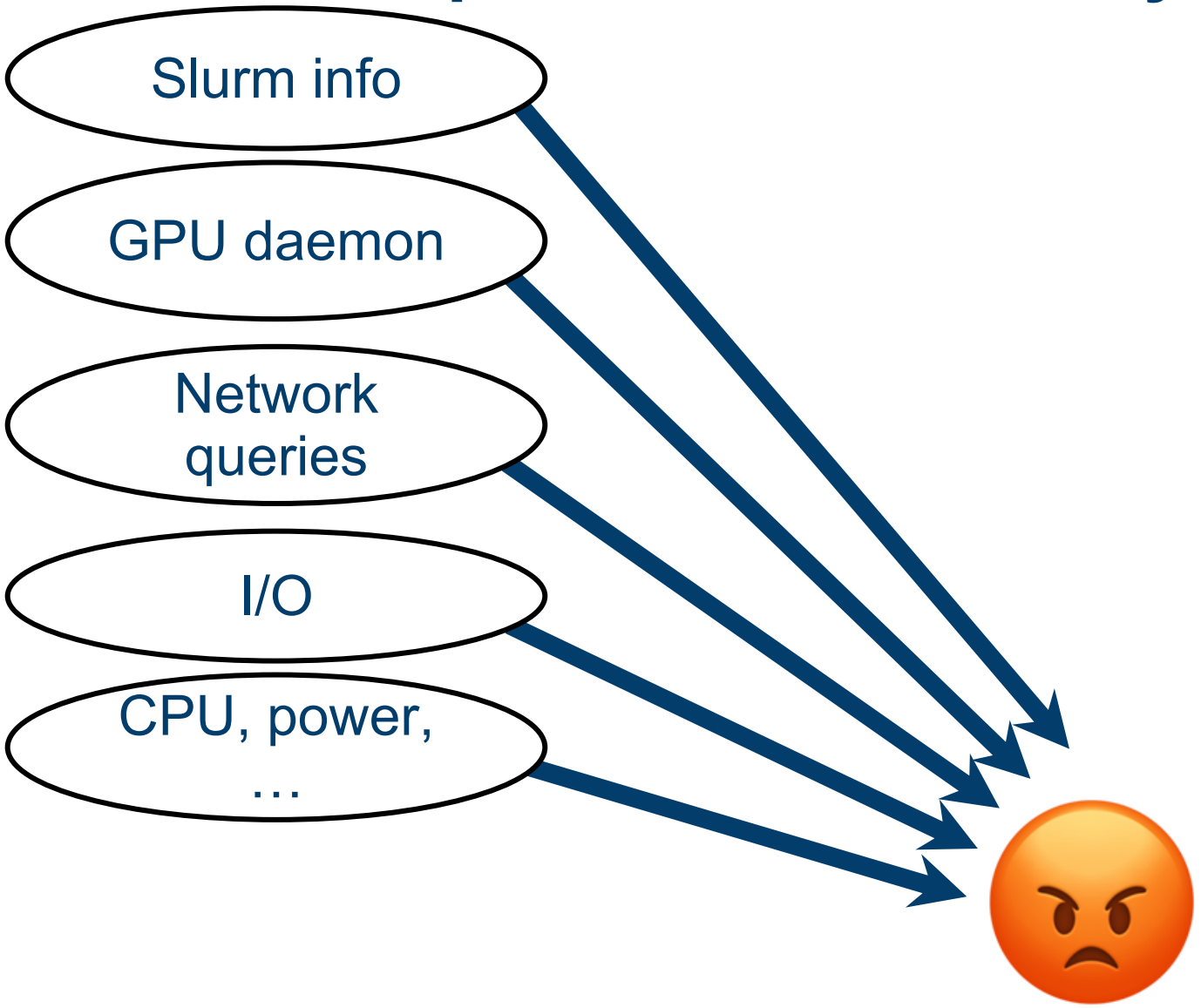
What is operational data analytics? Why do ~~we~~ users need it?



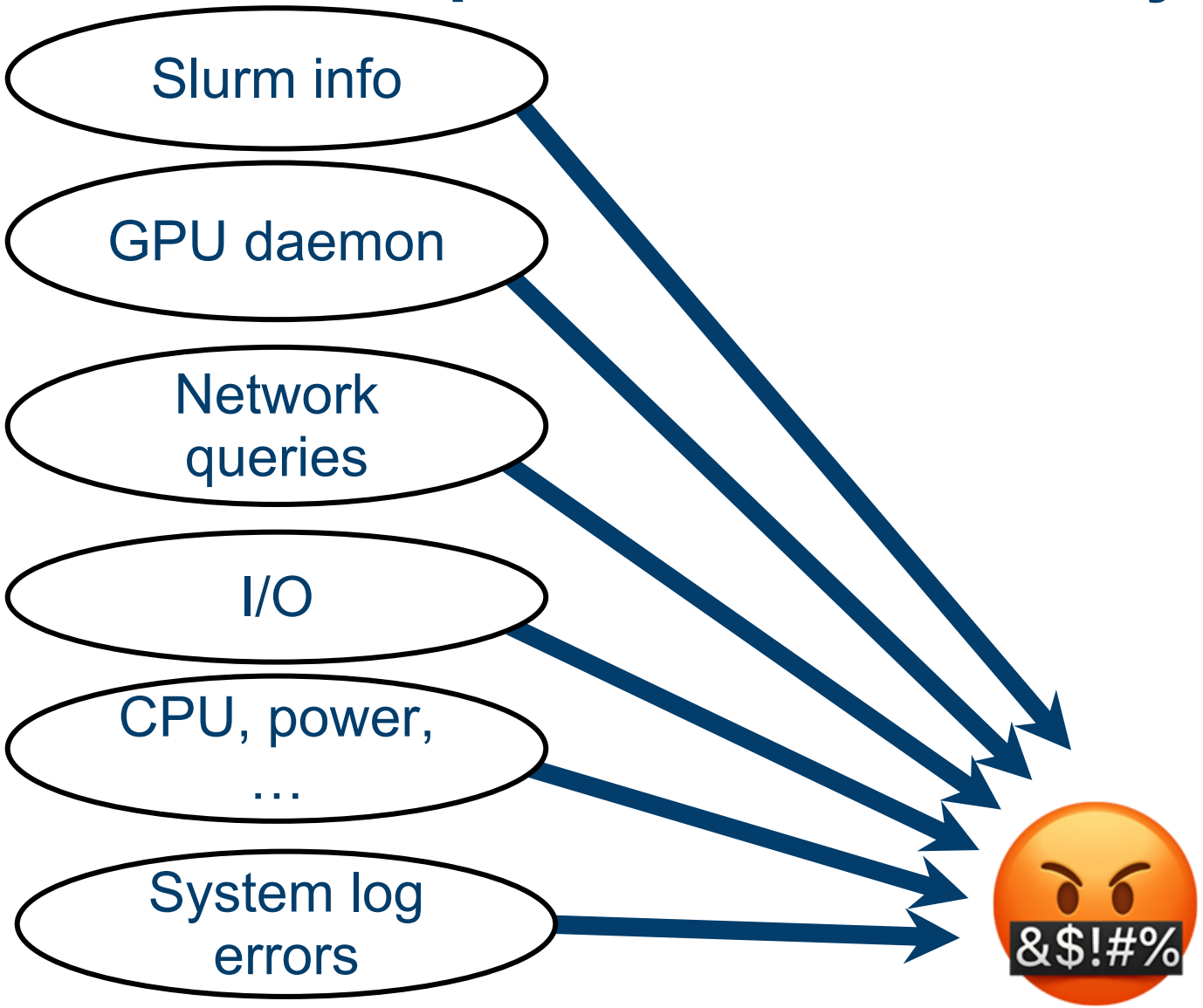
What is operational data analytics? Why do ~~we~~ users need it?



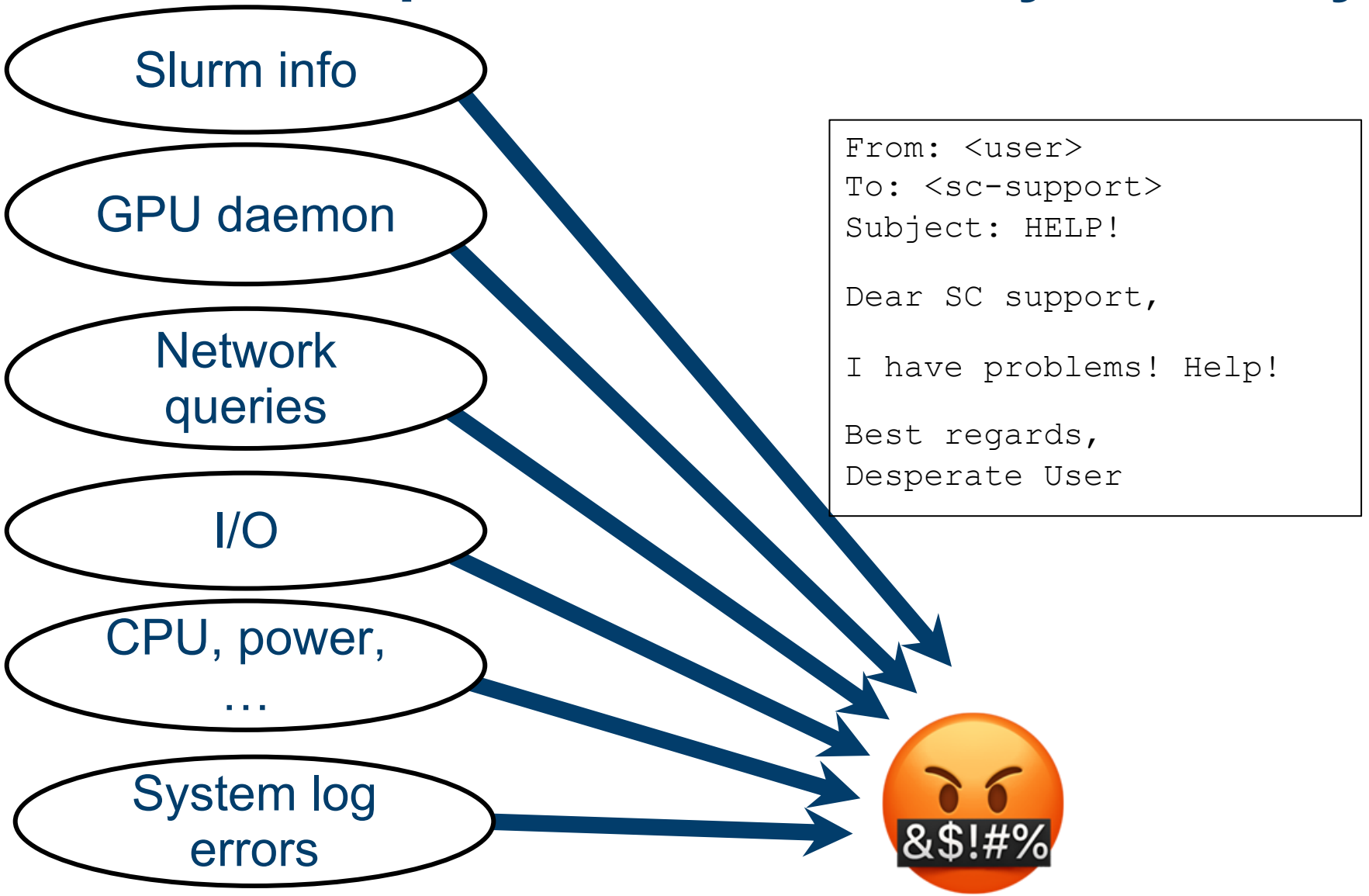
What is operational data analytics? Why do ~~we~~ users need it?



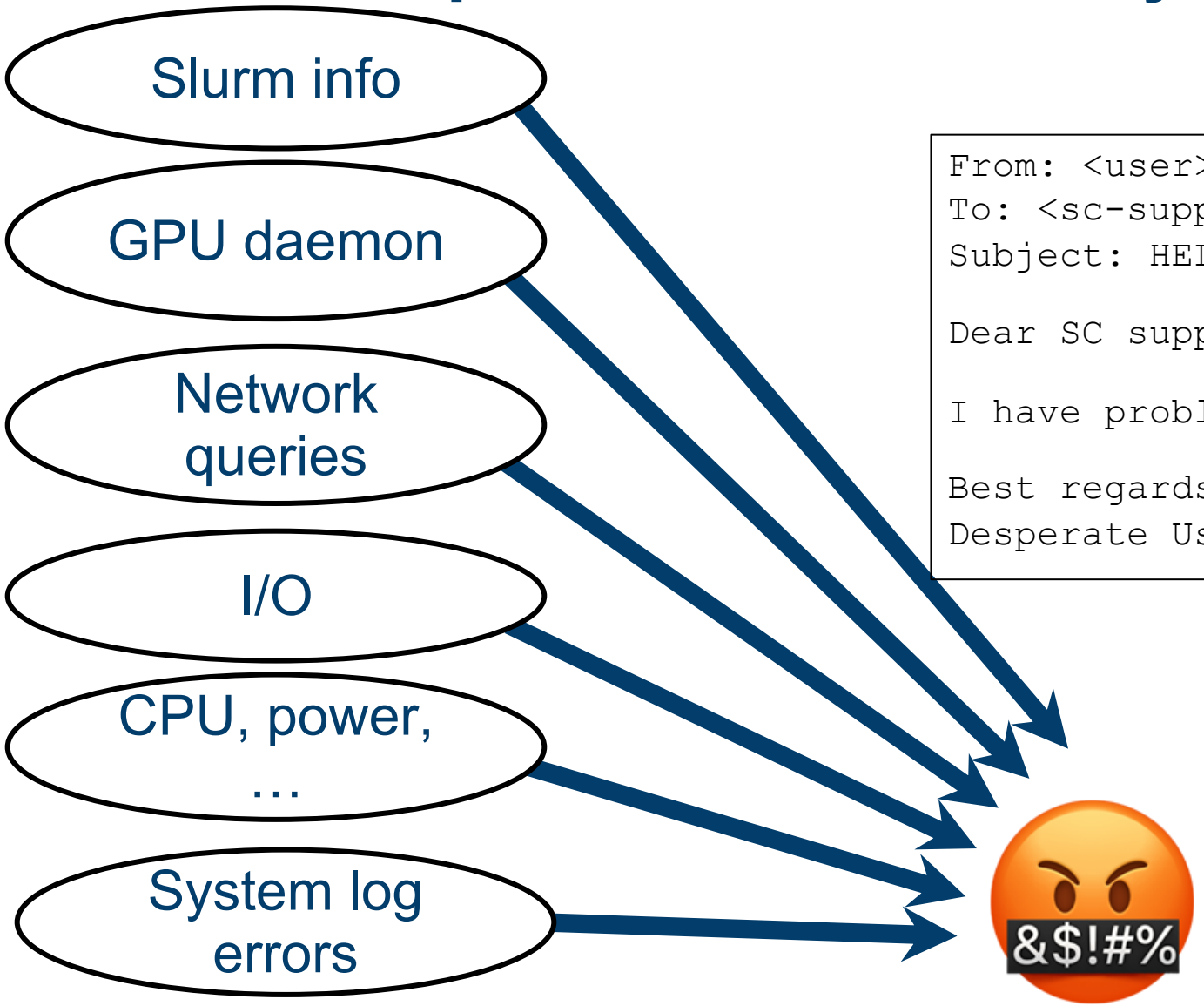
What is operational data analytics? Why do ~~we~~ users need it?



What is operational data analytics? Why do ~~we~~ users need it?



What is operational data analytics? Why do ~~we~~ users need it?




```
From: <user>
To: <sc-support>
Subject: HELP!

Dear SC support,

I have problems! Help!

Best regards,
Desperate User
```

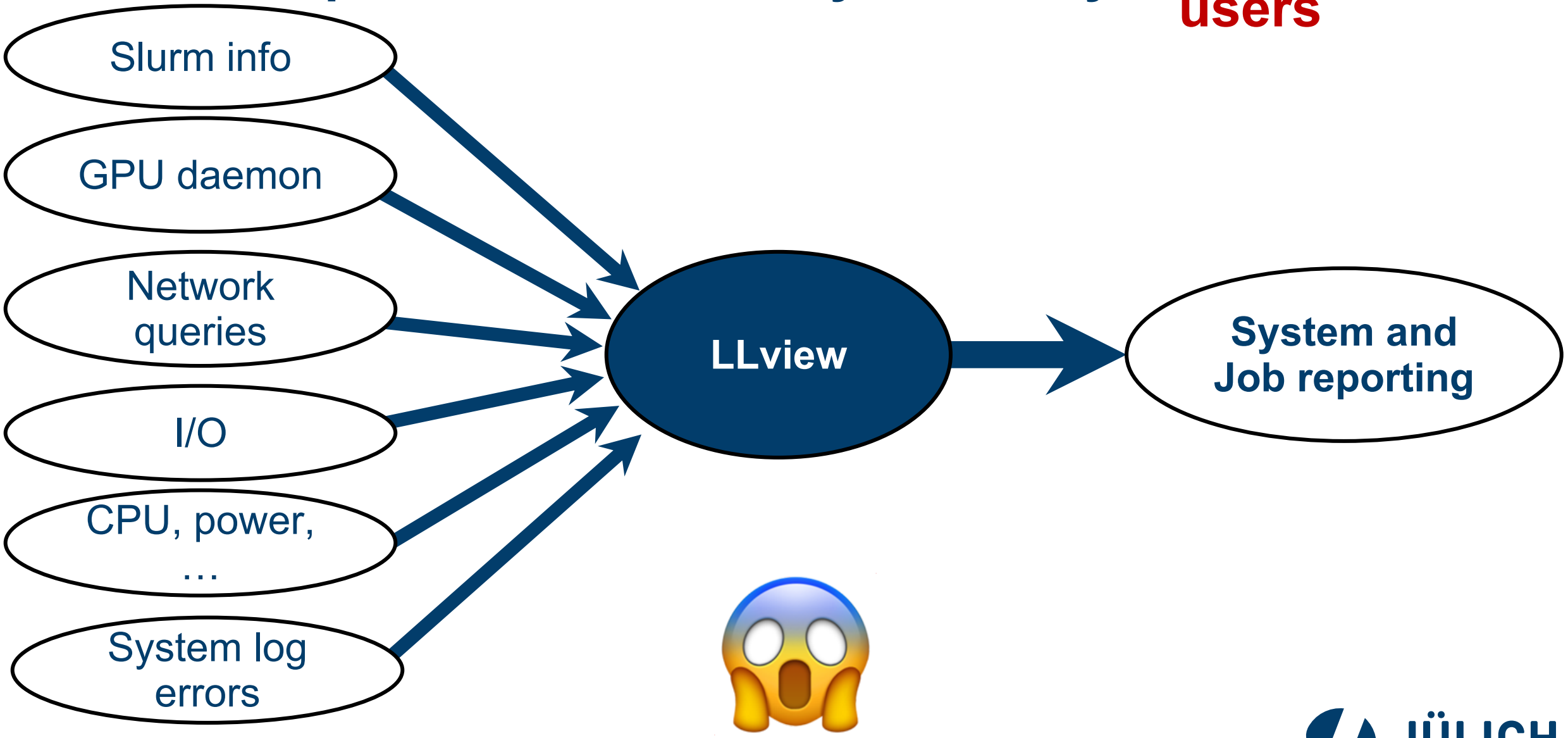
```
From: <sc-support>
To: <admins>
Subject: Fw: HELP!

Dear Admins,

We have problems! Help!

Best regards,
Desperate Support
```

What is operational data analytics? Why do ~~we~~ users need it?



What is operational data analytics? Why do ~~we~~ users need it?

Slurm info

GPU daemon

Network queries

I/O

CPU, power, ...

System log errors

Supporters

Admins

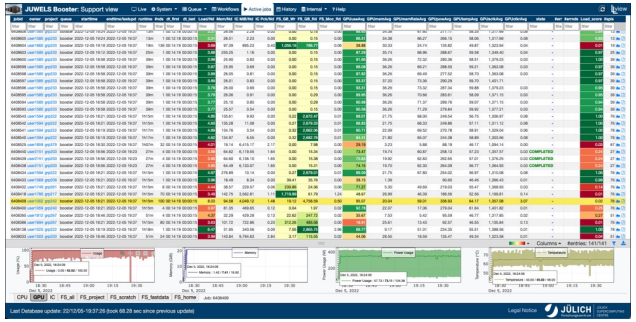
LLview

System and Job reporting

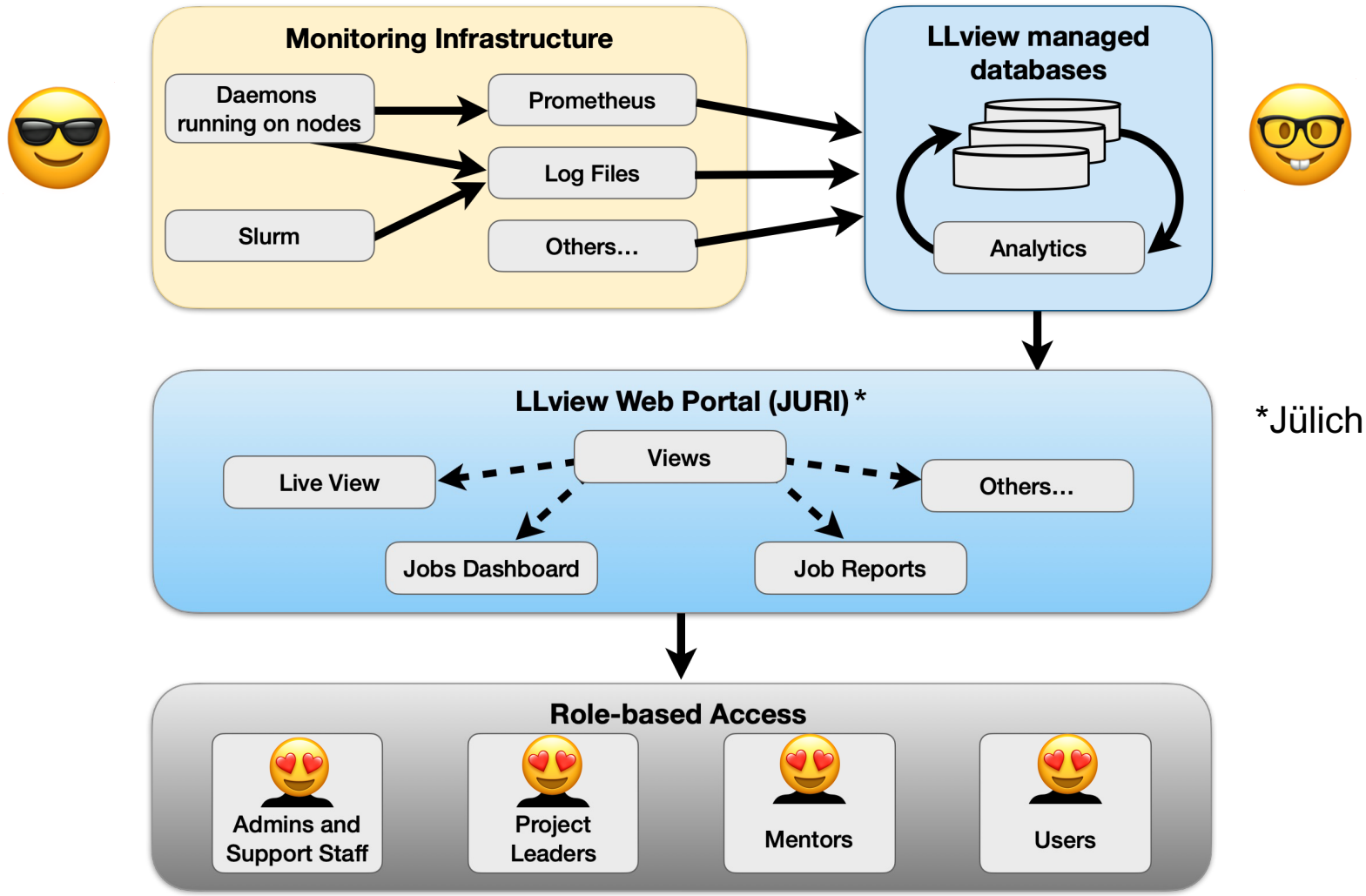
Updates every ~1min using existing sources

→ Negligible overhead!

→ Scalable!



Simplified Scheme of LLview



*Jülich Reporting Interface

Features of LLview

- Near-real time information (also in HTML and PDF reports)
- Role-based access (Users, PI/PAs, mentors, support)
- Modular/extendable design
- Separation between data processing and visualisation

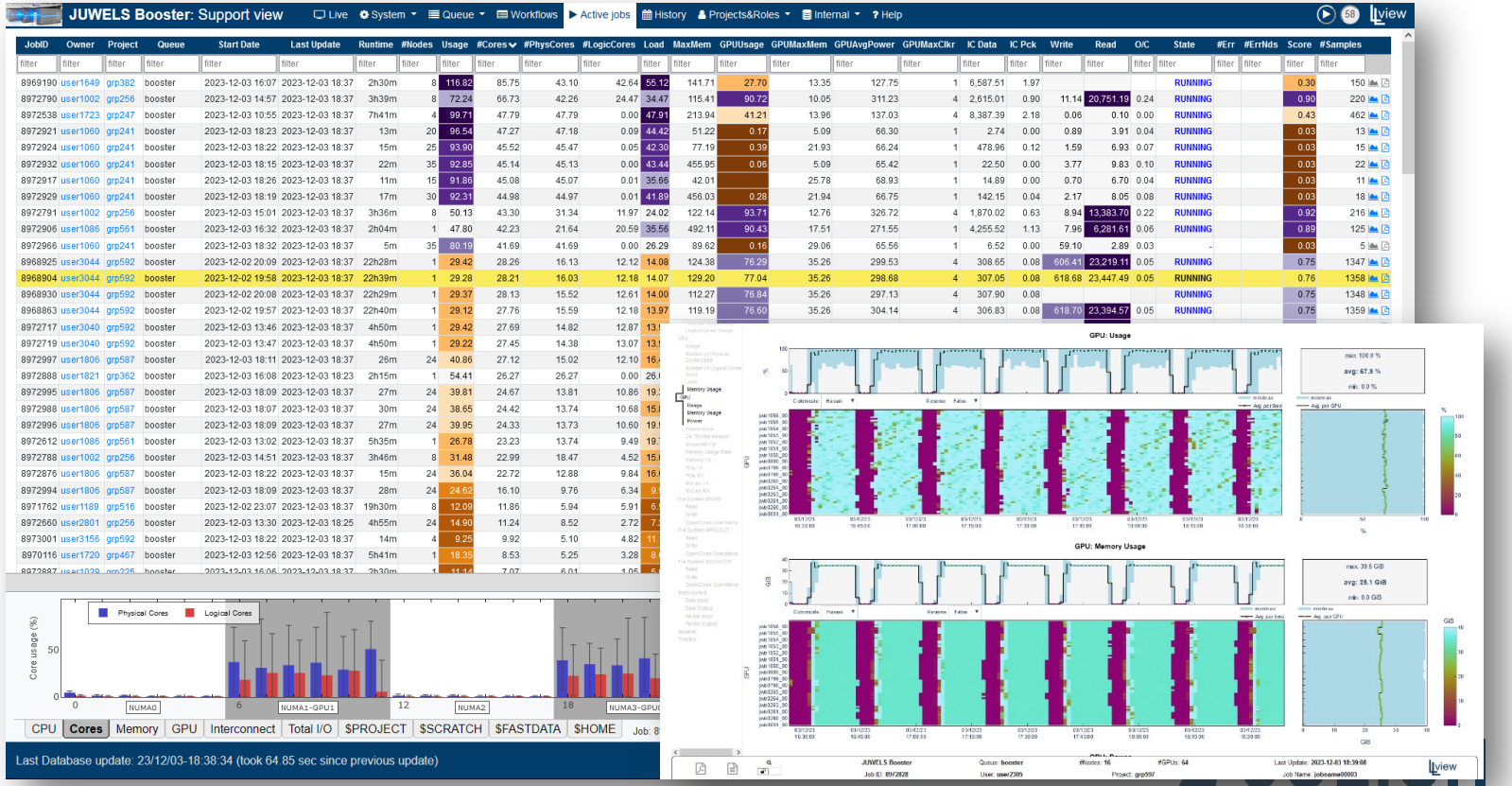
Current version: 2.3.2

Job information

- Basic job information
- Computing time spent/estimation
- Job Status
- CPU: Load, Core, Memory usage
- GPU: Usage, Power, Temperature, …
- I/O statistics
- Network traffic
- Node list (with interconnected cells)
- System error messages
- Timeline

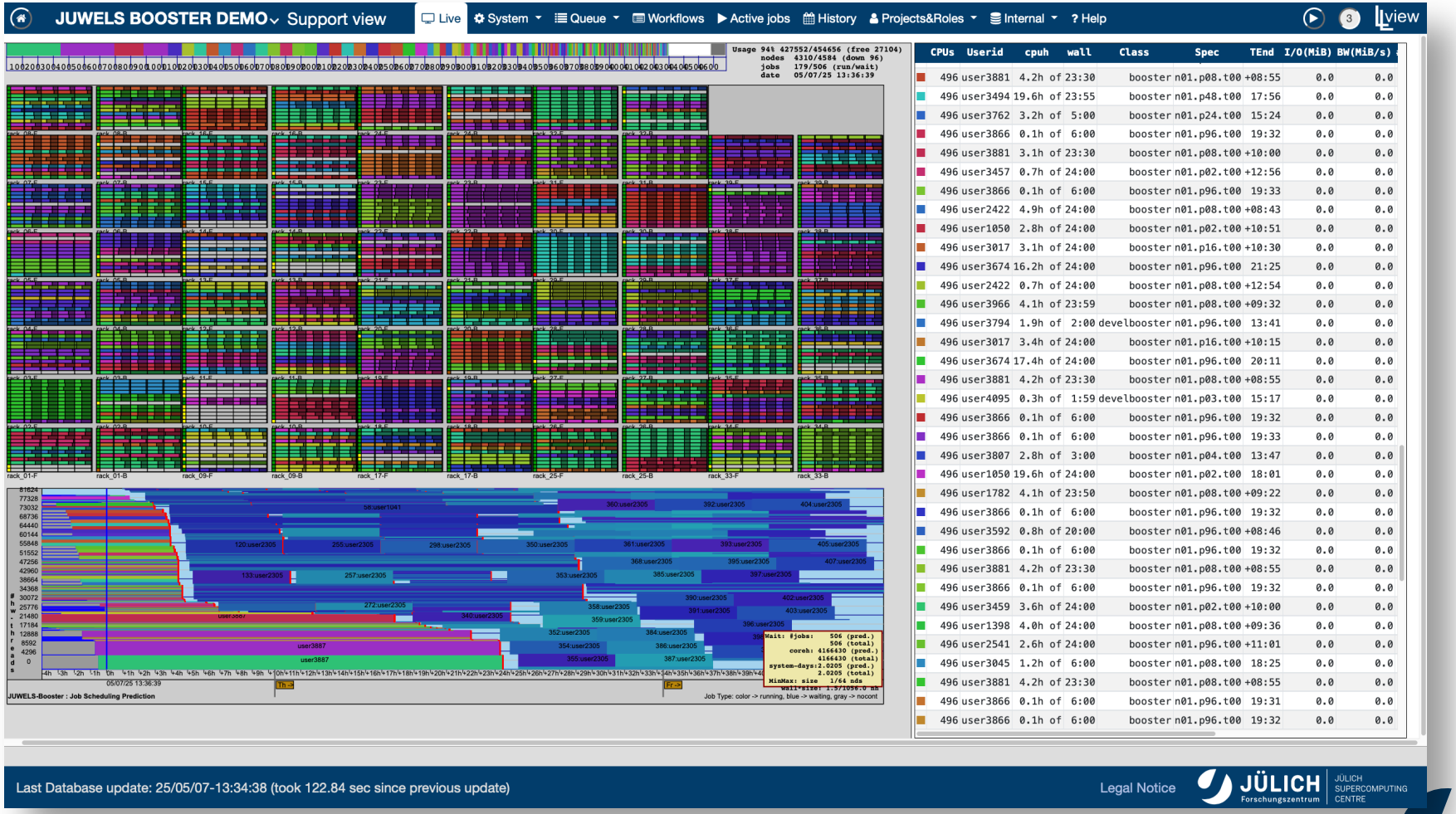
System information

- System usage overview
- Node usage, Errors



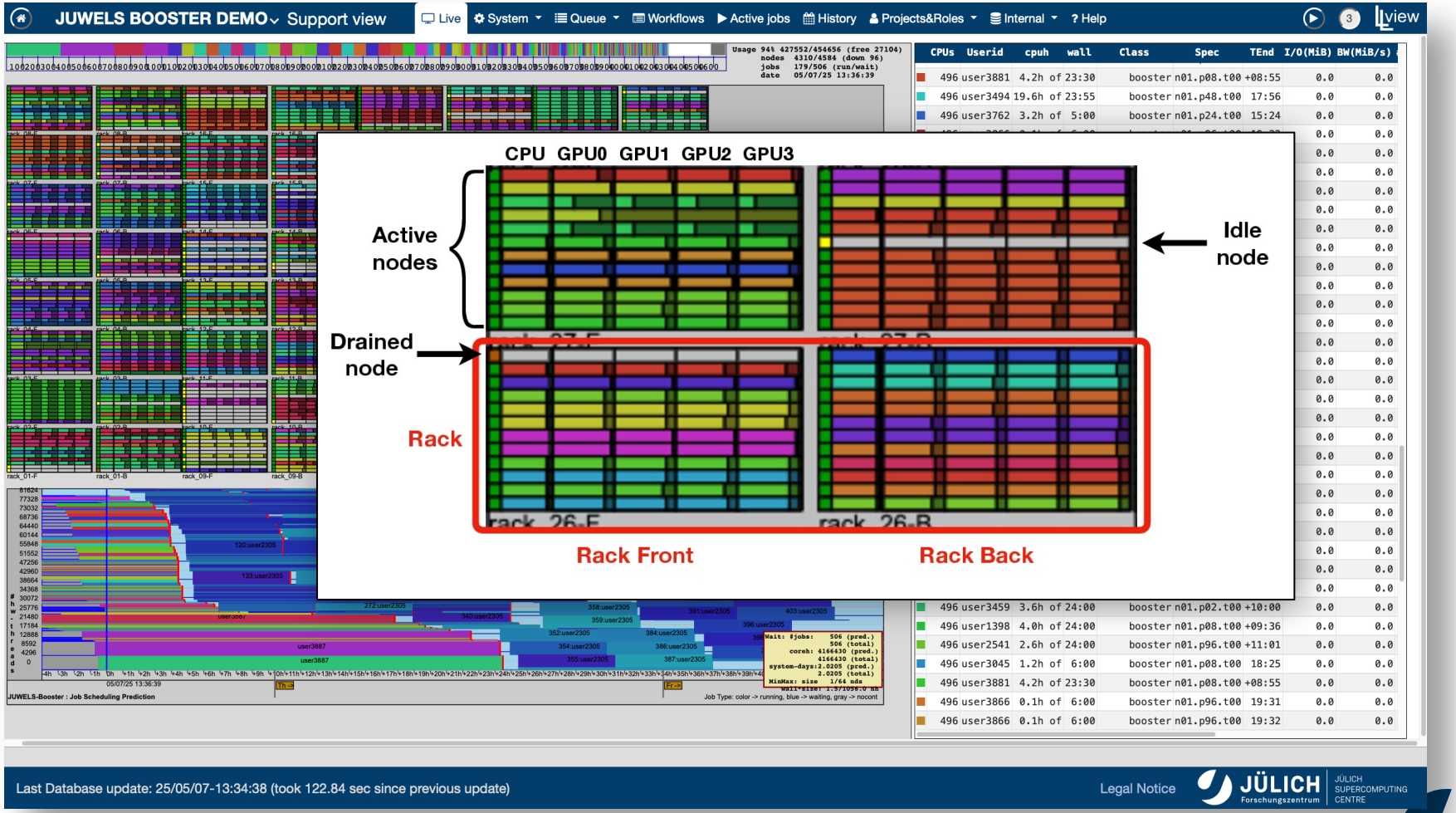
The Three Views and Their Analytical Use Cases

Live View



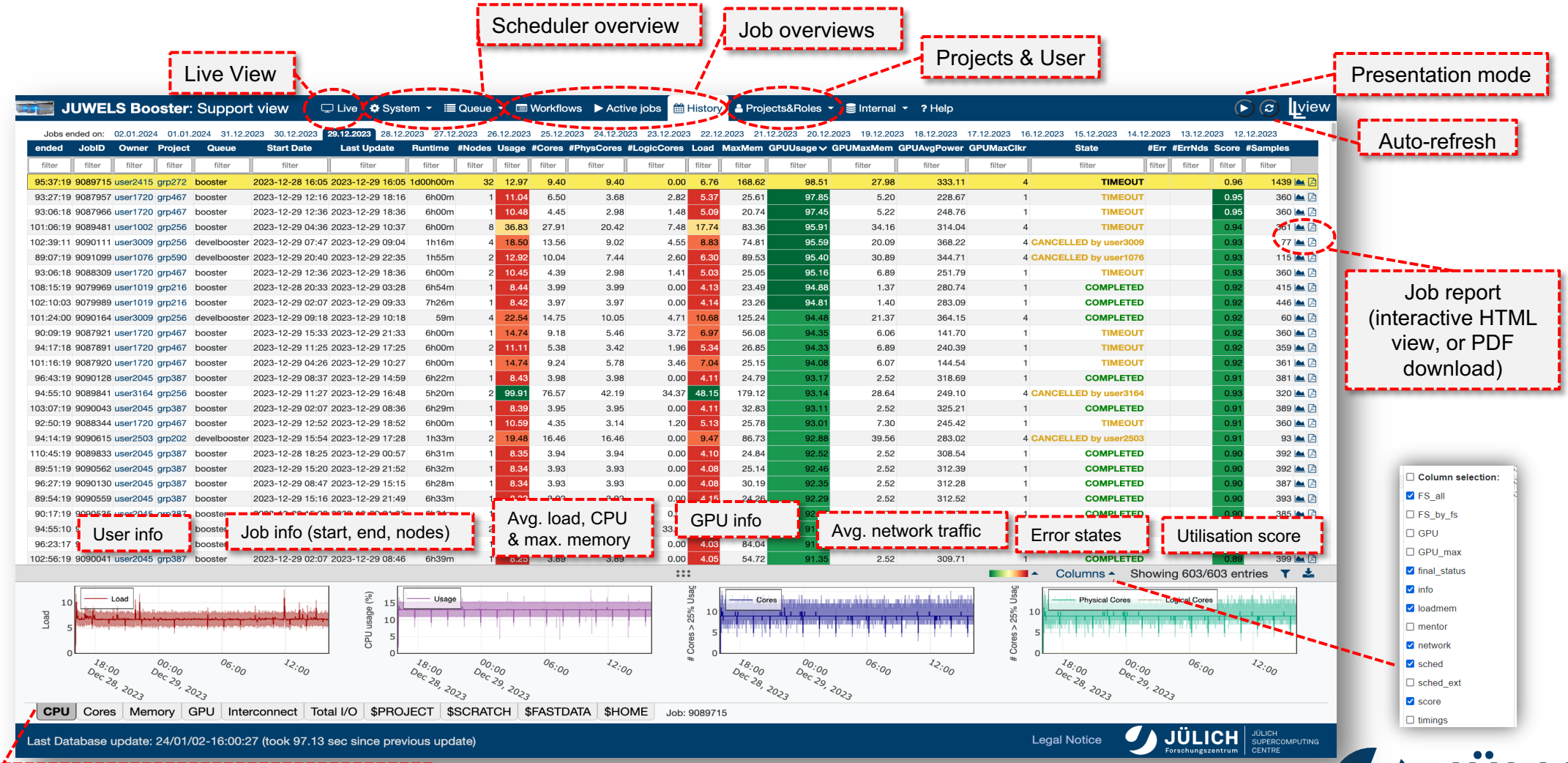
The Three Views and Their Analytical Use Cases

Live View



The Three Views and Their Analytical Use Cases

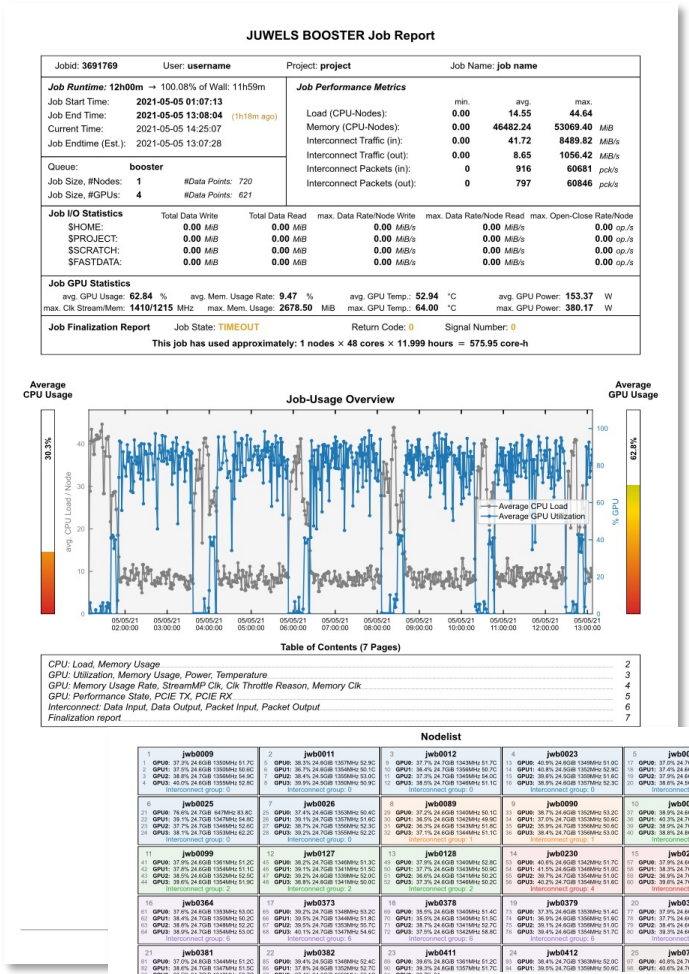
Jobs Dashboard

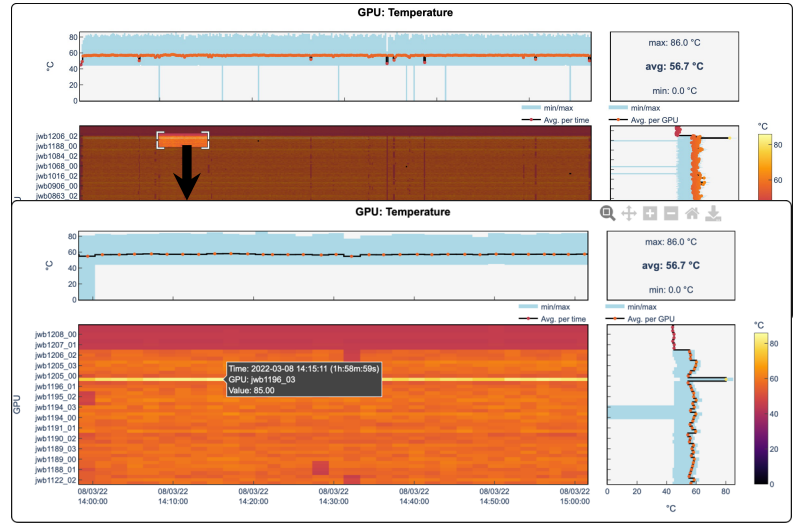


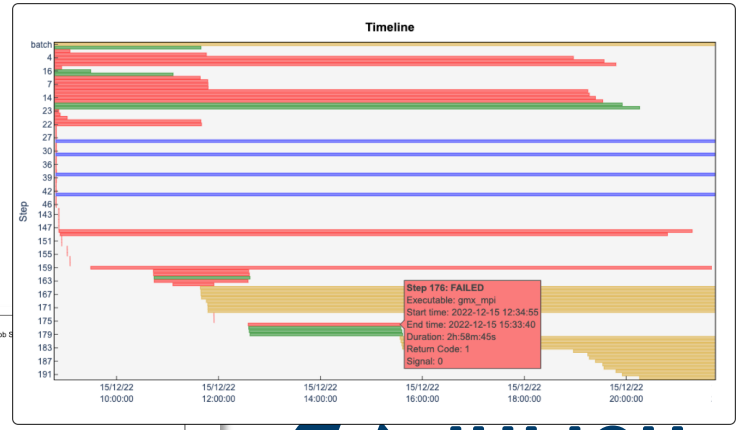
The Three Views and Their Analytical Use Cases

Jobs Reports

- Basic job information
- General average usage
- Computing time spent/estimation
- Status
- CPU, Core
  - Usage, Load, Memory usage
- GPU
  - Usage, Power, Temperature, …
- I/O
- Network traffic
- Nodelist (with interconnected cells)
- Error messages
- Timeline

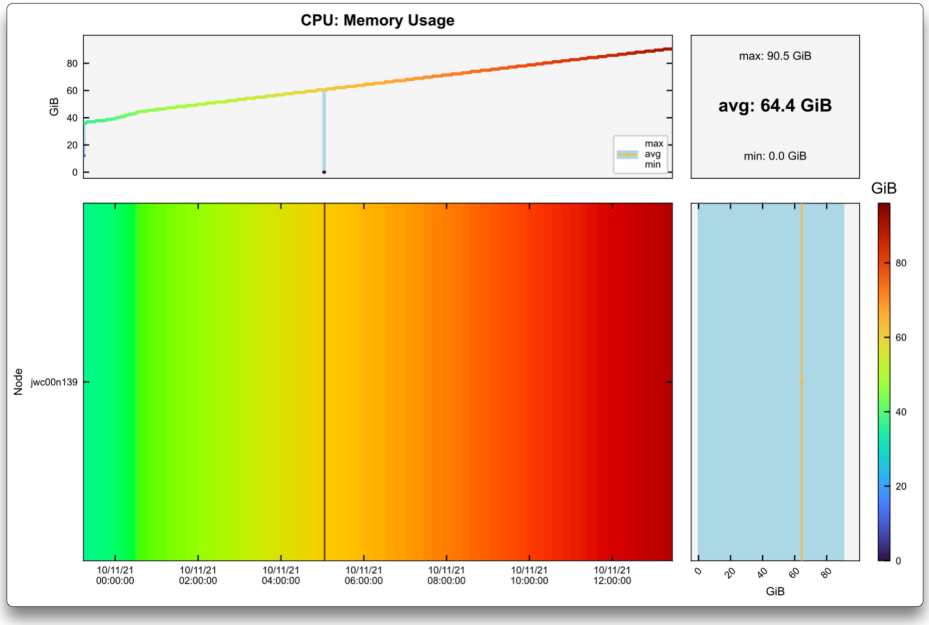


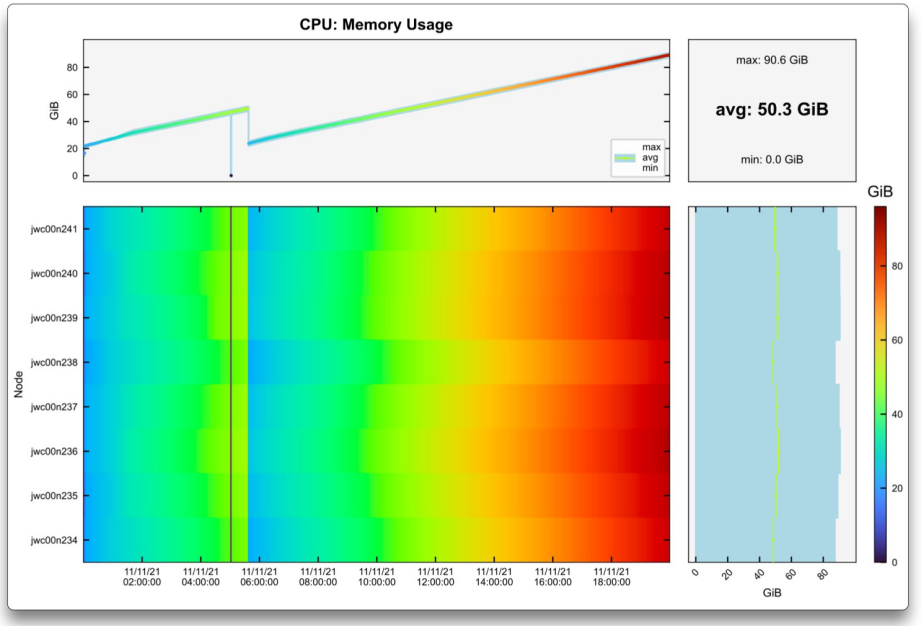


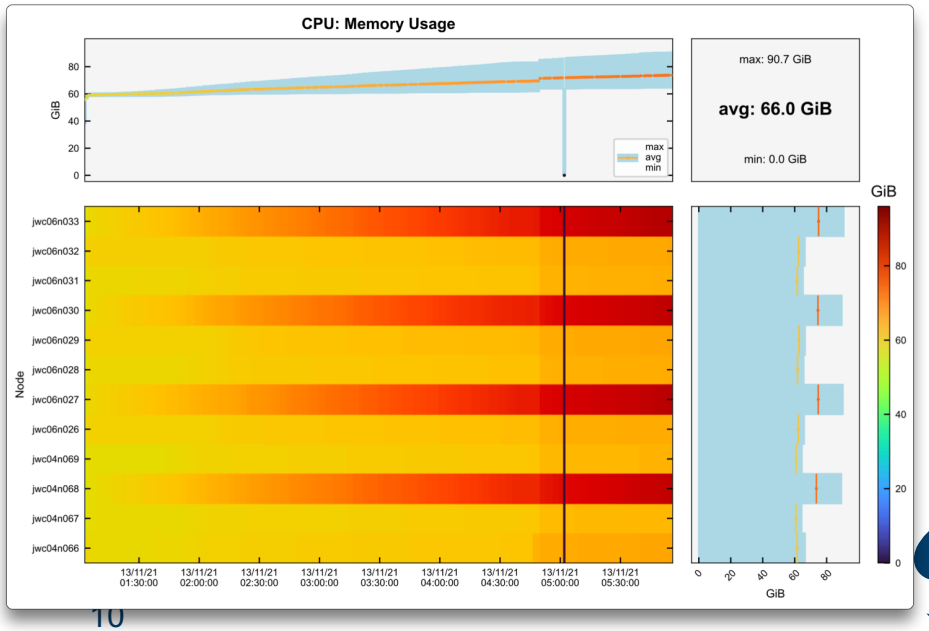


Examples

Memory leak







Examples

System Error Logs

Job Finalization Report	Job State: **FAILED**	Return Code: **9**	Signal Number: **0**
System Error Report	# Msgs: **2** # Nodes: **1**	**(Out-of-memory)**	

This job has used approximately: 1 nodes × 48 cores × 0.186 hours = 8.93 core-h

System Error Report

Msgs: **2** # Nodes: **1**

Error Messages:

```
2025-04-24T19:21:29+0200 jwb0163.juwels kernel: python invoked oom-killer:
gfp_mask=0x140dca(GFP_HIGHUSER_MOVABLE!__GFP_COMP!__GFP_ZERO), order=0, oom_score_adj=0
2025-04-24T19:21:29+0200 jwb0163.juwels kernel: Out of memory: Killed process 1556003 (python) total-
vm:902572524kB, anon-rss:496358356kB, file-rss:89856kB, shmem-rss:19200kB, UID:xxxxx pgtables:973944kB
oom_score_adj:0
```

Examples

System Error Logs

Job Finalization Report	Job State: **TIMEOUT**	Return Code: **0**	Signal Number: **0**
System Error Report	# Msgs: **2** # Nodes: **1**	**(Flipping Link)**	

This job has used approximately: 16 nodes × 48 cores × 11.973 hours = 9195.26 core-h

System Error Report

Msgs: **2** # Nodes: **1**

Error Messages:

```
2025-06-02T03:24:56+0200 jwb0155.juwels pshealthcheck[184436]: [EE] ERROR: mlx5_2 local port (LID
4667, port 1): LinkDownedCounter = 8 (threshold = 1)
2025-06-02T03:24:56+0200 jwb0155.juwels pshealthcheck[184436]: [EE] ERROR: mlx5_2 remote port (LID
297, port 34): LinkDownedCounter = 8 (threshold = 1)
```

Examples

System Error Logs

Job Finalization Report	Job State: **FAILED**		Return Code: **0**	Signal Number: **9**
System Error Report	# Msgs: **59**	# Nodes: **59**	**(Node Error)**	

This job has used approximately: 60 nodes × 48 cores × 11.328 hours = 32624.64 core-h

System Error Report

Msgs: **59** # Nodes: **59**

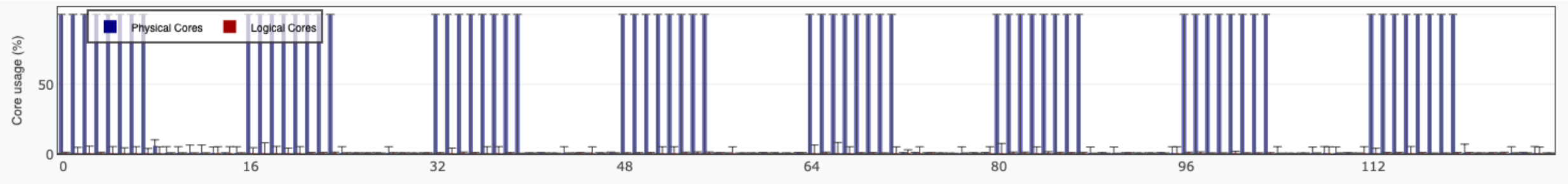
Error Messages:

```
2025-06-03T10:50:14+0200 jwb0011.juwels psid[63607]: psslurm: nodeDownAlloc: node 3764 in allocation
11584692 state A_RUNNING is down
2025-06-03T10:50:15+0200 jwb0025.juwels psid[80835]: psslurm: nodeDownAlloc: node 3764 in allocation
11584692 state A_EPILOGUE is down
2025-06-03T10:50:15+0200 jwb0036.juwels psid[62353]: psslurm: nodeDownAlloc: node 3764 in allocation
11584692 state A_EPILOGUE is down
2025-06-03T10:50:15+0200 jwb0041.juwels psid[64695]: psslurm: nodeDownAlloc: node 3764 in allocation
11584692 state A_EPILOGUE is down
2025-06-03T10:50:15+0200 jwb0054.juwels psid[62378]: psslurm: nodeDownAlloc: node 3764 in allocation
11584692 state A_EPILOGUE is down
```

Examples

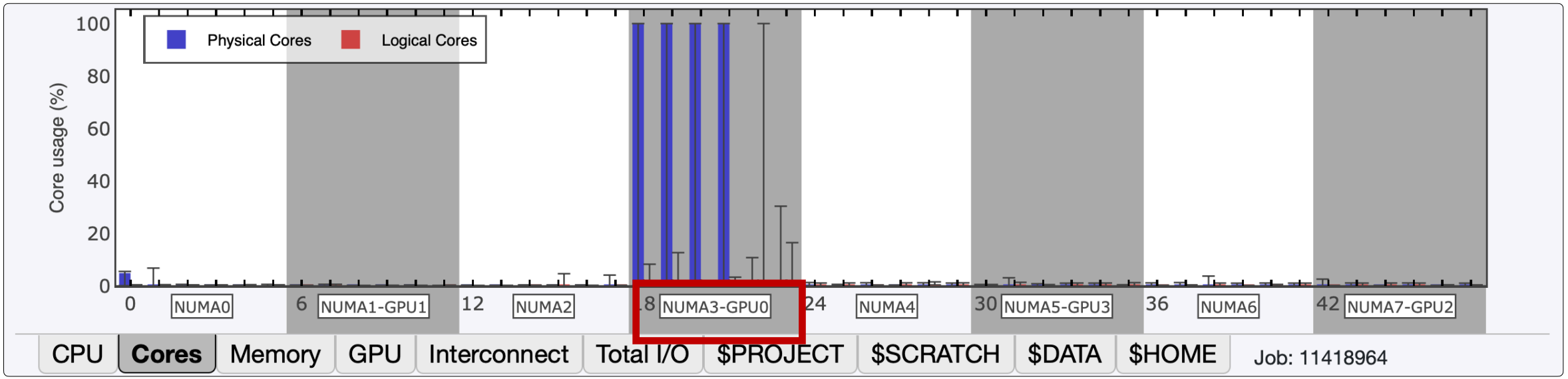
Inadequate Resource Usage

Usage	Cores	#PhysCores	#LogicCores	Load
<50	filter	filter	filter	filter
2.90	0.01	0.01	0.0	512.26
1.37	0.37	0.27	0.1	477.35
4.07	0.01	0.01	0.0	395.68
0.68	0.00	0.00	0.00	158.31
3.00	3.73	3.73	0.00	149.31
18.90	23.98	23.98	0.00	128.64
10.49	0.00	0.00	0.00	128.32
18.91	24.00	24.00	0.00	127.77
18.86	23.94	23.94	0.00	127.68
18.88	23.98	23.98	0.00	127.68
30.34	78.75	78.75	0.00	124.96



Examples

Inadequate Resource Usage

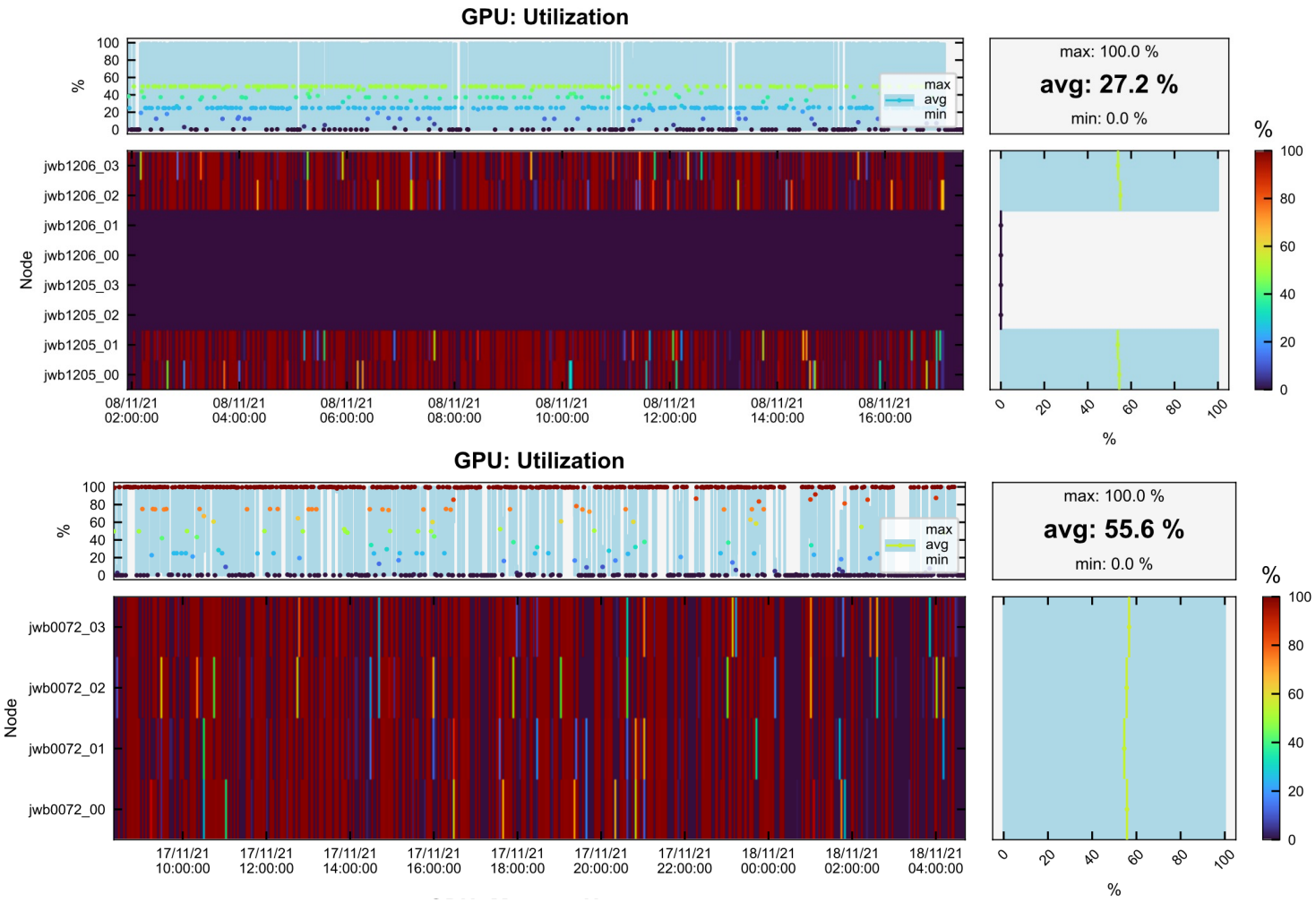


Examples

Inadequate Resource Usage

2 nodes (8 GPUs) using only 2 GPUs each

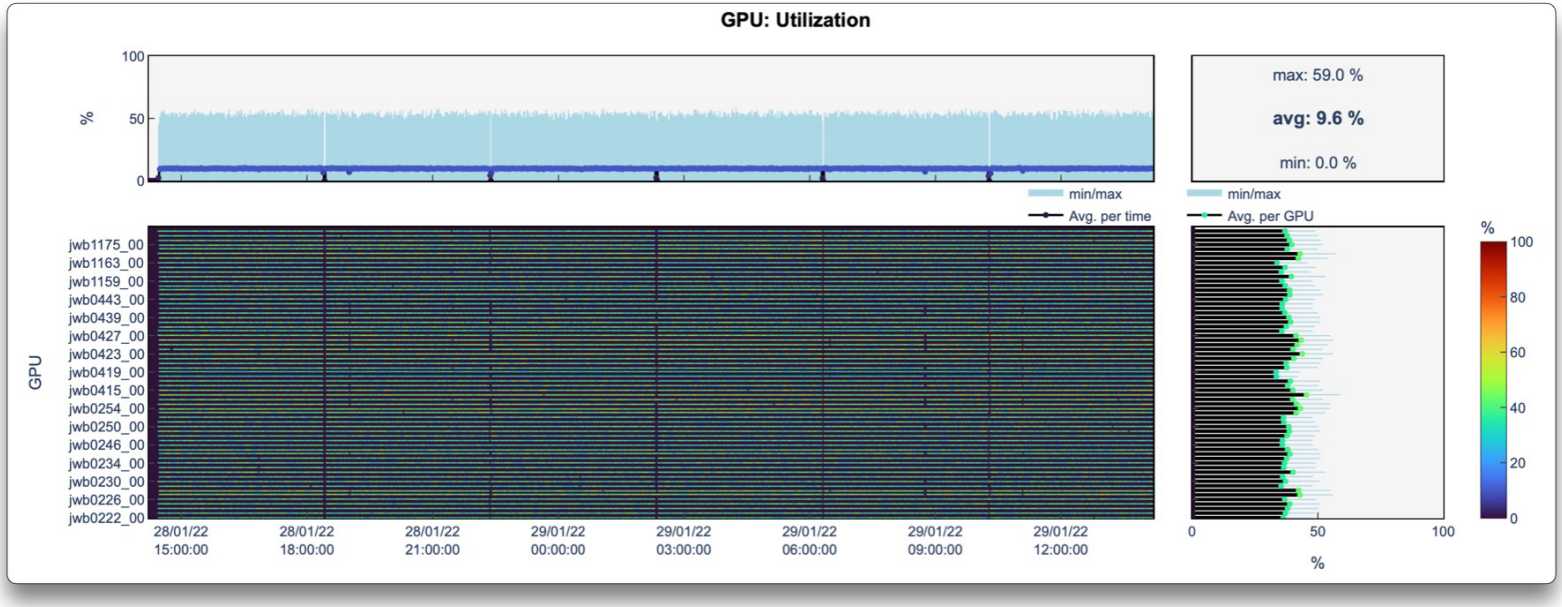
1 nodes using all 4 GPUs



Examples

Inadequate Resource Usage

1 GPU per node

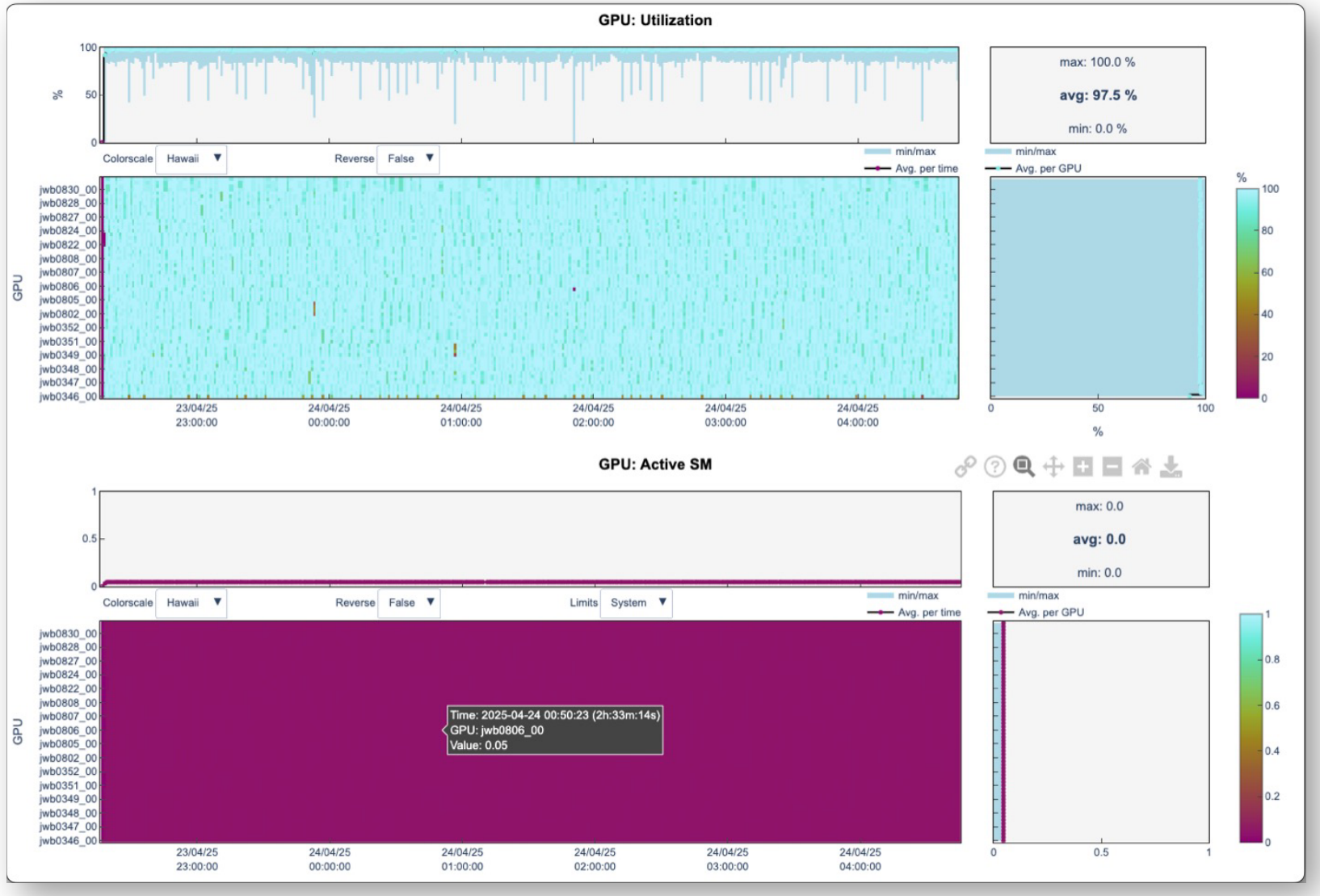


Examples

Inadequate Resource Usage

GPU Utilisation ~100%

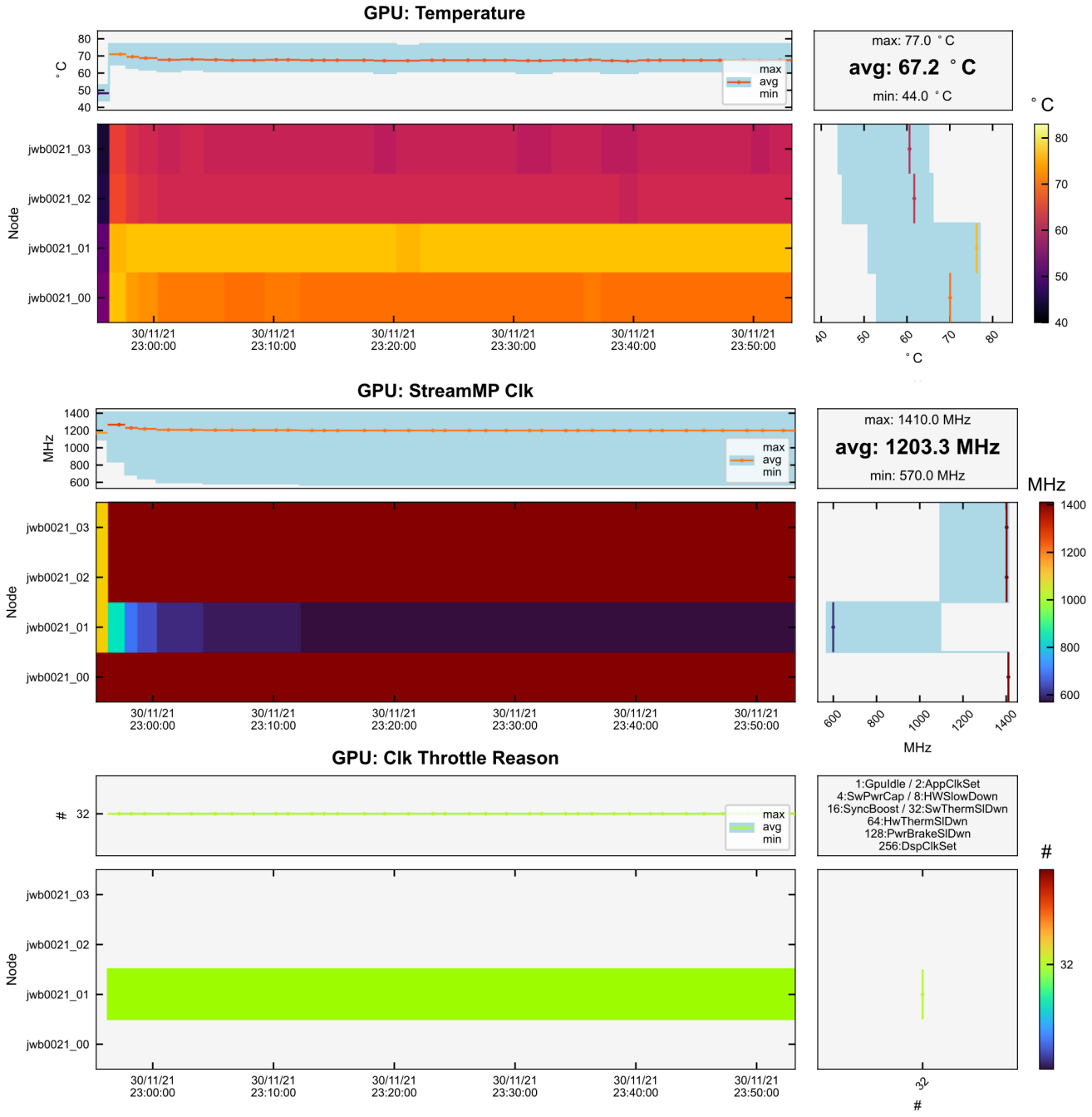
GPU Active SM ~5%



Examples

Hardware Issues

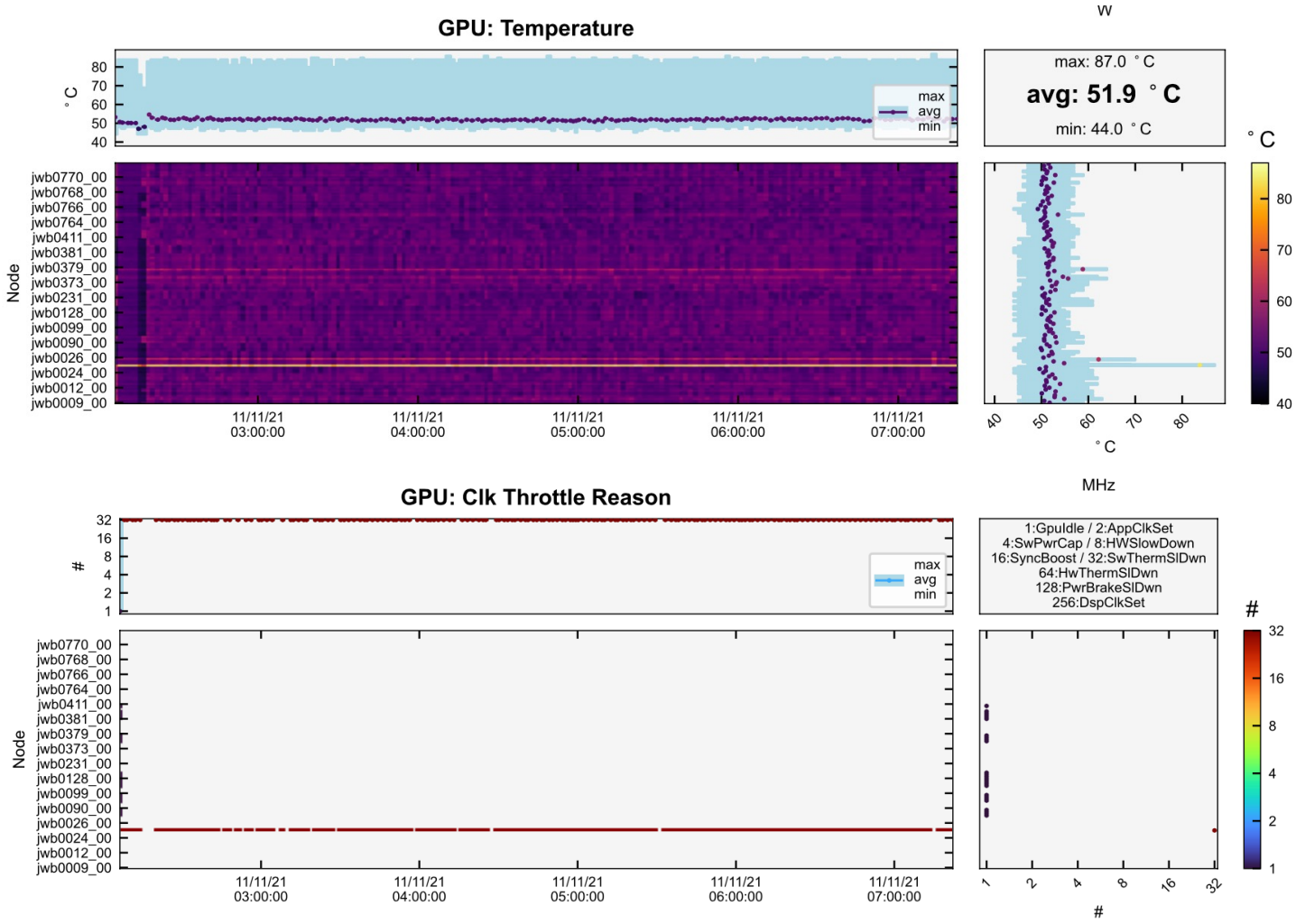
4 GPUs:



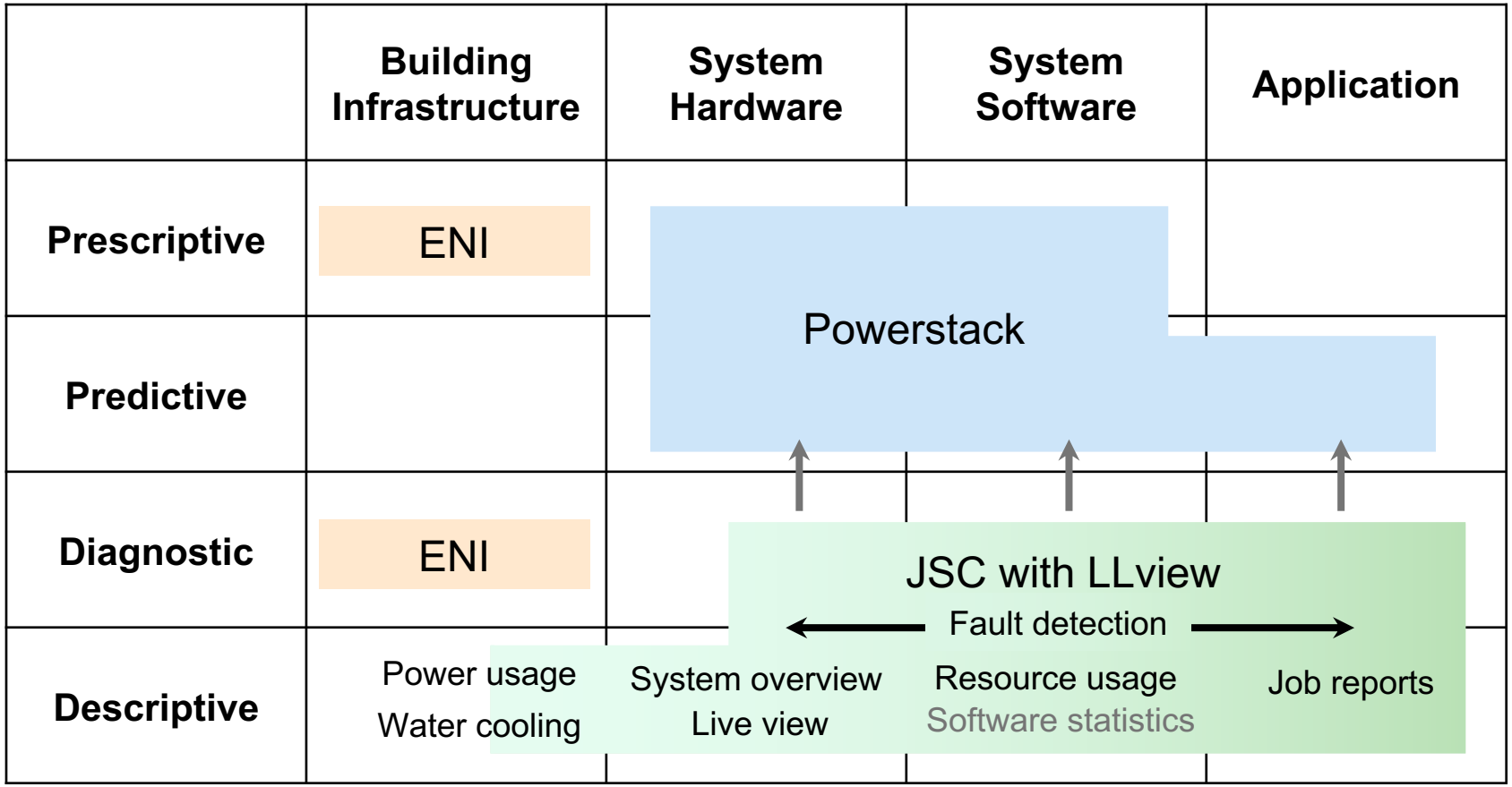
Examples

Hardware Issues

128 GPUs:

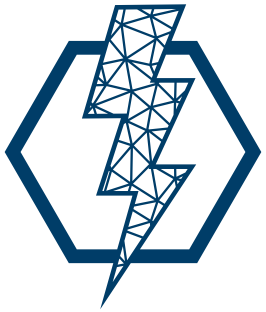


Classifying LLview



A. Netti et al., A Conceptual Framework for HPC Operational Data Analytics, 2021 IEEE International Conference on Cluster Computing (CLUSTER), 596 (2021)

Outlook



JUPITER

eupex.eu

SEANERGYS

(Starting in June 2025)

- Faster updates (for JUPITER)
- Generalisations, new metrics and data sources
- Queue view and analysis
- REST-API
- Long-term analysis
- Workflow and project reports
- …

Conclusions

Thank you!

LLview…

- …helps to identify (user and system) software and hardware issues
- …enables users find problems on their own
- …brings support to a higher level
- …reduces the need for communication between users and support as well as support and admins
- …improves the overall system utilisation

- Open source GPLv3 Licence github.com/FZJ-JSC/llview
- Website: llview.fz-juelich.de
- Contact: llview.jsc@fz-juelich.de

Presentation DOI: 10.5281/zenodo.15681176