



Energie-Management für einen HPC-Cluster mit Slurm

Using Slurm`s Energy Saving capabilities to improve energy efficiency for a small HPC cluster

Einführung

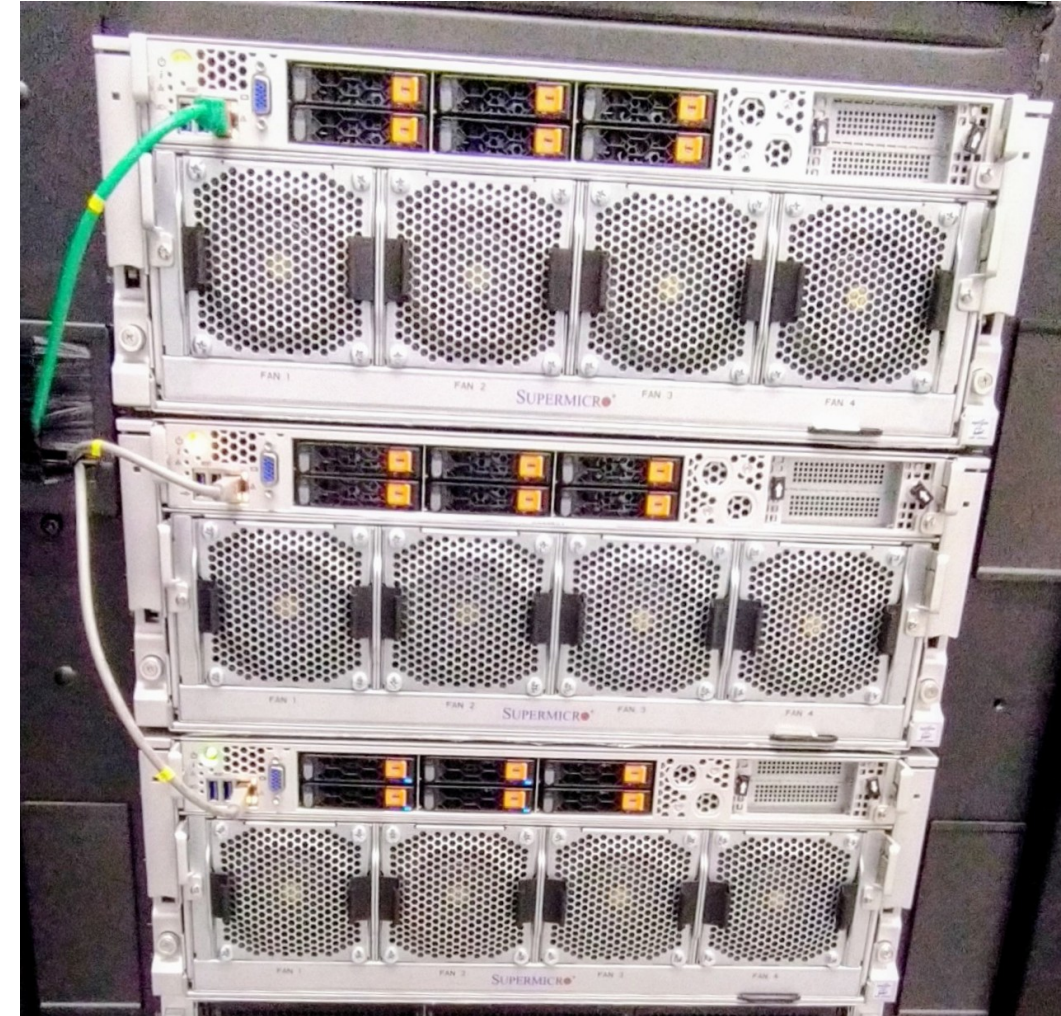
Gliederung

- Einführung
- Warum?
- Was?
- Wie?
- Wann?
- Wirkungen
- Nebenwirkungen

Einführung

HPC22-Cluster

- 48 Rechenknoten
 - work, big, gpu, interactive
- ein Login-Knoten, ein Management-System
- Stromversorgung - 1+1 redundant
- Netzteile
 - 650 W (work, interactive)
 - 1920 W (big)
 - 5000 W (gpu)





Einführung

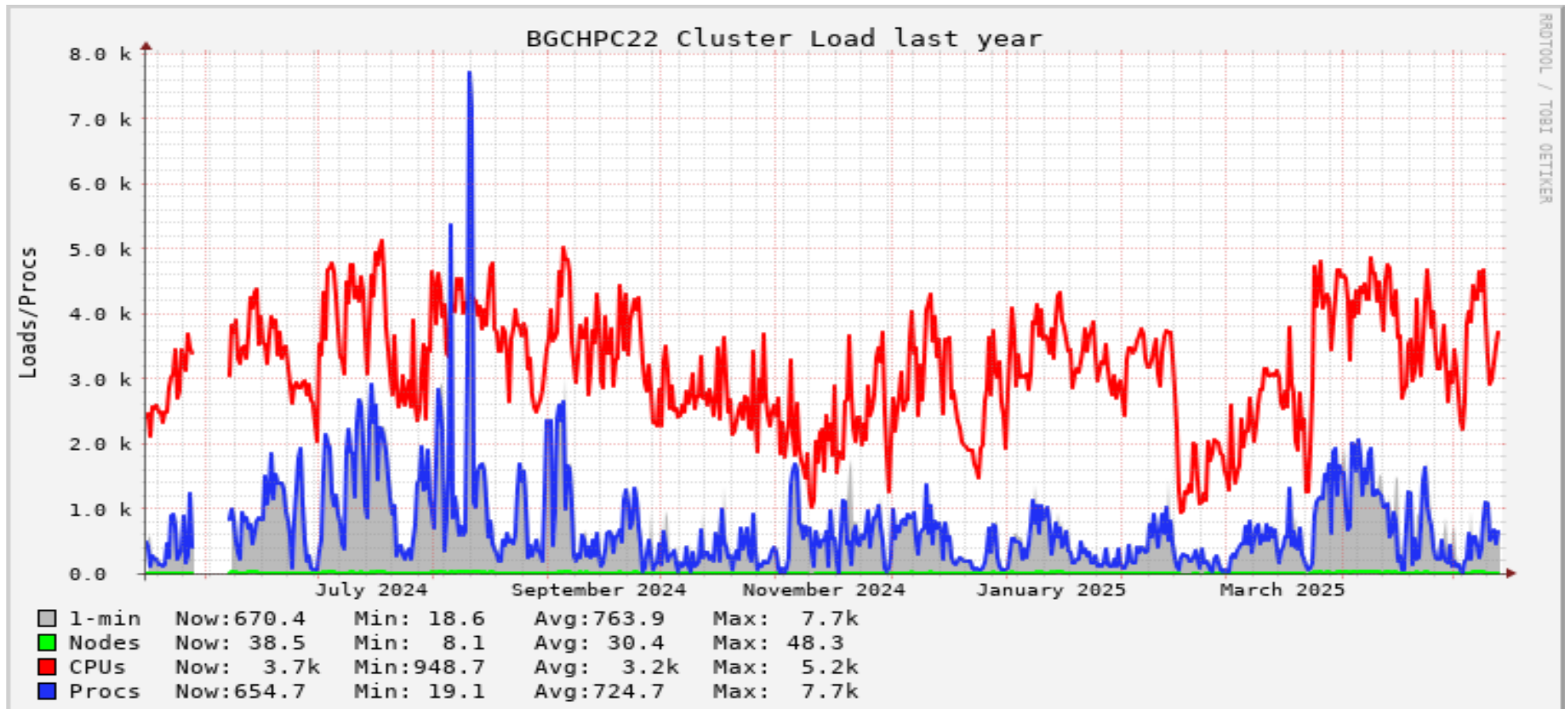
Ausstattung

- CPU
- RAM
- Auslastung

node-r5-he13	128	128	100.0	64.71	0	0	2012	1216	ALLOCATED
node-r5-he15	128	128	100.0	128.28	0	0	2012	1600	ALLOCATED+RESERVED
node-r5-he17	64	64	100.0	24.17	0	0	501	390	ALLOCATED
node-r5-he18	64	63	98.4	20.36	0	0	501	491	MIXED
node-r5-he19	64	64	100.0	25.28	0	0	501	390	ALLOCATED
node-r5-he20	64	64	100.0	17.18	0	0	501	390	ALLOCATED
node-r5-he21	64	64	100.0	24.71	0	0	501	390	ALLOCATED
node-r5-he22	64	50	78.1	50.06	0	0	501	500	MIXED
node-r5-he23	64	0	0.0	0.00	0	0	501	0	IDLE+POWERED_DOWN
node-r5-he24	64	0	0.0	0.01	0	0	501	0	IDLE+POWERED_DOWN
node-r5-he25	64	0	0.0	0.00	0	0	501	0	IDLE+POWERED_DOWN
node-r5-he26	64	0	0.0	0.00	0	0	501	0	IDLE+POWERED_DOWN
node-r5-he27	64	0	0.0	0.00	0	0	501	0	IDLE+POWERED_DOWN
node-r5-he28	64	0	0.0	0.00	0	0	501	0	IDLE+POWERED_DOWN
node-r5-he29	64	0	0.0	0.01	0	0	501	0	IDLE+POWERED_DOWN
node-r5-he30	64	0	0.0	0.00	0	0	501	0	IDLE+POWERED_DOWN
node-r5-he31	64	40	62.5	11.91	0	0	501	244	MIXED
node-r5-he32	64	60	93.8	6.00	0	0	501	480	MIXED
node-r5-he33	64	60	93.8	6.00	0	0	501	480	MIXED
node-r5-he34	64	62	96.9	11.57	0	0	501	142	MIXED
node-r5-he35	64	64	100.0	11.84	0	0	501	346	ALLOCATED
node-r5-he36	64	56	87.5	20.26	0	0	501	341	MIXED
node-r5-he37	64	60	93.8	6.00	0	0	501	480	MIXED
node-r5-he38	64	60	93.8	6.02	0	0	501	480	MIXED
node-r5-he39	64	58	90.6	29.96	0	0	499	453	MIXED
node-r6-he01	256	0	0.0	3.27	0	0	1005	0	IDLE
node-r6-he05	256	128	50.0	6.86	0	0	1005	100	MIXED

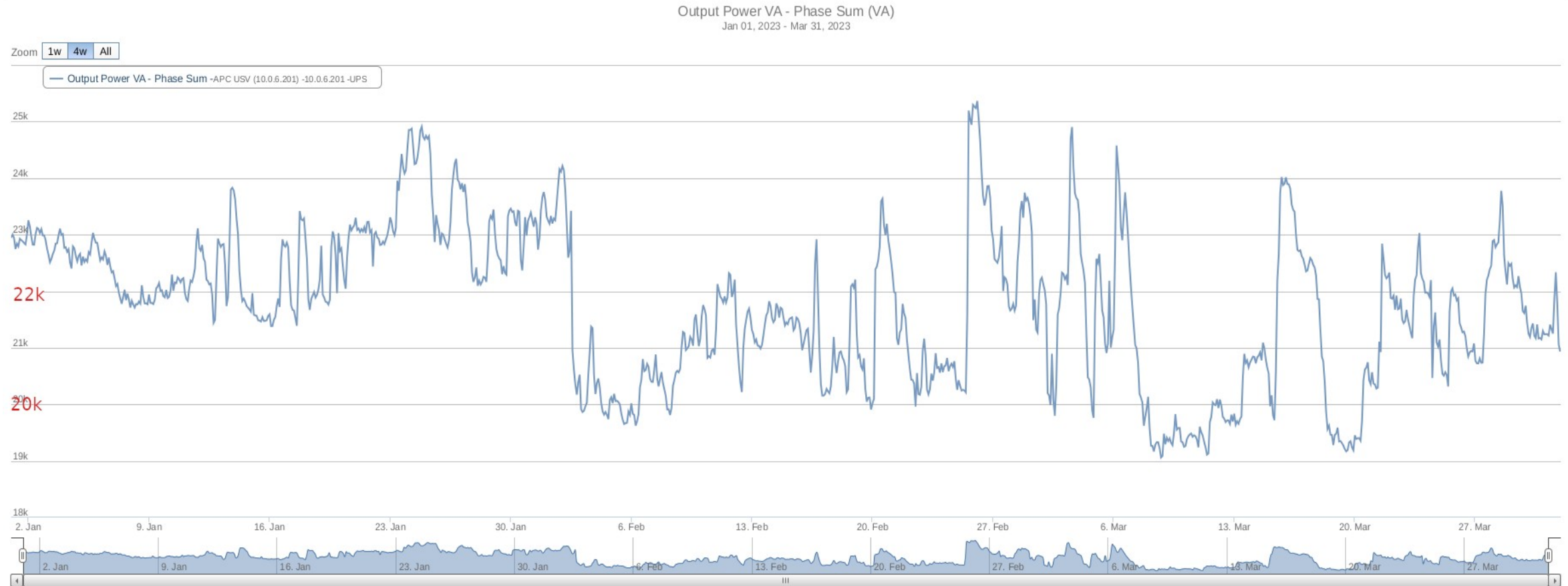
Warum?

Auslastung – schwankt!



Warum?

Hohe Grundlast (USV-Output, VA)



Was?

Verschwendung vermeiden!

- **Leerlaufende Rechenknoten abschalten**
- **???**

Wie?

Slurm

- **Simple Linux Utility for Resource Management**
- **Workload-Manager und Job-Scheduler**
- **Energiespar-Funktionen**
 - “Slurm provides an integrated mechanism for nodes being suspended (powered down, placed into power saving mode) and resumed (powered up, restored from power saving mode) on demand or by request.”[1]

Wie?

Slurm-Konfiguration - /etc/slurm/slurm.conf

```
SuspendTime=21600          # seconds, 6 h
SuspendRate=1              # per minute
ResumeRate=1               # per minute
SuspendProgram=/etc/slurm/slurm-suspend-skript.sh
ResumeProgram=/etc/slurm/slurm-resume-skript.sh
SuspendTimeout=120         # seconds
ResumeTimeout=1200         # seconds
SuspendExcNodes=node-r6-he05,node-r7-he01,node-r5-he19,node-r5-he20,node-r5-
he21,node-r5-he01,node-r5-he03,node-r5-he39,node-r5-he17,node-r5-he18
SuspendExcParts=interactive,jupyter
```

Wie?

SuspendProgram

```
#!/bin/bash
...
pwrcmd="chassis power soft"
...
cmdpref="ipmitool -I lanplus -U root -f $ipmiwfile "
...
if [ -z $ipmikey ]
then
    _cmd="$cmdpref -H $ip $pwrcmd"
else
    _cmd="$cmdpref -H $ip -y $ipmikey $pwrcmd"
fi
```

Wie?

ResumeProgram

```
#!/bin/bash
```

```
...
```

```
pwrcmd="chassis power on"
```

```
...
```

```
cmdpref="ipmitool -I lanplus -U root -f $ipmipwfile "
```

```
...
```

```
if [ -z $ipmikey ]
```

```
then
```

```
    _cmd="$cmdpref -H $ip $pwrcmd"
```

```
else
```

```
    _cmd="$cmdpref -H $ip -y $ipmikey $pwrcmd"
```

```
fi
```

Wann?

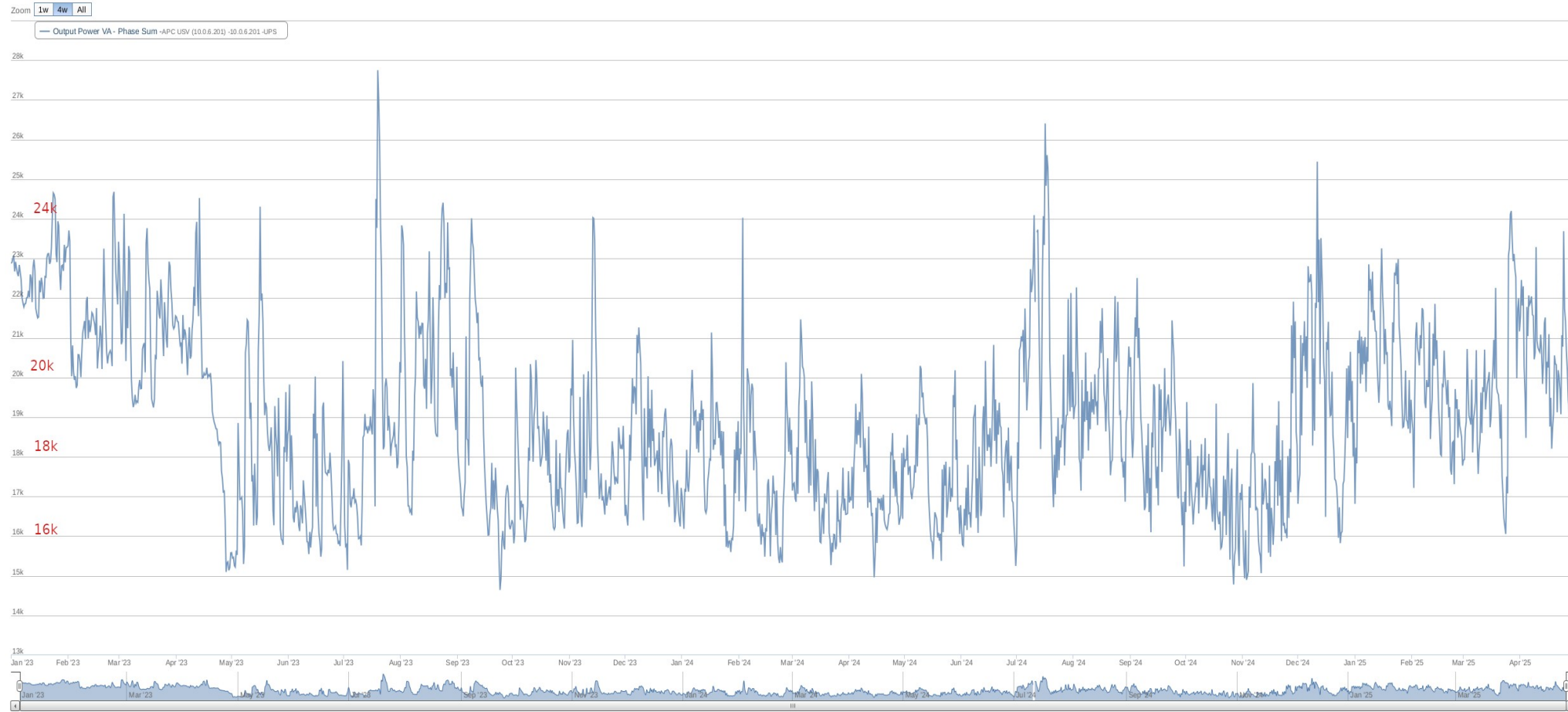
April 2023

- **25.04.2023 – Überlast auf einer PDU, manuelles Abschalten von 10 Rechnenknoten**
- **27.04.2023 – Einführung Energie-Management**

Wirkungen

Verbrauch reduziert (10-15%)

USV-Output, VA



Nebenwirkungen

Nutzer

- positiv bis neutral

Ausfälle

- 1 x GPU-Knoten
 - CPU defekt
 - Keine Verbindung zum Energie-Management

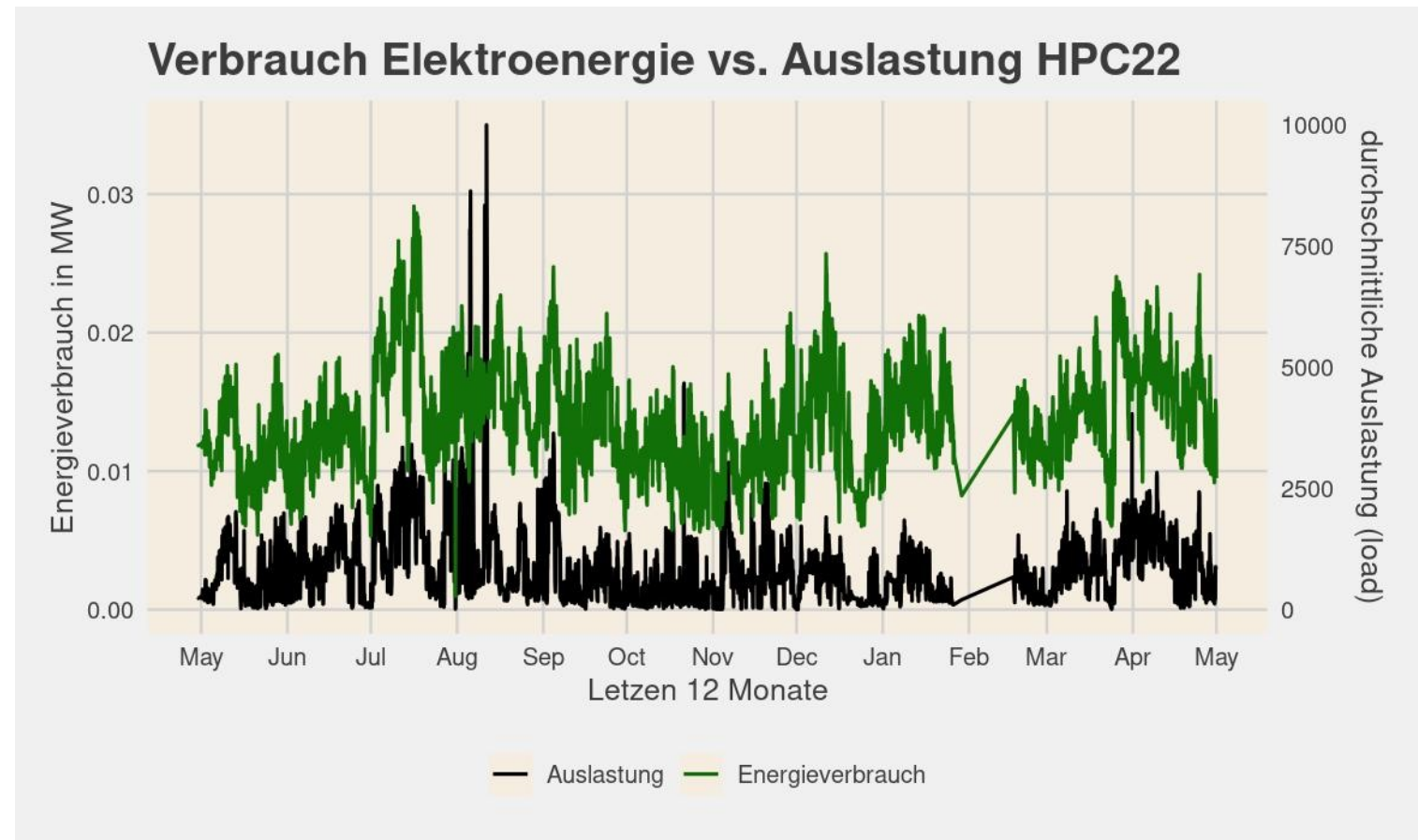




Nebenwirkungen

Monitoring Energieverbrauch

- in MAX



Quellen

Online

- **[1] Slurm Power Saving Guide, https://slurm.schedmd.com/power_save.html**

Bilder

- **Mahmoud Shora, sysGen GmbH**