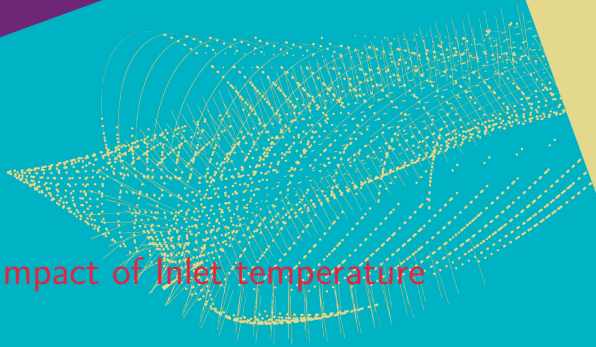
Nikhef



# Power and Performance Impact of Inlet temperature

## Sensor monitoring at scale

E Kooistra a66@nikhef.nl

April 16, 2024

# Origin story

## The hotter the better?

1. Last HEPIX there where some comments on power usage vs inlet
2. Talked about this with our facilities team and offerd to test this out
3. Facilities has been pushing for quite a while to increase inlet themperature

# Why?

Assumption is the mother of all ....

1. Inlet temperature PuE effect
2. Total power usage?
3. Single socket/Dual socket
4. Hot standby vs Both load
5. 80+ Rating of PSUs

# How?

## Or ipmi_exporter is my new best friend

1. Prometheus exporter
2. Polls sensors, dcmi, chassis and sel
3. Flexible filtering of sensor states
4. Operating system independent

# What?

Read out all the sensors!

| ID | Name | Type | Reading | Units | Event |
|---|---|---|---|---|---|
| 1 | Inlet_Temp | Temperature | 26.00 | C | 'OK' |
| 2 | Outlet_Temp | Temperature | 29.00 | C | 'OK' |
| 3 | CPU_Temp | Temperature | 34.00 | C | 'OK' |
| 4 | DIMMG1_Temp | Temperature | 37.00 | C | 'OK' |
| 5 | DIMMG2_Temp | Temperature | 38.00 | C | 'OK' |
| 6 | CPU_VR_Temp | Temperature | 30.00 | C | 'OK' |
| 7 | SOC_VR_Temp | Temperature | 37.00 | C | 'OK' |
| 8 | VDD_ABCD_Temp | Temperature | 35.00 | C | 'OK' |
| 9 | VDD_EFGH_Temp | Temperature | 35.00 | C | 'OK' |
| 10 | PSU1_Temp | Temperature | 26.00 | C | 'OK' |
| 11 | PSU2_Temp | Temperature | 27.00 | C | 'OK' |
| 12 | Riser1_GPU_Temp | Temperature | 31.00 | C | 'OK' |
| 13 | Riser2_GPU_Temp | Temperature | 30.00 | C | 'OK' |
| 17 | PCIe1_Inlet_Temp | Temperature | 29.00 | C | 'OK' |
| 18 | PCIe2_Inlet_Temp | Temperature | 31.00 | C | 'OK' |
| 22 | Riser_Int_Temp | Temperature | 26.00 | C | 'OK' |
| 23 | Riser_1_Temp | Temperature | 26.00 | C | 'OK' |
| 24 | Riser_2_Temp | Temperature | 25.00 | C | 'OK' |
| 26 | Front_NVME_Temp | Temperature | 31.00 | C | 'OK' |

# What?

Read out all the sensors!

| ID | Name | Type | Reading | Units | Event |
|---|---|---|---|---|---|
| 29 | CPU_VDDCR | Voltage | 0.74 | V | 'OK' |
| 30 | SOC_VDDCR | Voltage | 0.82 | V | 'OK' |
| 31 | VDD_MEM_ABCD | Voltage | 1.21 | V | 'OK' |
| 32 | VDD_MEM_EFGH | Voltage | 1.21 | V | 'OK' |
| 33 | VDD_VPP_ABCD | Voltage | 2.50 | V | 'OK' |
| 34 | VDD_VPP_EFGH | Voltage | 2.54 | V | 'OK' |
| 35 | P5V_VDD_RearMID | Voltage | 5.09 | V | 'OK' |
| 36 | P3V_BAT | Voltage | 3.16 | V | 'OK' |
| 37 | VDDCR_SOC_DUAL | Voltage | 0.90 | V | 'OK' |
| 38 | VDD_5_DUAL | Voltage | 5.04 | V | 'OK' |
| 39 | VDD_33_RUN | Voltage | 3.35 | V | 'OK' |
| 40 | VDD_VTT_EFGH | Voltage | 0.59 | V | 'OK' |
| 41 | VDD_18_RUN | Voltage | 1.81 | V | 'OK' |
| 42 | VDD_18_DUAL | Voltage | 1.83 | V | 'OK' |
| 43 | VDD_5_RUN | Voltage | 5.17 | V | 'OK' |
| 44 | P12V_RUN | Voltage | 11.96 | V | 'OK' |
| 45 | VDD_VTT_ABCD | Voltage | 0.59 | V | 'OK' |
| 46 | P1V15_BMC | Voltage | 1.16 | V | 'OK' |
| 47 | P1V2_DDR_BMC | Voltage | 1.20 | V | 'OK' |
| 48 | VDD_33_DUAL | Voltage | 3.37 | V | 'OK' |

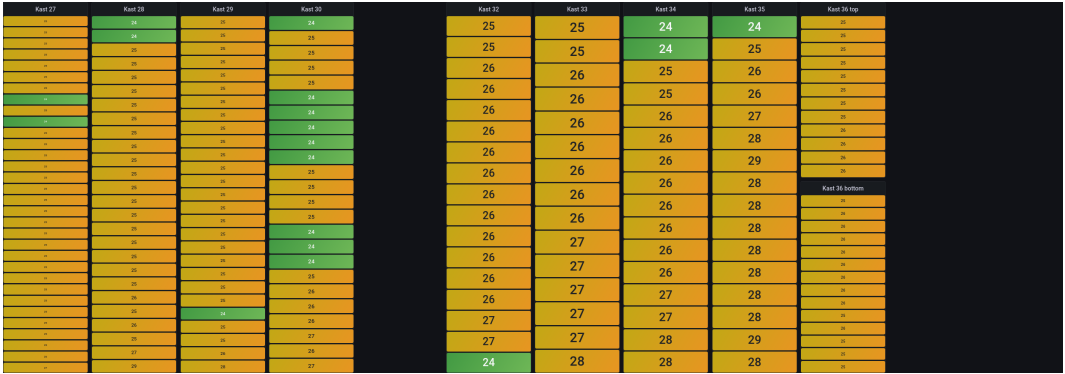
# What?

Read out all the sensors!

| ID | Name | Type | Reading | Units | Event |
|---|---|---|---|---|---|
| 63 | PSU1_PIN | Other Units Based Sensor | 138.00 | W | 'OK' |
| 64 | PSU1_POUT | Other Units Based Sensor | 124.00 | W | 'OK' |
| 65 | PSU1_VIN | Voltage | 237.90 | V | 'OK' |
| 66 | PSU1_VOUT | Voltage | 12.10 | V | 'OK' |
| 67 | PSU1_IIN | Current | 0.65 | % A | 'OK' |
| 68 | PSU2_PIN | Other Units Based Sensor | 120.00 | W | 'OK' |
| 69 | PSU2_POUT | Other Units Based Sensor | 116.00 | W | 'OK' |
| 70 | PSU2_VIN | Voltage | 234.24 | V | 'OK' |
| 71 | PSU2_VOUT | Voltage | 12.10 | V | 'OK' |
| 72 | PSU2_IIN | Current | 0.55 | % A | 'OK' |
| 73 | CPU_Power | Other Units Based Sensor | 53.00 | W | 'OK' |
| 74 | MEM_Power | Other Units Based Sensor | 20.00 | W | 'OK' |
| 75 | Total_Power | Other Units Based Sensor | 240.00 | W | 'OK' |
| 76 | Airflow_rate | Other Units Based Sensor | 84.00 | CFM | 'OK' |

# Grafana time!

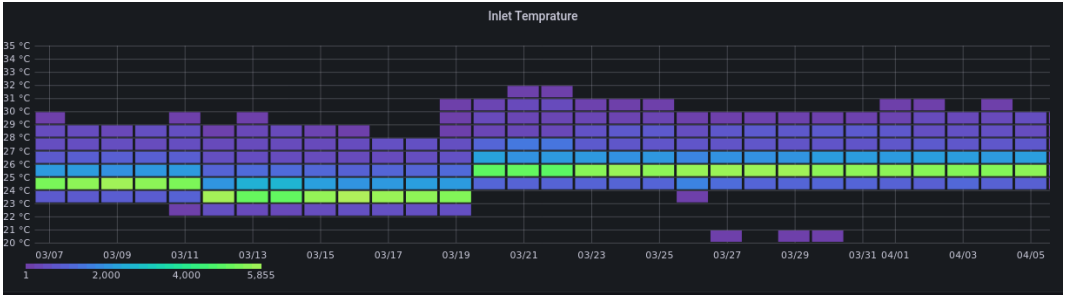
## Pretty plots make everything easier




The higher in the rack the lower the inlet temperature

# Grafana time!

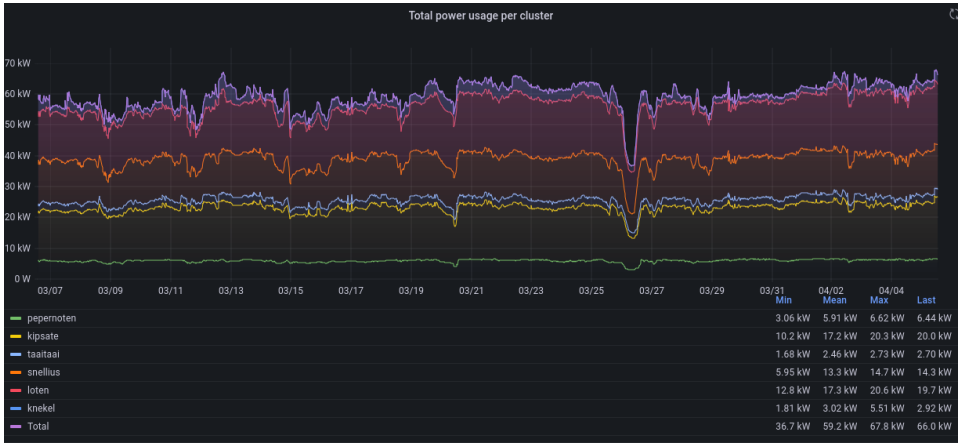
## Pretty plots make everything easier




Inlet air temp is a lower threshold

# Grafana time!

## Pretty plots make everything easier




Job usage influences power usage a lot

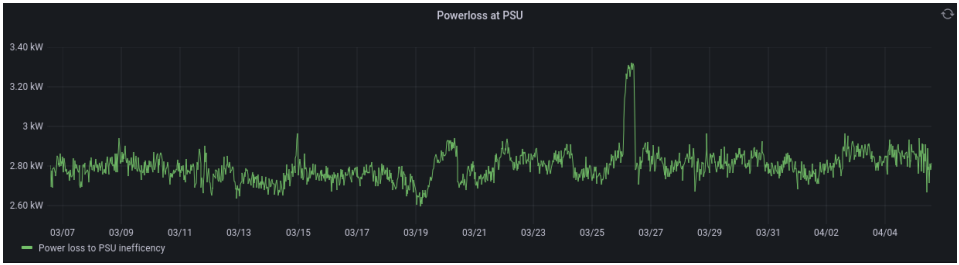
# Grafana time!

## Pretty plots make everything easier


1U boxes (yellow) use more power then 2U (green)

# Grafana time!

## Pretty plots make everything easier




Hot standby reduces efficency

# 80+ ?

## Or how its not that easy

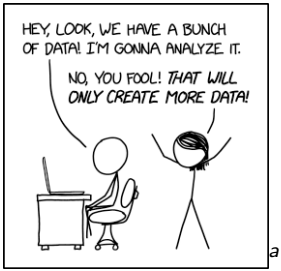
- Power supply ratings are not straight forward
- Hot / cold standy options differ per vendor
- Power feed balancing to prevent cascading failure
- Inrush currents can be still an issue

| Level/load | 10% | 20% | 50% | 100% |
|---|---|---|---|---|
| Gold | - | 88% | 92% | 88% |
| Platinum | - | 90% | 94% | 91% |
| Titanium | 90% | 94% | 96% | 91% |

# Future

## Collect all the data!

- Intergrate Node exporter data
- Export job scheduler data
- Power efficency of a job
- Communicate with vendors for more sensor probe points
- This is still an ongoing project.
- 128 core Bergamo nodes getting delivered soon




[a]

[a] https://xkcd.com/2582/