Monitoring ALICE Analysis Facility infrastructure

operation and visibility

Ádám Pintér
Wigner RCP – Datacenter
05/31/2024

Short introduction

- Started in 2012, CERN Tier-0 for 7y
- 4 rooms (~400 racks)
- WSCLAB (AF, Tier-2, GPULAB, Virgo/LIGO)
- ALICE Analysis Facility ~ 9 active racks
- Other IT projects (HUN-REN Cloud, HUN-REN KININ, HUN-REN ARP)

Capacity

- ALICE Analysis Facility
- Start: November of 2020
- * 3700 vCPU
- * 7.5 TB RAM
- * 2.3 PB storage (raw)
- OS layer: datacenter
- Service layer: Gergely Barnaföldi, Gábor Bíró & their team

Hardware aging

- Monitoring is an important tool to keep infrastructure healthy and to prolong cluster lifetime.

Defining checks

- Specific environments
- We selected: Sensu
- Server types (worker node, storage)
- Basic checks, advanced and complex
- In almost real time

- keepalive
- check-cpu, check-ram, swap-memory, temperature
- RAID-root-partition, RAID1-status, smart-short-test, disk-read-write-speed
- interface-status, check-internet-access, firewall-access
- switch-ping, switch-PSU, switch-traffic, master-switch-traffic

☐	**ce061.alice-af.wigner.hu** *agent • centos/amd64*	compute • entity:ce061.alice-af.wigner.hu server • compute-90	CLASS	SUBSCRIPTION	last seen *seconds ago*
☐	**ce064.alice-af.wigner.hu** *agent • centos/amd64*	compute • entity:ce064.alice-af.wigner.hu server • compute-93			last seen *seconds ago*
☐	**ce065.alice-af.wigner.hu** *agent • centos/amd64*	compute • entity:ce065.alice-af.wigner.hu server • compute-94			last seen *just now*
☐	**ce066.alice-af.wigner.hu** *agent • centos/amd64*	compute • entity:ce066.alice-af.wigner.hu server • compute-95			last seen *seconds ago*
☐	**ce067.alice-af.wigner.hu** *agent • centos/amd64*	compute • entity:ce067.alice-af.wigner.hu server • compute-96			last seen *seconds ago*
☐	**compute-04.maas** *agent • centos/amd64*	compute • entity:compute-04.maas server • compute-4			last seen *seconds ago*
☐	**compute-05.maas** *agent • centos/amd64*	compute • entity:compute-05.maas			last seen *67 napja*
☐	**compute-06.maas** *agent • centos/amd64*	compute • entity:compute-06.maas			last seen *67 napja*
☐	**compute-11.maas** *agent • centos/amd64*	compute • entity:compute-11.maas			last seen *67 napja*
☐	**compute-81.maas** *agent • centos/amd64*	compute • entity:compute-81.maas server • compute-81			last seen *just now*
☐	**eos-fst17.alice-af.wigner.hu** *agent • centos/amd64*	compute • entity:eos-fst17.alice-af.wigner.hu server • storage-97			last seen *just now*
☐	**eos-fst18.alice-af.wigner.hu** *agent • centos/amd64*	compute • entity:eos-fst18.alice-af.wigner.hu server • storage-98			last seen *seconds ago*
☐	**eos-fst19.alice-af.wigner.hu** *agent • centos/amd64*	compute • entity:eos-fst19.alice-af.wigner.hu server • storage-99			last seen *just now*
☐	**eos-fst20.alice-af.wigner.hu** *agent • centos/amd64*	compute • entity:eos-fst20.alice-af.wigner.hu server • storage-100			last seen *just now*
☐	**mgmt-23** *agent • ubuntu/amd64*	backend • compute • entity:mgmt-23 server • mgmt-23			last seen *seconds ago*

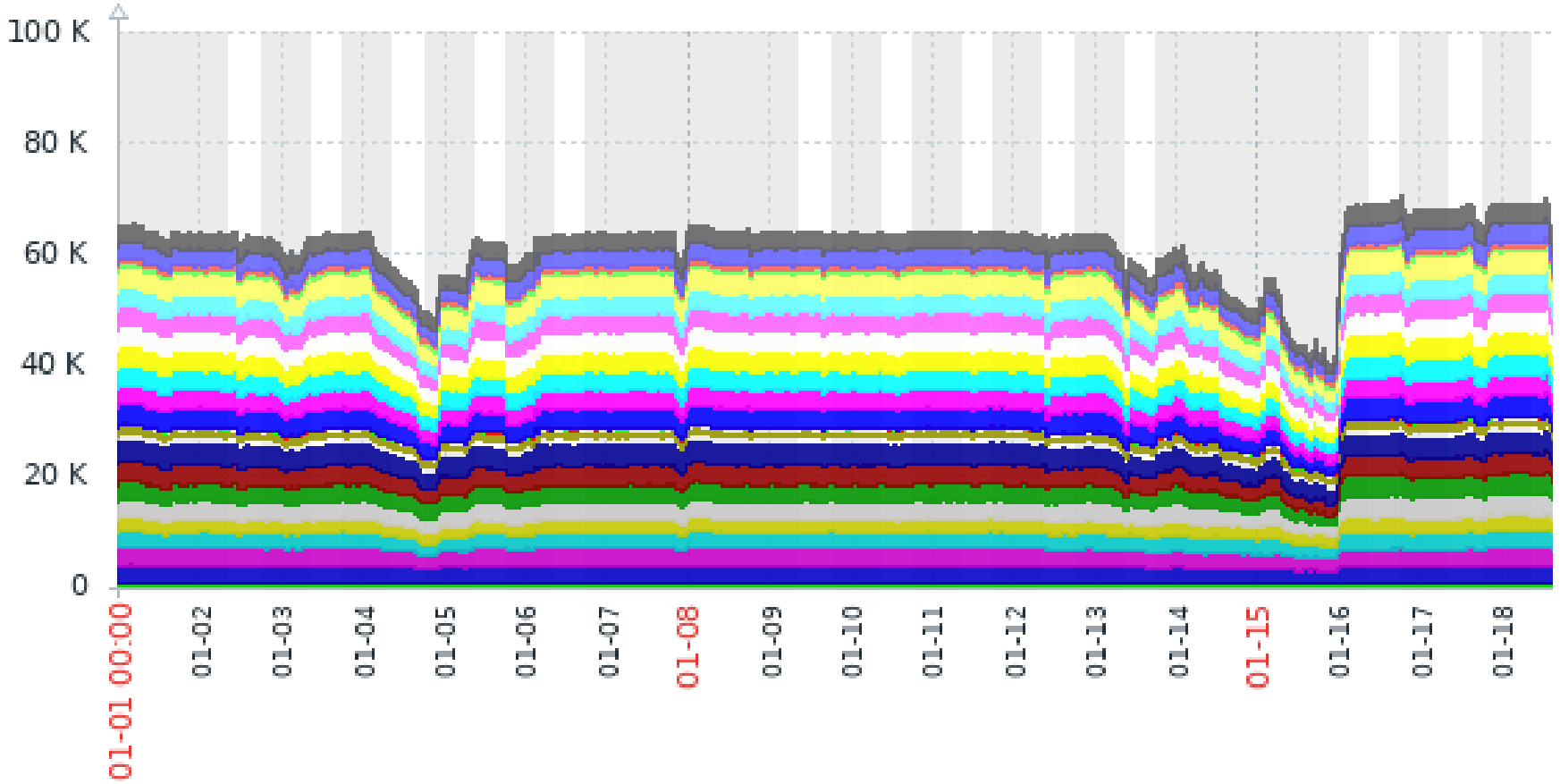
Total rows: 21

Network separation

- Dedicated VLAN network for monitoring
- We selected 10G interface, it is necessary to change to 1G interface
- In order to minimize interference with real worker node traffic

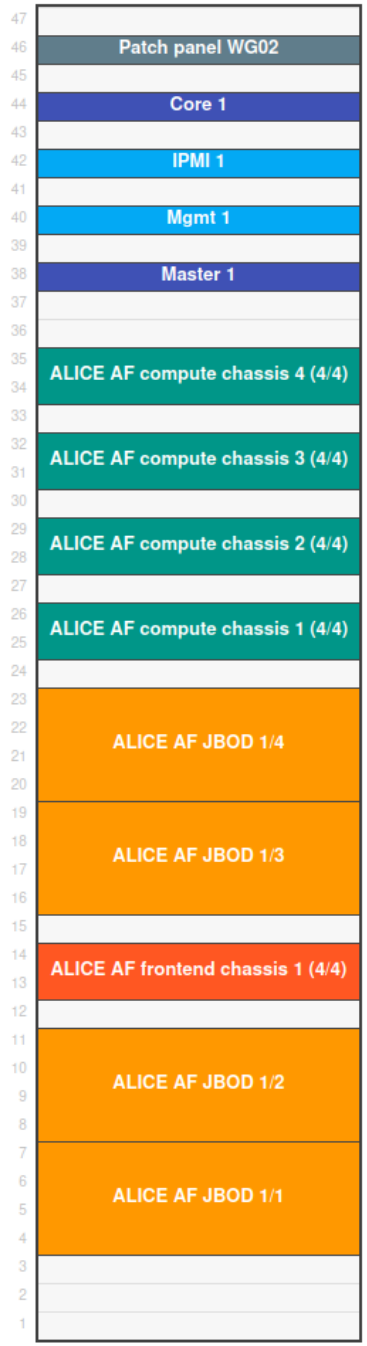
Power usage of cluster

- Power consumption and electricity bills are important factors nowadays
- In order to see the details, we are using a visualization solution for power usage statistics based on each rack


100 K
80 K
60 K
40 K
20 K
0
01-01 00:00
01-02
01-03
01-04
01-05
01-06
01-07
01-08
01-09
01-10
01-11
01-12
01-13
01-14
01-15
01-16
01-17
01-18

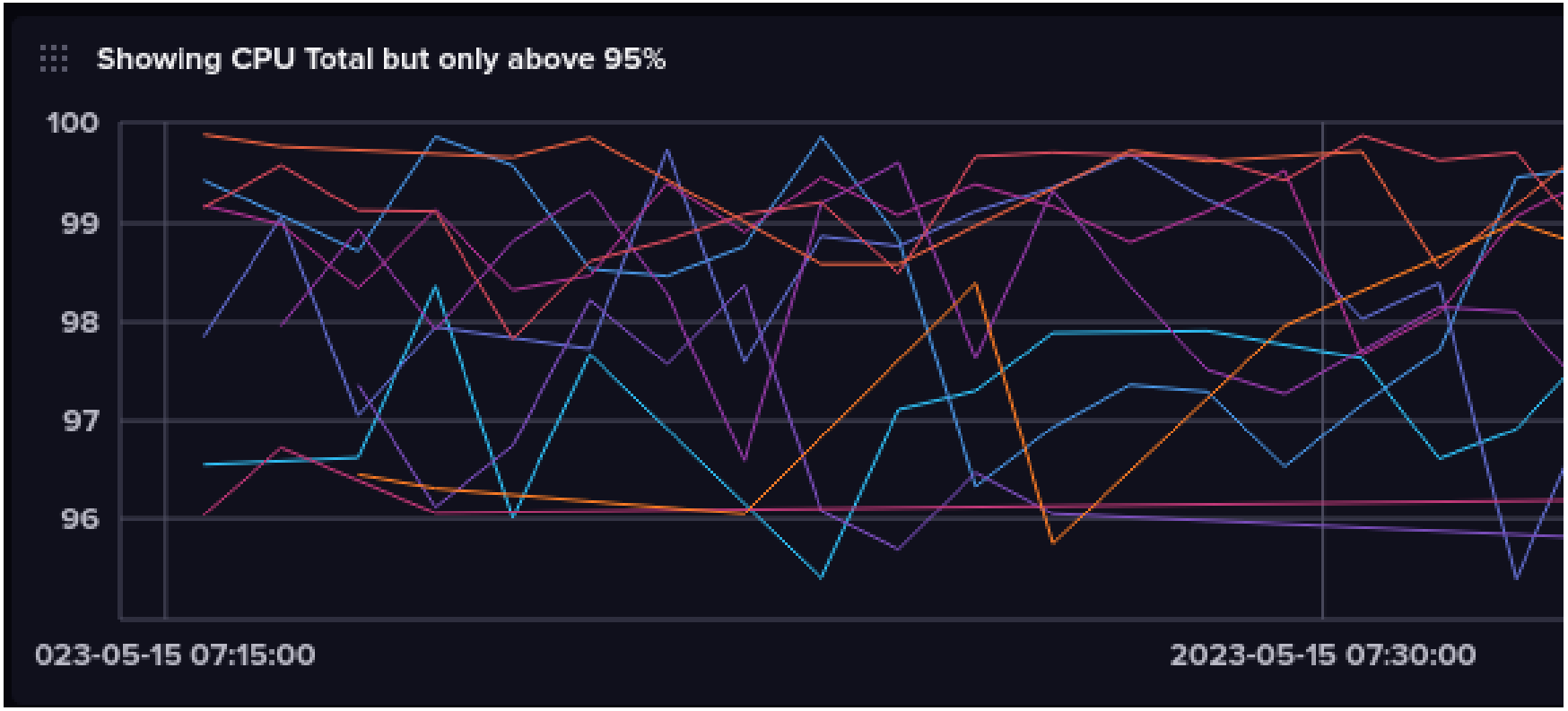
Scaling up

- Scale up the monitoring system efficiently
- We are using automation tools for node preparation and installation (MAAS & Ansible)
- 100+ server entities: inventory
- NetBox for (rack, server, switch, IPAM, cabling)


47
46
Patch panel WG02
45
44
Core 1
43
42
IPMI 1
41
40
Mgmt 1
39
38
Master 1
37
36
35
ALICE AF compute chassis 4 (4/4)
34
33
32
ALICE AF compute chassis 3 (4/4)
31
30
29
ALICE AF compute chassis 2 (4/4)
28
27
26
ALICE AF compute chassis 1 (4/4)
25
24
23
ALICE AF JBOD 1/4
22
21
20
19
ALICE AF JBOD 1/3
18
17
16
15
14
ALICE AF frontend chassis 1 (4/4)
13
12
11
ALICE AF JBOD 1/2
10
9
8
7
ALICE AF JBOD 1/1
6
5
4
3
2
1

Event and data history

- Historical data is also very valuable
- We integrated a database solution into our monitoring workflow: InfluxDB
- Continuous disk tests (S.M.A.R.T.)


Showing CPU Total but only above 95%
100
99
98
97
96
023-05-15 07:15:00
2023-05-15 07:30:00

Roadmap for future dev. (I.)

- Scheduled backup for monitoring data
- Proper alerting based on pre-defined warning and critical levels
- Iterative time-based optimization for running checks
- HTCondor service monitoring

Roadmap for future dev. (II.)

- Next infrastructure: Virgo/LIGO & GPULAB
- Monitoring GPU RAM
- GPU-utilization
- GPU card temperature

- New smart monitoring ideas for complex cases

Thank you

If you have any questions:

- IT team: Tádé Szlachányi, Gergő Miklós, Norbert Kóródi, Artúr Dummel, Gábor Haraszti
- In person (KFKI Campus – building 18)
- pinter.adam@wigner.hun-ren.hu