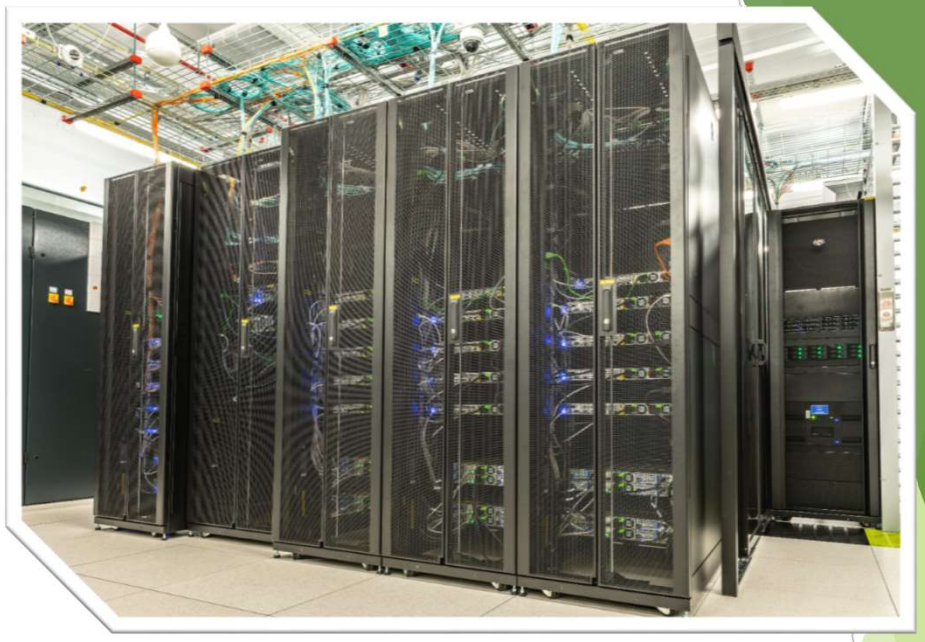
Реальные ситуации, когда анализ данных о работе СКЦ помог существенно улучшить качество его работы

Шкандыбин Ю.Н.,
Малютин А.К.,

Докладчик: Антон Малютин

Skoltech
Skolkovo Institute of Science and Technology

ZHORES cluster



HPC cluster with Infiniband EDR backend

- About 1 PFLOP/s peak performance
- More than 100 compute nodes
- 29 GPU nodes, 114 GPUs total
- 0,5 PB parallel file system and 0,7 PB tape library
- 86.8 kW power consumption
- #8 at Russian top50

ZHORES
SUPERCOMPUTER
HIGH PERFORMANCE COMPUTING AND BIG DATA
SKOLTECH CDISE

Our statistic collection and analyzis tools

Zabbix

- Hosts: IPMI statistic SNMP + Linux staistic Zabbix agent
- Slurm: scripts implemented to Zabbix agent

Elasticsearch + Kibana (1 instance):

- Slurm

Elasticsearch + Kibana (deprecated instance):

- Syslog
- Snoopy (user commands monitor)

Slurm account manager

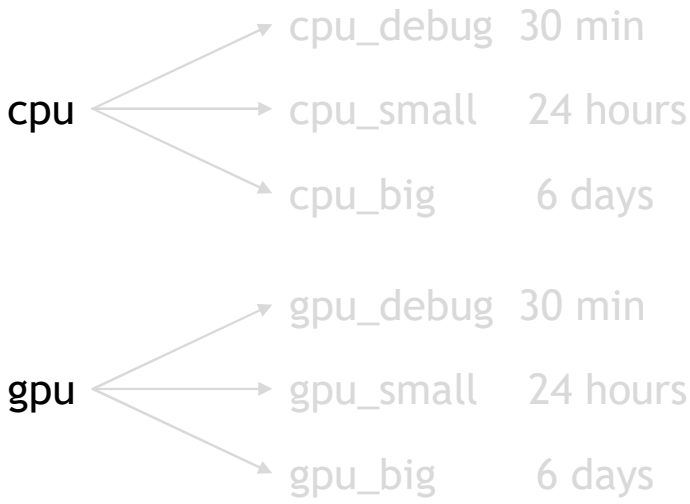
How can we use analysis results

- Change queues/partitions configuration
- Add new applications to the cluster
- Adjust compute node configuration (HW, Linux, system software)
- Buy an additional hardware
- Improve purchase and cluster development strategies

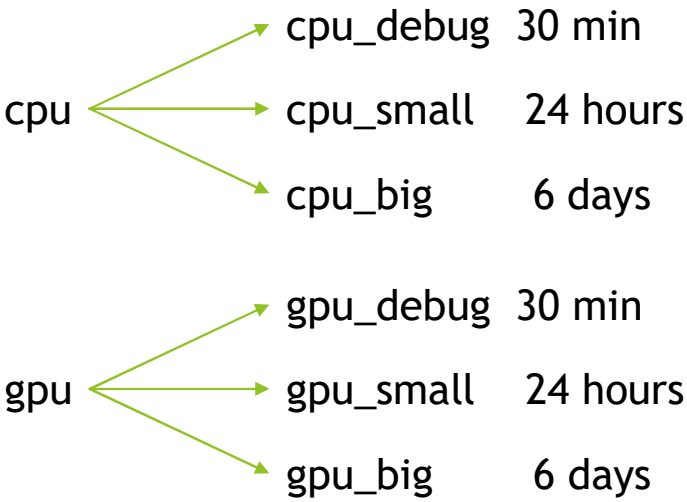
How can we use analysis results

- Change queues/partitions configuration
- Add new applications to the cluster
- Adjust compute node configuration (HW, Linux, system software)
- Buy an additional hardware
- Improve purchase and cluster development strategies

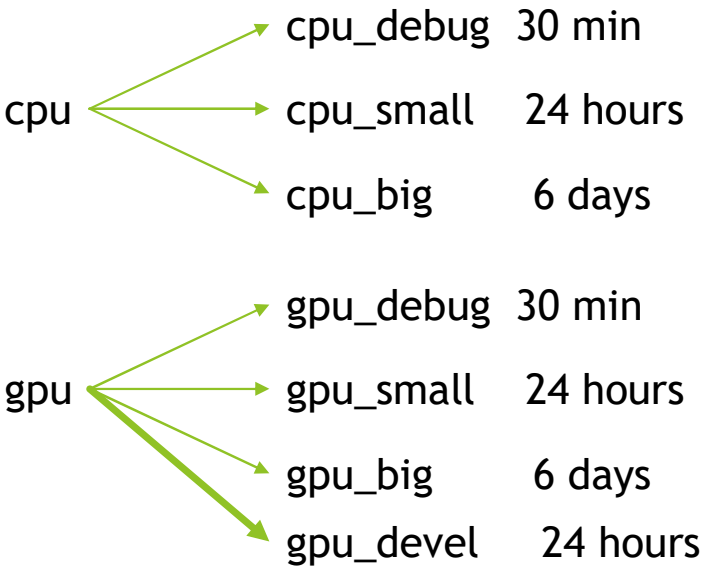
Change queues/partitions configuration



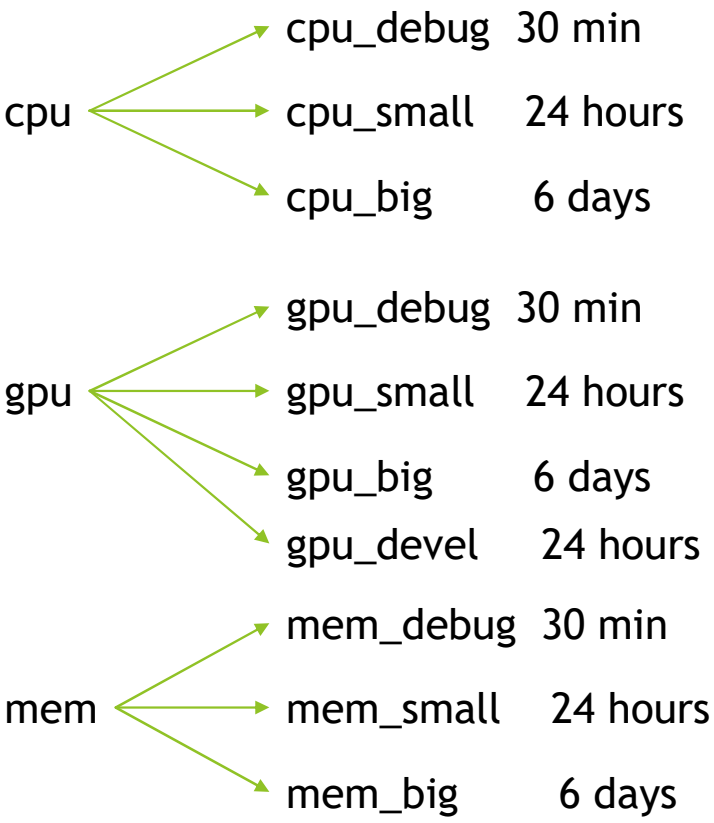
Change queues/partitions configuration



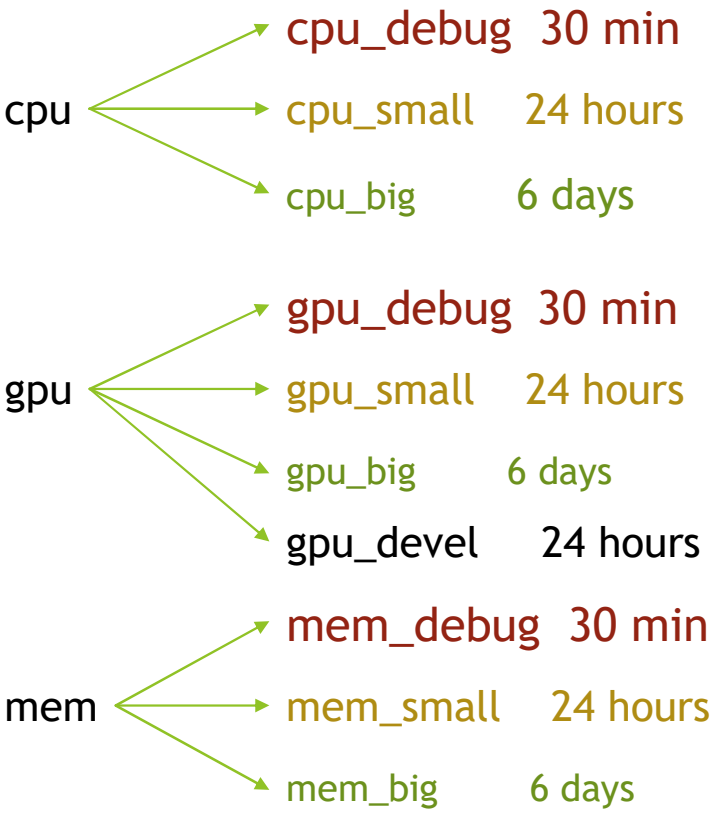
Change queues/partitions configuration



Change queues/partitions configuration

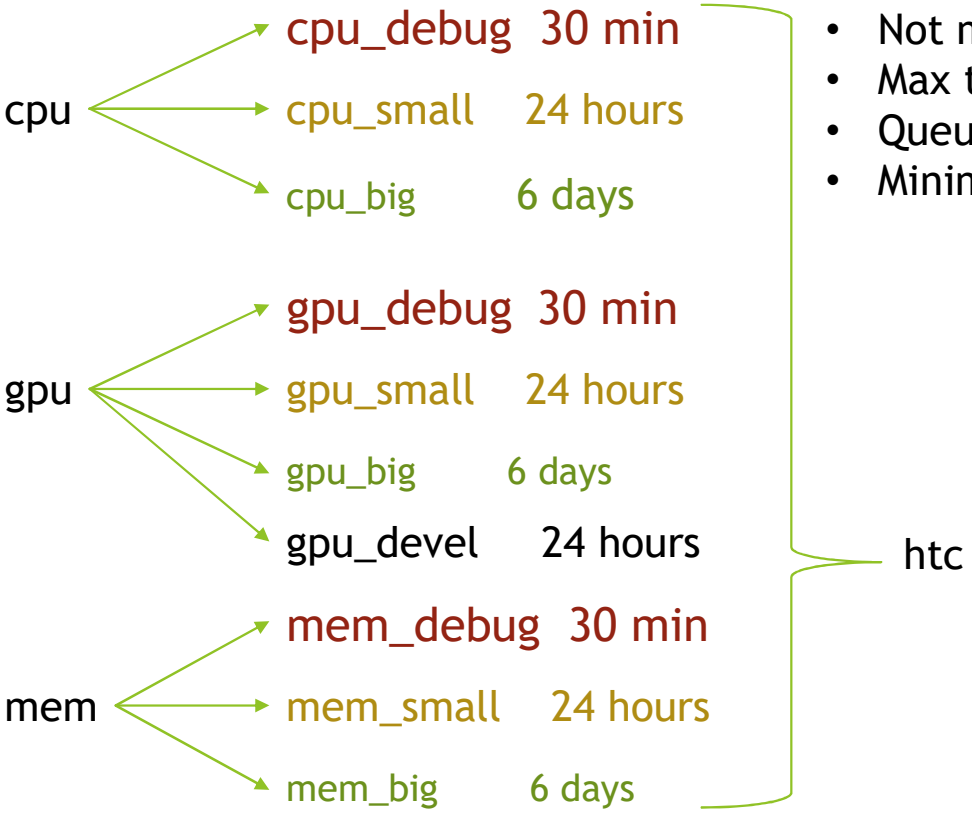


Change queues/partitions configuration



- Not more than ½ of partition resources per user
- Max tasks per user
- Queues priority
- Minimum memory limit per CPU

Change queues/partitions configuration



- Not more than ½ of partition resources per user
- Max tasks per user
- Queues priority
- Minimum memory limit per CPU

How can we use analysis results

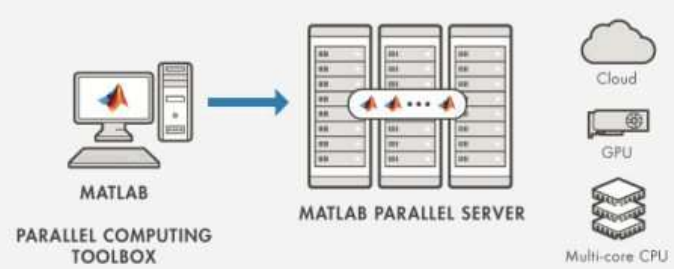
- Change queues/partitions configuration
- Add new applications to the cluster
- Adjust compute node configuration (HW, Linux, system software)
- Buy an additional hardware
- Improve purchase and cluster development strategies

Add new applications to the cluster

Due to popularity dedicated JupyterHub was deployed

A user can start a Python Jupyter Lab/Notebook with out SSH session to the cluster. In the same time he can run his jobs on any partitions from the web interface.

```
>> parpool(parcluster('HPC1'),100);
>> parfor i=1:3000
>>     c(i,:) = eig(rand(1000));
>> end
```



How can we use analysis results

- Change queues/partitions configuration
- Add new applications to the cluster
- Adjust compute node configuration (HW, Linux, system software)
- Buy an additional hardware
- Improve purchase and cluster development strategies

Adjust compute node configuration (HW, Linux, system software)

- Update Linux kernel
- Update drivers
- Adjust BIOS settings (Hyper threading, power consumptions, etc)

How can we use analysis results

- Change queues/partitions configuration
- Add new applications to the cluster
- Adjust compute node configuration (HW, Linux, system software)
- Buy an additional hardware
- Improve purchase and cluster development strategies

Buy an additional hardware

A special server with 10 blades with 4 CPUs and equipped large amount of memory was bought. Each blade has 80 cores and up to 3TB of memory.

It covers jobs:
- with high memory needs up to 3TB;
- with problems belong spreading load to several nodes.

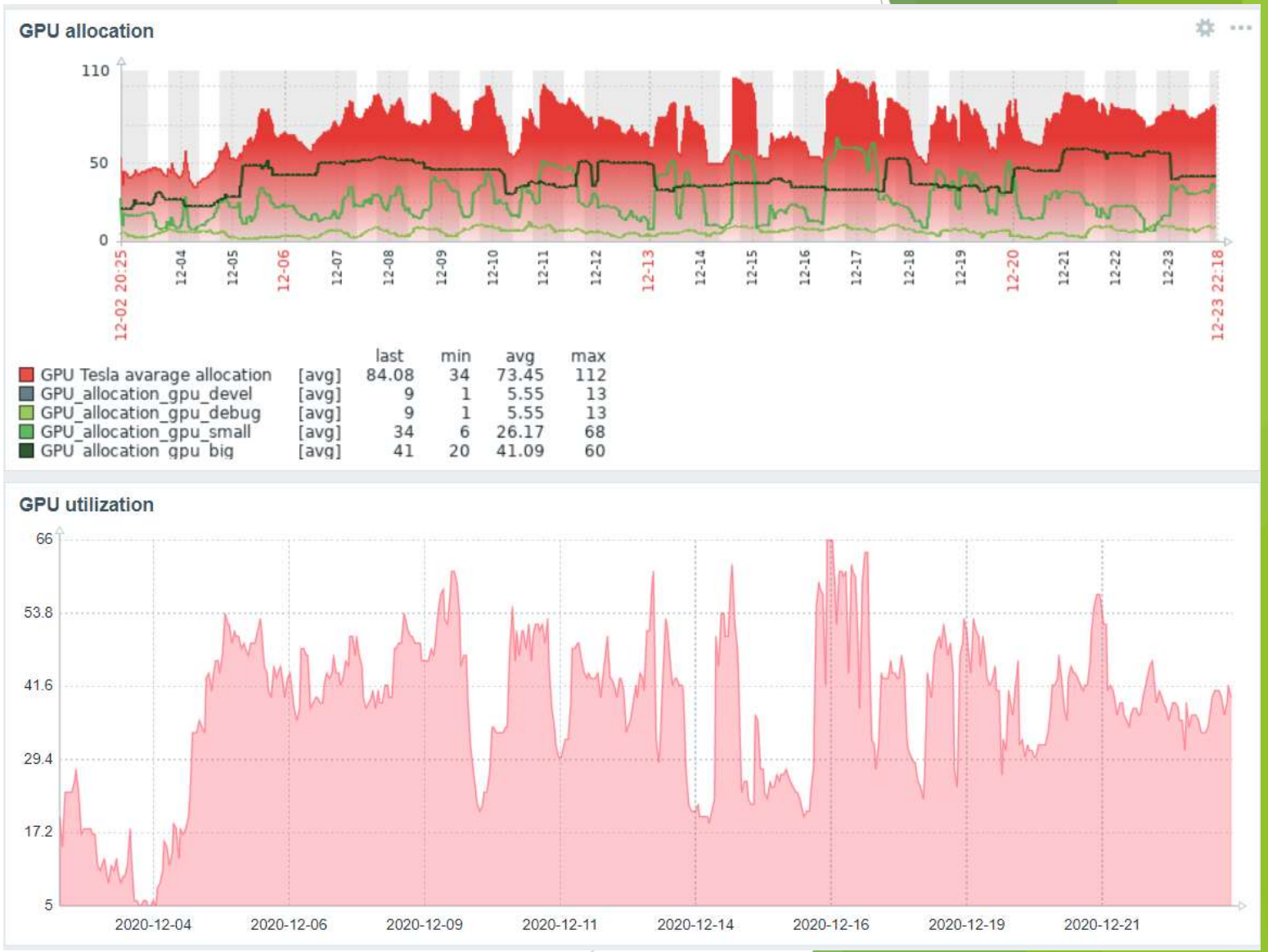
How can we use analysis results

- Change queues/partitions configuration
- Add new applications to the cluster
- Adjust compute node configuration (HW, Linux, system software)
- Buy an additional hardware
- Improve purchase and cluster development strategies

Improve purchase and cluster development strategies

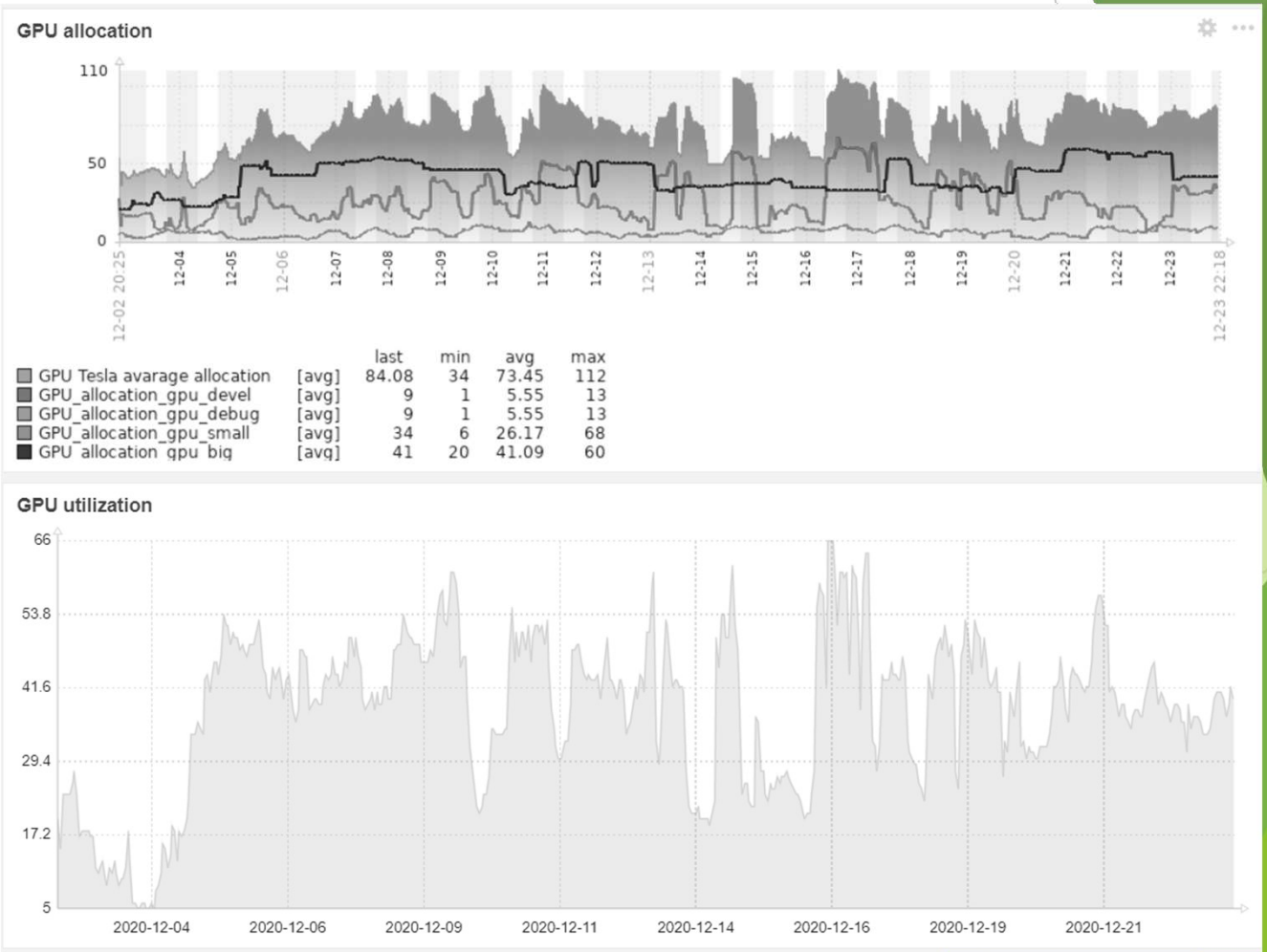
- GPU is a key reso
- Partitioning In t research lculations will not be affected.

GPU efficiency



GPU efficiency and human problem

- User education
- User education
- User education
- User education



Q&A