



מכון
ויצמן
למדע

WEIZMANN
INSTITUTE
OF SCIENCE

Monitoring small HPC with Zabbix

Choose Your Weapon

Mark Vilensky
Scientific Computing Manager
Dept. of Chemical and Biological Physics
Weizmann Institute of Science

Riga, October 2023



מכון
ויצמן
למדע

WEIZMANN
INSTITUTE
OF SCIENCE

Monitoring small HPC with Zabbix

About us

Chemistry Faculty

- Five departments
- 14 heavy usage groups, ~50 active users
- Long HPC history (clusters, Altix)
- Faculty cluster was planned in 2009 and built in 2010, and it is continuously growing





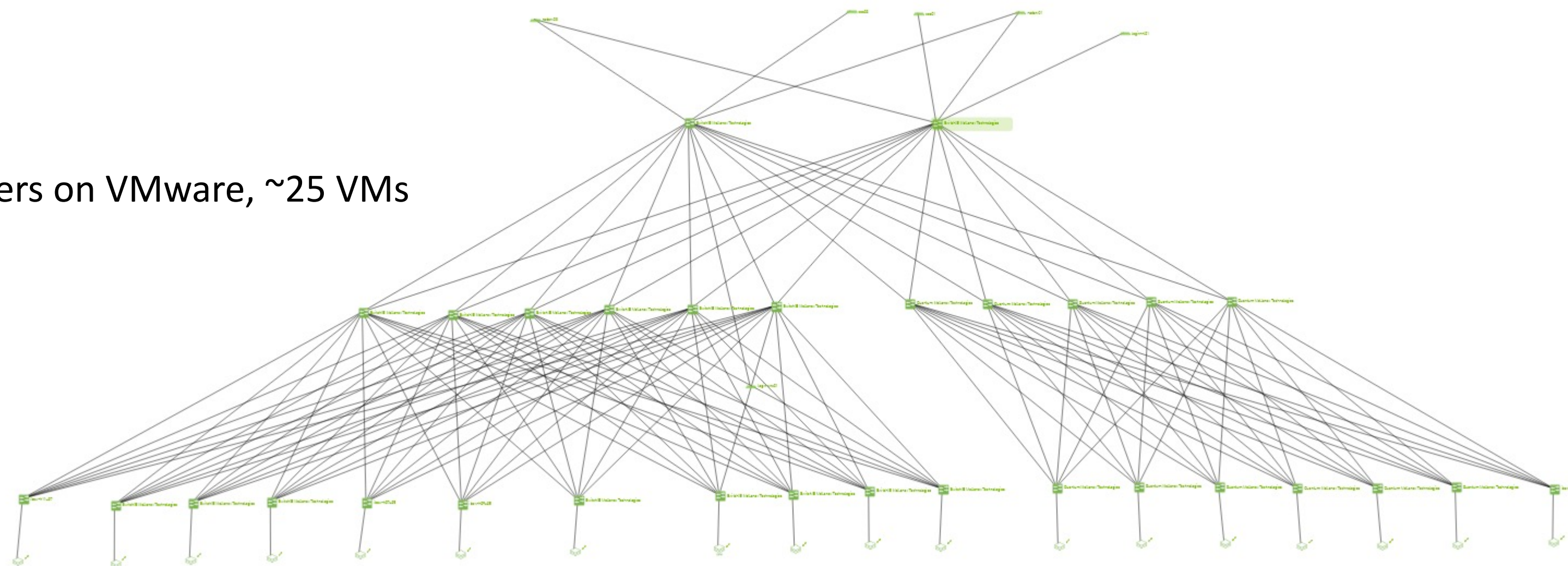
מכון
ויצמן
למדע

WEIZMANN
INSTITUTE
OF SCIENCE

Monitoring small HPC with Zabbix

Cluster topology

- 584 dual socket compute nodes
- Memory ranges between 4 to 16 GB/core
- Total 25,652 cores
- Two Infiniband partitions 5:1 blocking and non-blocking
- GPFS/DDN
- CentOS 7.9
- PBS Professional
- Management servers on VMware, ~25 VMs
- Staff = 3!





מכון
ויצמן
למדע

WEIZMANN
INSTITUTE
OF SCIENCE

Monitoring small HPC with Zabbix

Hardware

- Compute:
 - HPE Apollo 2000 gen 10
 - HPE ProLiant DL360 gen 9
 - Dell C6420
 - Dell C6520
 - Dell R750
- Login nodes
- Infiniband switches: EDR, HDR
- Infiniband adapters: EDR, HDR/100
- Ethernet switches – HPE 5406R zl2, Aruba 2930M, Aruba 6300M
- Storage – DDN 7990 embedded
- Head nodes – VMware on 2 x Lenovo SR630 V2, Lenovo DE4000H





מכון
ויצמן
למדע

WEIZMANN
INSTITUTE
OF SCIENCE

History

- Operating system level
 - Nagios
 - Passive checks
 - Ganglia
- Hardware
 - HP SIM
 - HPE OneView
 - Dell OME / Cloud IQ
- HPE switches

Problems

- Nagios R.I.P
- Ganglia R.I.P since 2018
 - Python 2
- HPE switches
- Useless alerts
- Flooding alerts
- Complicated setup
- Unexpected faults



מכון
ויצמן
למדע

WEIZMANN
INSTITUTE
OF SCIENCE

Monitoring small HPC with Zabbix

Requirements

- SCALABILITY!
- Auto discovery
- Host groups
- Custom checks
- Only useful alerts
- Smooth operation
- Picture for the management



Alternatives

- Zabbix
- Prometheus
- Commercial?

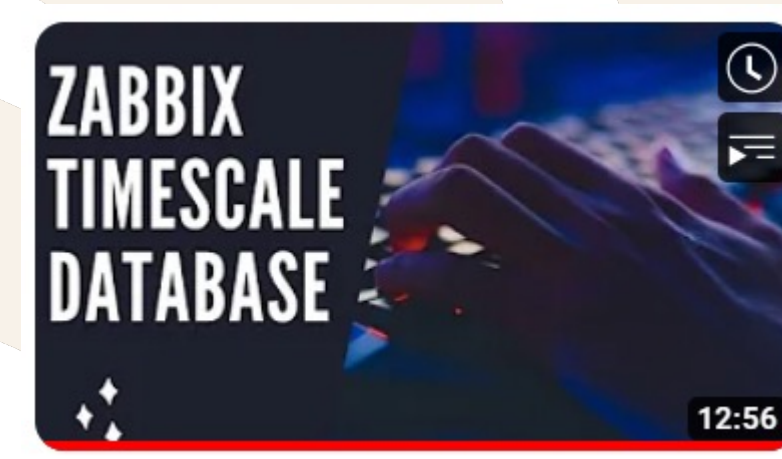
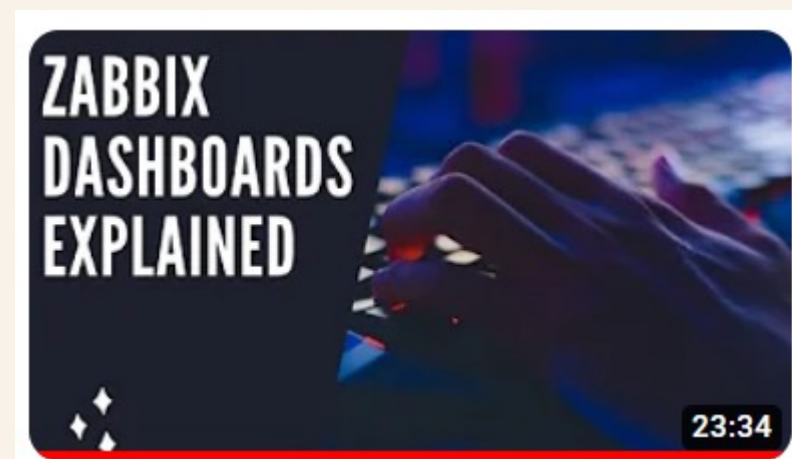


מכון
ויצמן
למדע

WEIZMANN
INSTITUTE
OF SCIENCE

Monitoring small HPC with Zabbix

A secret link <https://www.youtube.com/@DmitryLambert>





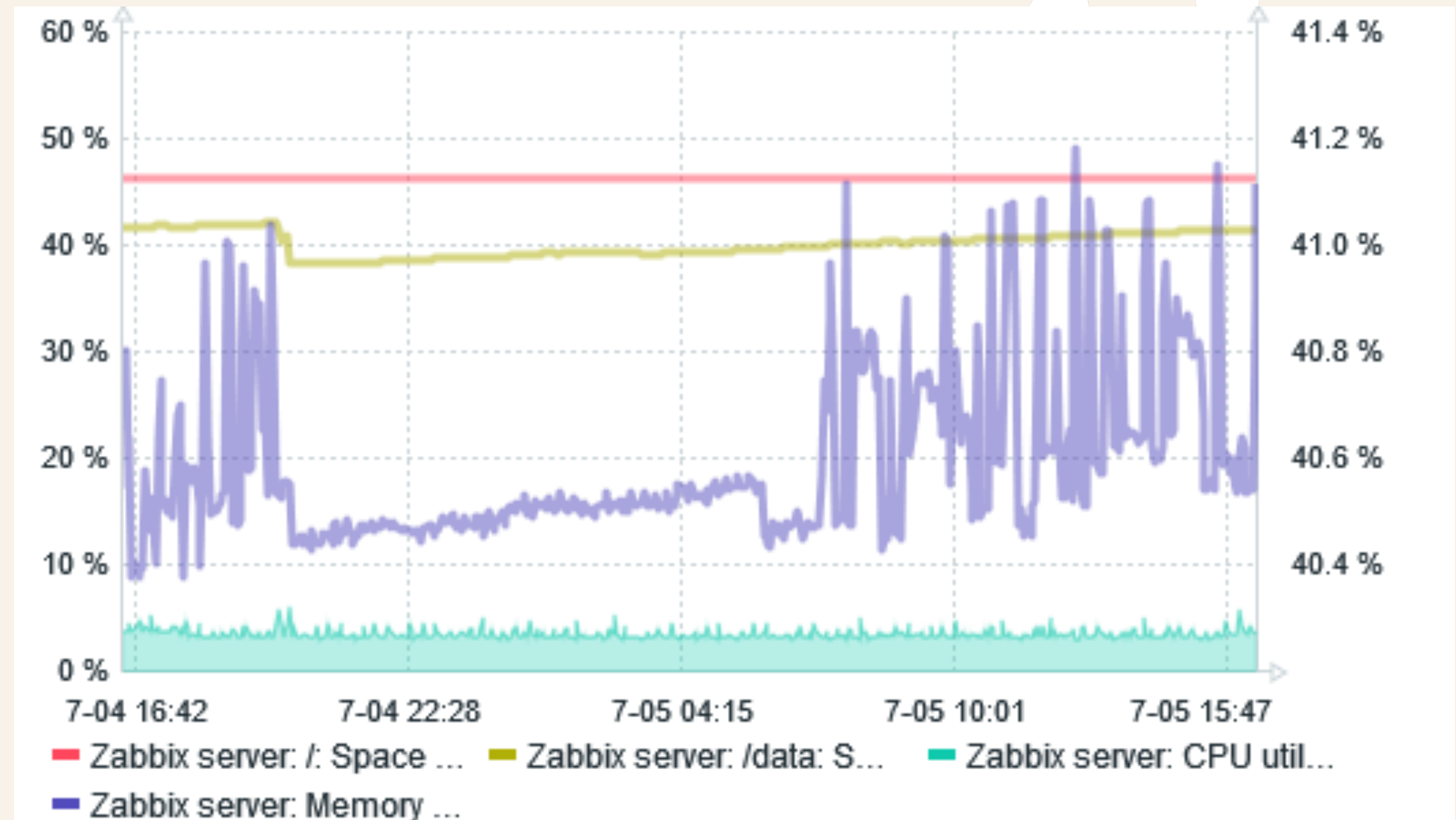
מכון
ויצמן
למדע

WEIZMANN
INSTITUTE
OF SCIENCE

Monitoring small HPC with Zabbix

Zabbix server

- Version: 6.0.19
- 16 vCPU
- 32 GB RAM
- 105 GB disk for OS
- 200 GB disk for data
- 3 NIC
- Rocky 8.7
- **PostgreSQL 14.6.1**
- **Timescale DB 2.8.1**
- History – ~~14~~ days 180 days
- Trends – ~~60~~ days 1 year



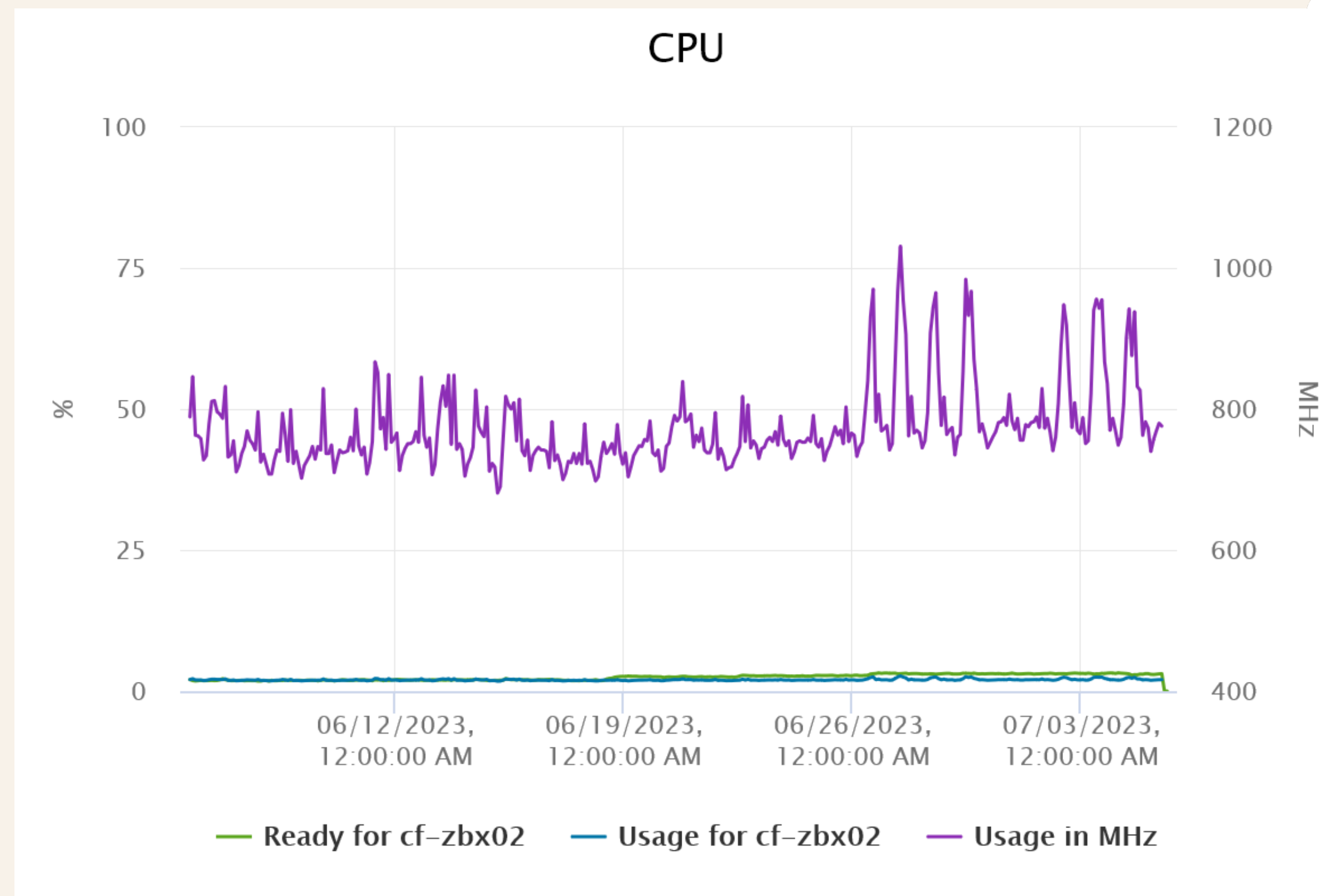


מכון
ויצמן
למדע

WEIZMANN
INSTITUTE
OF SCIENCE

Monitoring small HPC with Zabbix

Zabbix server - vCenter view (one month)



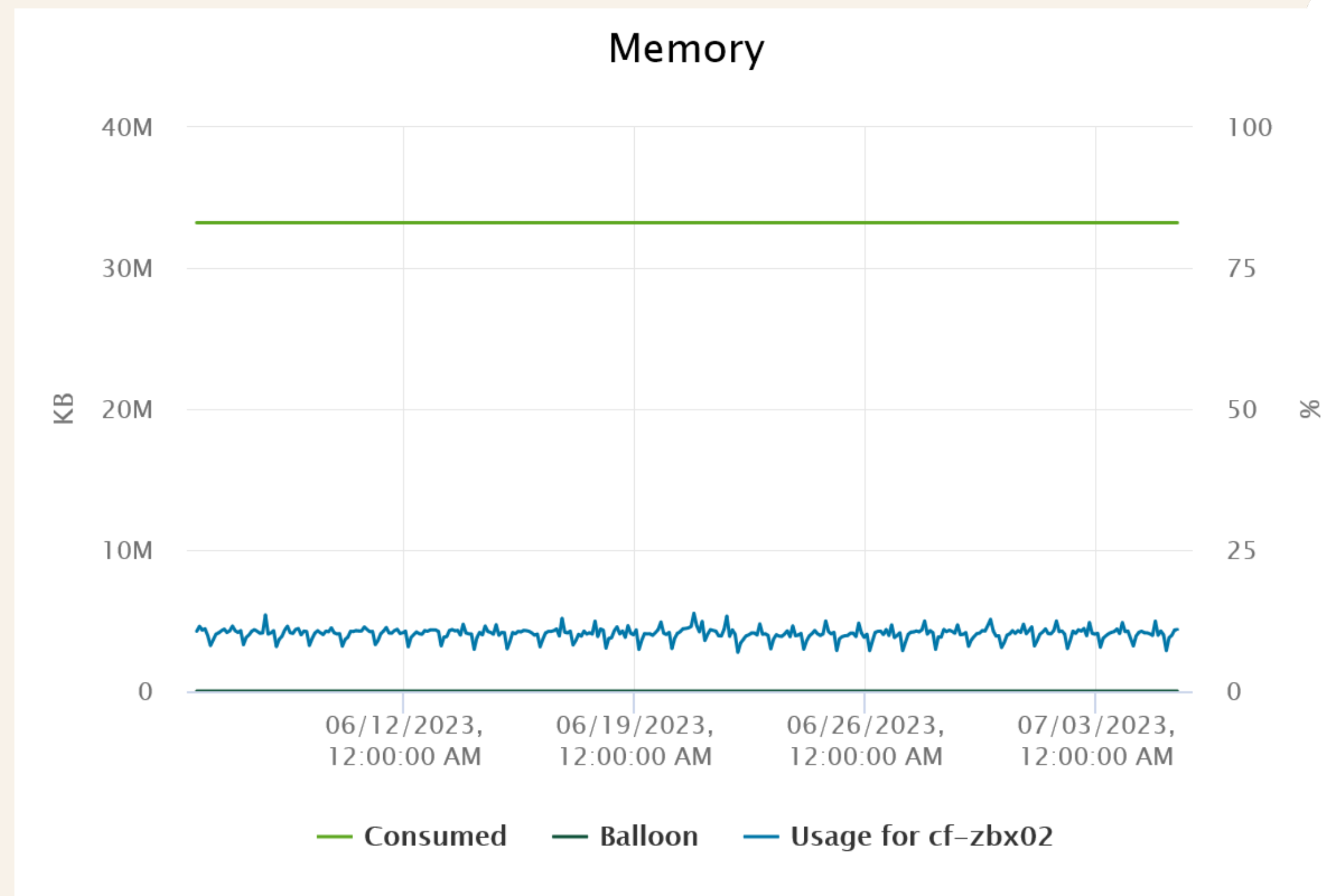


מכון
ויצמן
למדע

WEIZMANN
INSTITUTE
OF SCIENCE

Monitoring small HPC with Zabbix

Zabbix server - vCenter view (one month)



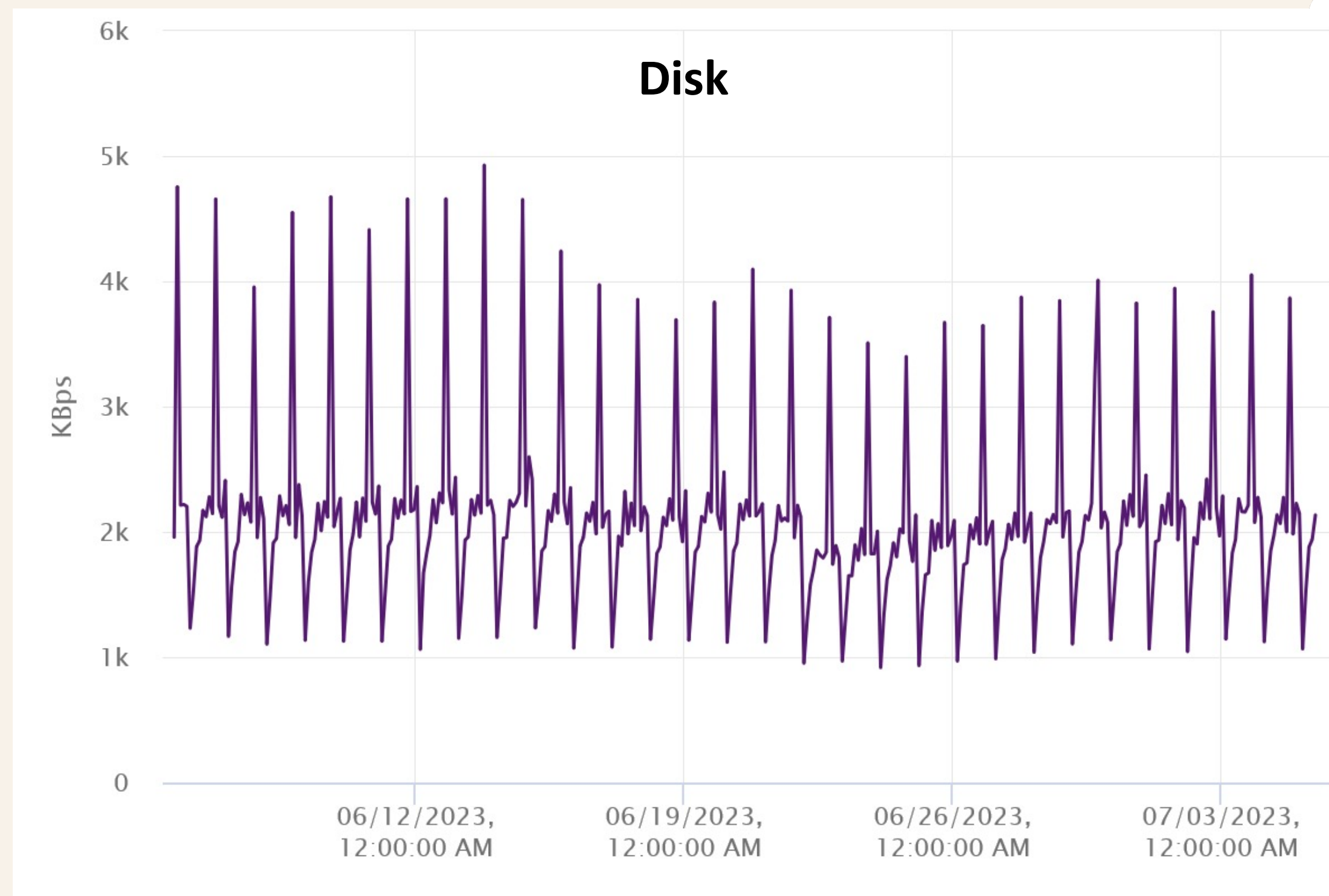


מכון
ויצמן
למדע

WEIZMANN
INSTITUTE
OF SCIENCE

Monitoring small HPC with Zabbix

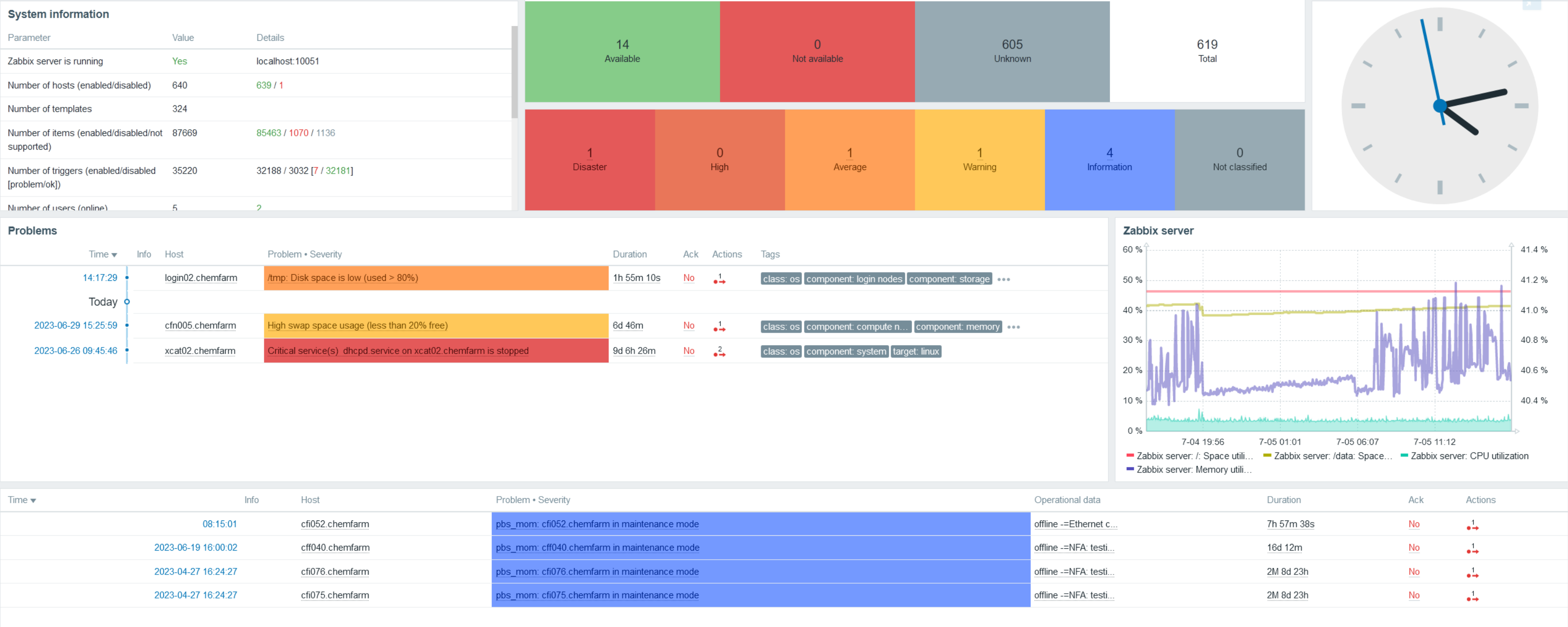
Zabbix server - vCenter view (one month)



How many metrics?

System information		
Parameter	Value	Details
Zabbix server is running	Yes	localhost:10051
Number of hosts (enabled/disabled)	640	639 / 1
Number of templates	324	
Number of items (enabled/disabled/not supported)	87678	85463 / 1070 / 1145
Number of triggers (enabled/disabled [problem/ok])	35224	32192 / 3032 [50 / 32142]
Number of users (online)	5	2
Required server performance, new values per second	828.42	

Main Dashboard





מכון
ויצמן
למדע

WEIZMANN
INSTITUTE
OF SCIENCE

Monitoring small HPC with Zabbix

External scripts and custom checks and actions

- PBS
 - Offline nodes with comment are set to “maintenance”
- GPFS
 - Node health
 - Connectivity
 - Waiters
- FlexLM vendor and daemon status on the license servers
- Specific services on all infrastructure servers
- Repair Ethernet module on some compute nodes



Huston! Do we have a problem?

Problems						
Time ▼	Info	Host	Problem • Severity	Duration	Ack	Actions
14:17:29	•	login02.chemfarm	/tmp: Disk space is low (used > 80%)	2h 23m 50s	No	1 →
Today	○					
2023-06-29 15:25:59	•	cfn005.chemfarm	High swap space usage (less than 20% free)	6d 1h 15m	No	1 →
2023-06-26 09:45:46	•	xcat02.chemfarm	Critical service(s) dhcpd.service on xcat02.chemfarm is stopped	9d 6h 55m	No	2 →



מכון
ויצמן
למדע

WEIZMANN
INSTITUTE
OF SCIENCE

Monitoring small HPC with Zabbix

Host groups

- Compute
 - CPU
 - GPU
- Ethernet switches
- Infrastructure:
 - Login nodes
 - PBS
 - GPFS, e.g. NSD
 - Name and licenses servers
 - Provisioning (xCAT)
 - Monitoring servers

ZABBIX <<

zbx02

- Monitoring
- Services
- Inventory
- Reports
- Configuration
- Host groups**

<input type="checkbox"/>	License nodes	Hosts 3
<input type="checkbox"/>	Infrastructure	Hosts 36
<input type="checkbox"/>	GPU nodes	Hosts 10
<input type="checkbox"/>	GPFS nodes	Hosts 15

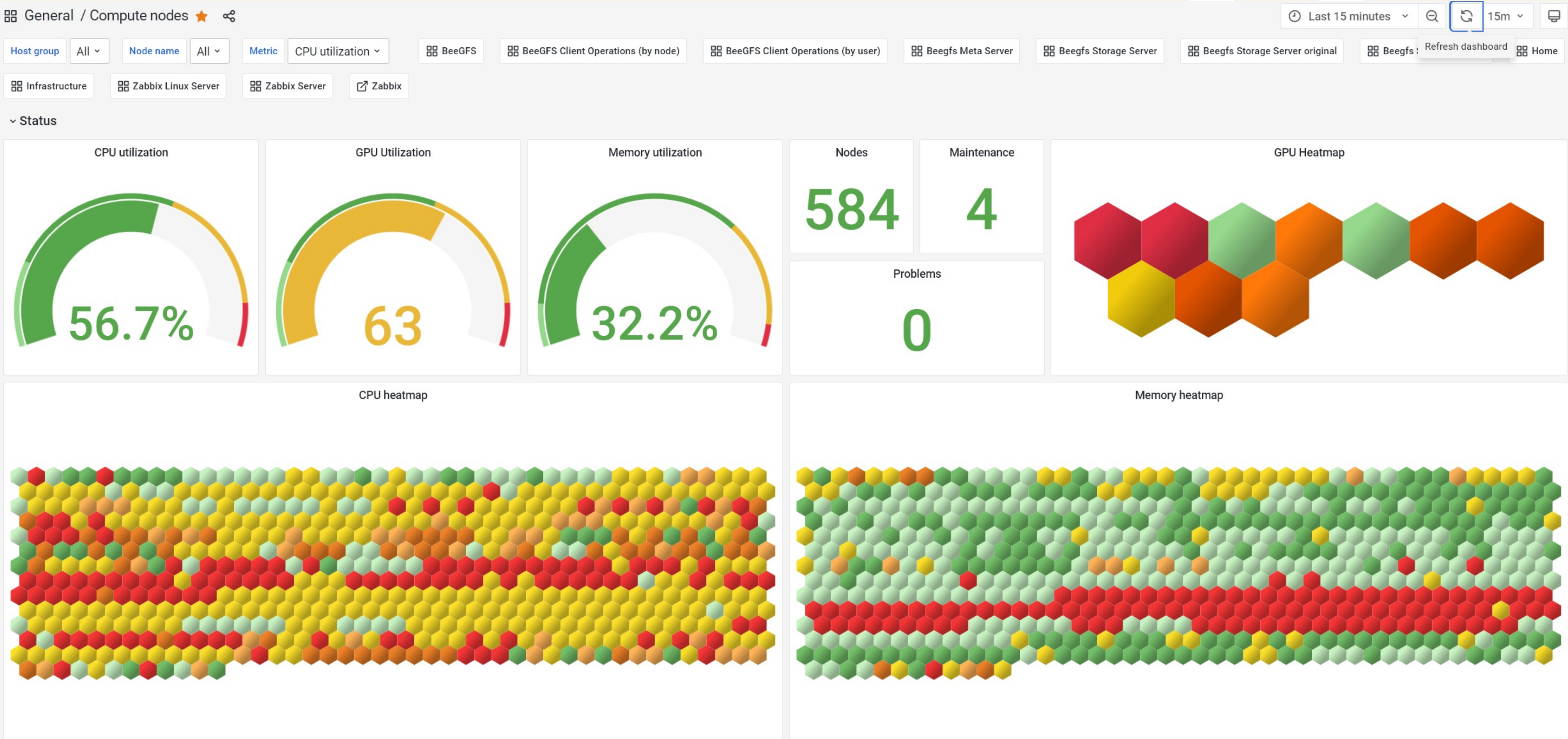
Media 1

Severity	Name	Operational data	Expression	Status
Average	Linux by Zabbix agent active: Load average is too high	Load averages(1m 5m 15m): {ITEM.LASTVALUE1} {ITEM.LASTVALUE3} {ITEM.LASTVALUE4}), # of CPUs: {ITEM.LASTVALUE2}	<code>min(/Compute node/system.cpu.load[all,avg1],5m)/last(/Compute node/system.cpu.num)>{\$LOAD_AVG_PER_CPU.MAX.WARN}</code> and <code>last(/Compute node/system.cpu.load[all,avg5])>0</code> and <code>last(/Compute node/system.cpu.load[all,avg15])>0</code>	Disabled
Average	Linux by Zabbix agent active: Lack of available memory	Available: {ITEM.LASTVALUE1}, total: {ITEM.LASTVALUE2}	<code>max(/Compute node/vm.memory.size[available],5m)<{\$MEMORY.AVAILABLE.MIN}</code> and <code>last(/Compute node/vm.memory.size[total])>0</code>	Disabled
Average	Linux by Zabbix agent active: High memory utilization Depends on: Compute node: Lack of available memory		<code>min(/Compute node/vm.memory.utilization,5m)>{\$MEMORY.UTIL.MAX}</code>	Disabled
Warning	Linux by Zabbix agent active: High CPU utilization Depends on: Compute node: Load average is too high	Current utilization: {ITEM.LASTVALUE1}	<code>min(/Compute node/system.cpu.util,5m)>{\$CPU.UTIL.CRIT}</code>	Disabled
Information	pbs_mom: {HOSTNAME} in maintenance mode		<code>find(/Compute node/pbs.status,20m,"like","-")=1</code> and <code>nodata(/Compute node/pbs.status,20m)=0</code>	Enabled
High	GPFS: Too long waiters on {HOSTNAME}		<code>last(/Compute node/gpfs.node.waiters.maxtime,#1)>0</code> and <code>nodata(/Compute node/gpfs.node.waiters.maxtime,61m)=0</code>	Enabled

- Why only “high” and “low”?
- Difference between
 - Compute – alive



A picture is worth a thousand words or, maybe \$\$\$...





מכון
והצמח
למדע

WEIZMANN
INSTITUTE
OF SCIENCE

Monitoring small HPC with Zabbix

Conclusions and Thanks



- Metrics outside Zabbix:
 - GPFS GUI
 - BeeGFS
 - UFM – Infiniband metrics
- Minor upgrade - smooth
- RHEL?
- Only two dashboards
- Not using reports
- Special thanks to Dr. Andrei Vasilev
- Questions



מכון
ויצמן
למדע

WEIZMANN
INSTITUTE
OF SCIENCE

Monitoring small HPC with Zabbix

Thank
you.