

advanced clustering
technologies, inc.

clustervisor

An Introduction to Version 1.x



advanced clustering
technologies, inc.

Version numbering

1.23.04-XXXX

Major version Release year Release month Build number

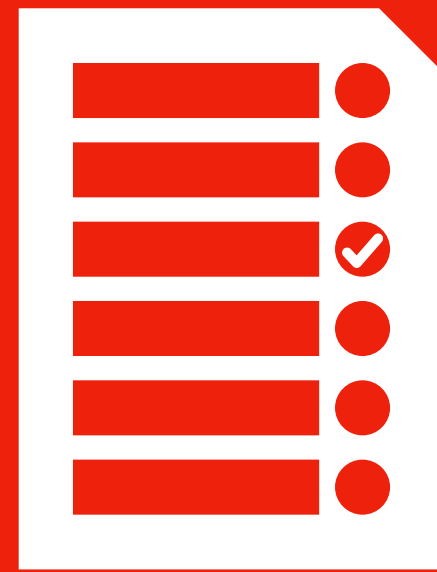
- Now much easier to determine if you are using latest version
- Compatibility between major equal major version numbers
- Easy dnf updates between releases of same major version
- Limit any breaking changes

What is ClusterVisor?

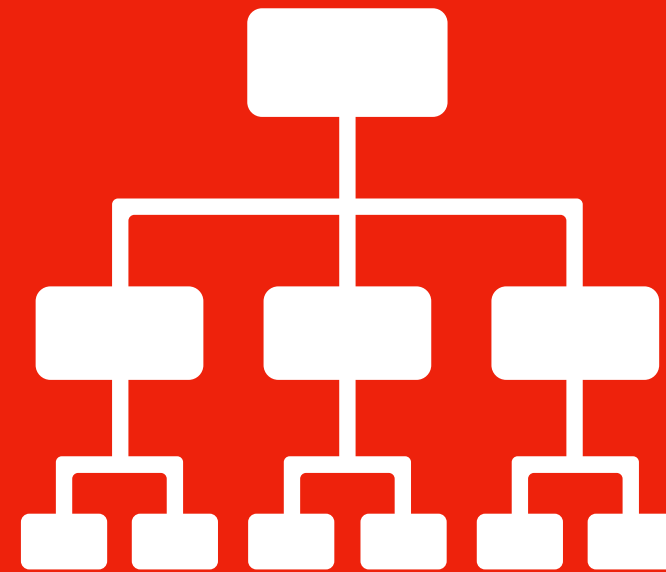
ClusterVisor provides easy to use interface to **deploy**, **provision**, **manage**, **monitor**, and **maintain** your cluster for its lifetime.



DEPLOY



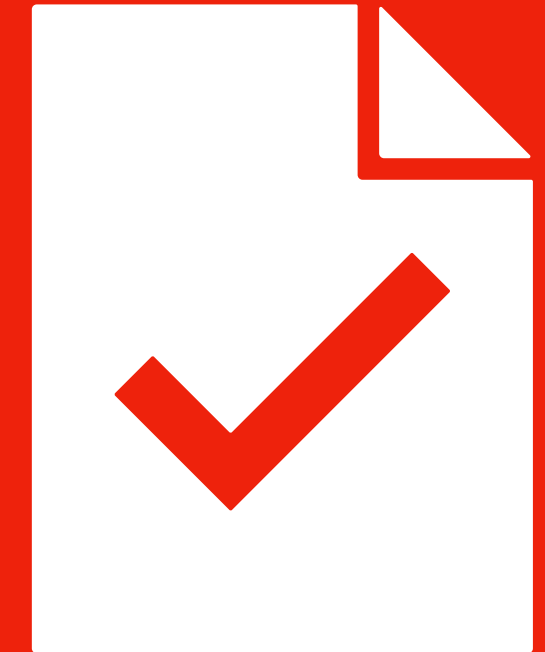
PROVISION



MANAGE



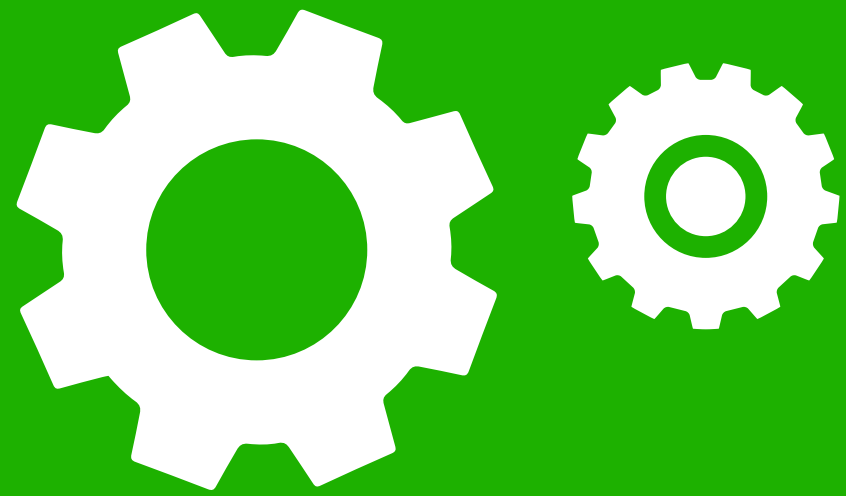
MONITOR



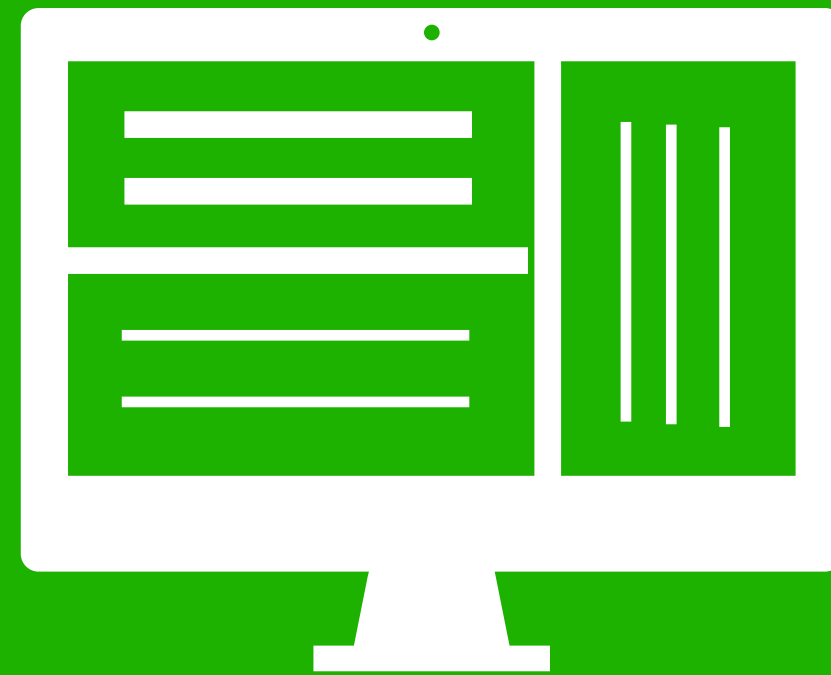
MAINTAIN

What's New?

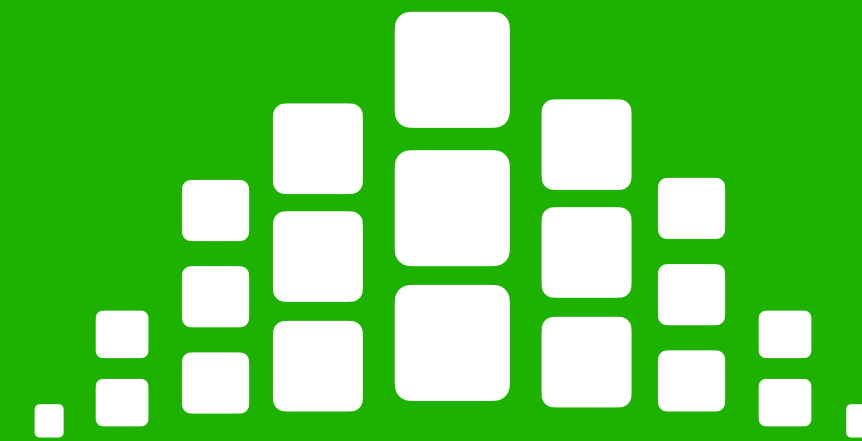
ClusterVisor 1.0 is a significant upgrade from the original version



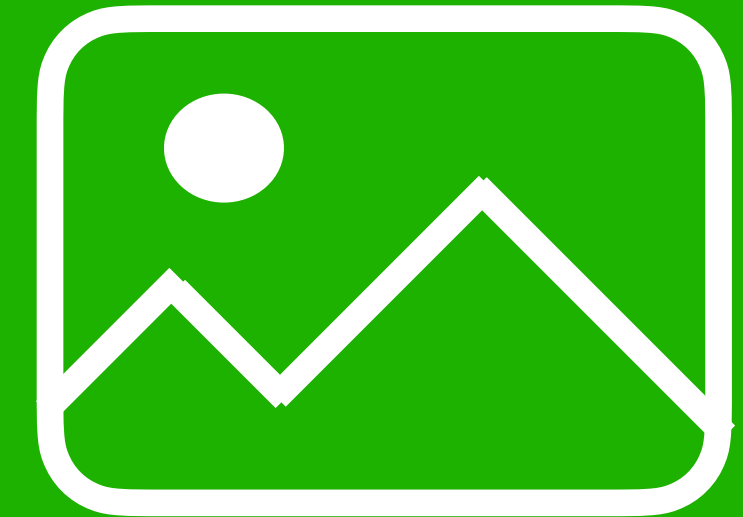
**A NEW AND
IMPROVED STATS
ENGINE**



**A REVAMPED
WEB UI**

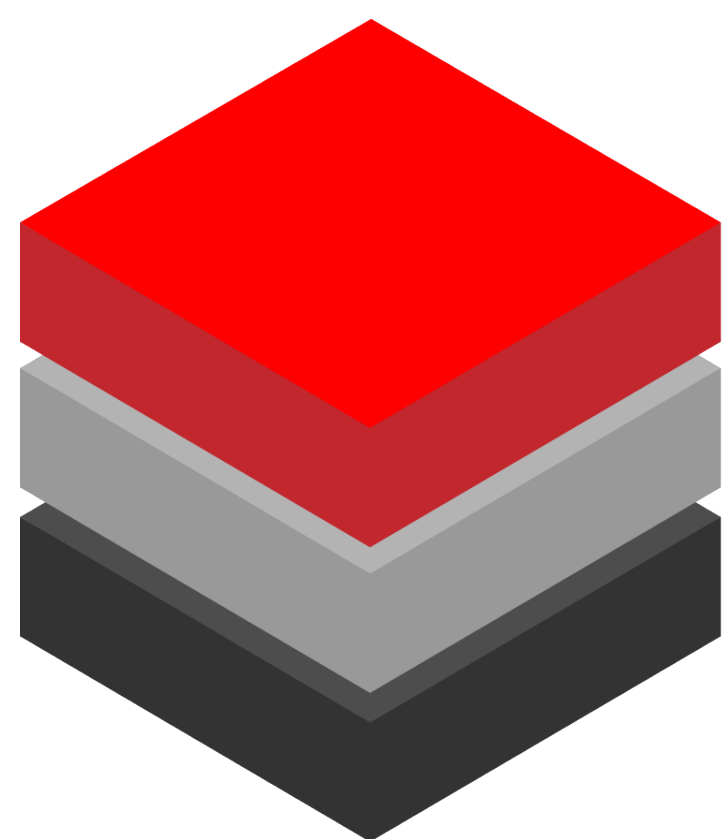


**INTEGRATION
WITH SLURM**

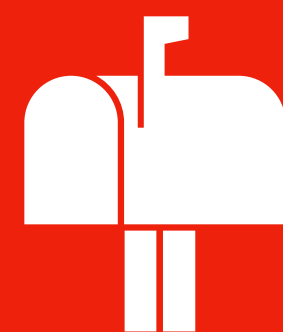


**NODE
PROVISIONING
ENHANCEMENTS**

And many, many more



clustervisor



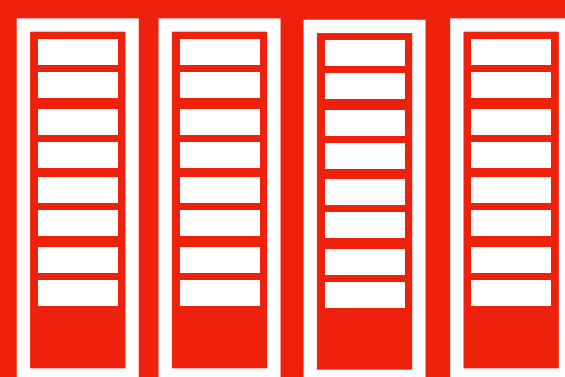
**FLEXIBLE
DELIVERY**



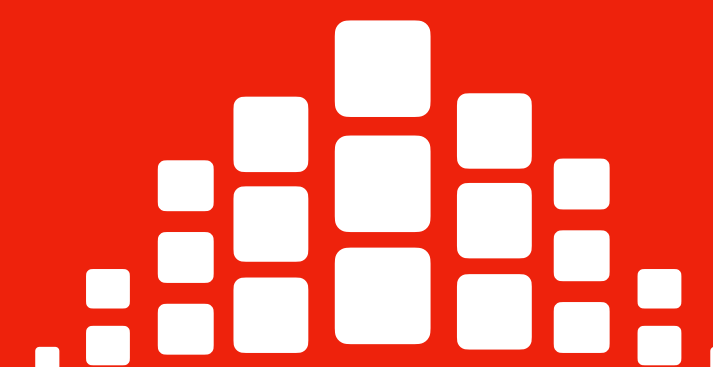
**CONFIGURATION
MANAGEMENT**



**STATISTICS &
MONITORING**



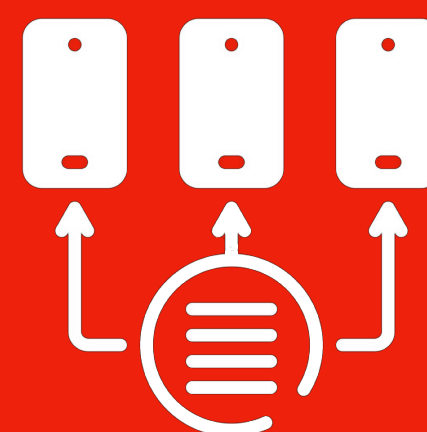
**RACK
DIAGRAMMING**



**INTEGRATION
WITH SLURM**



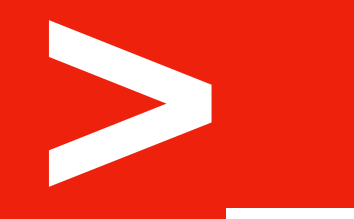
**CUSTOMIZABLE
DASHBOARDS**



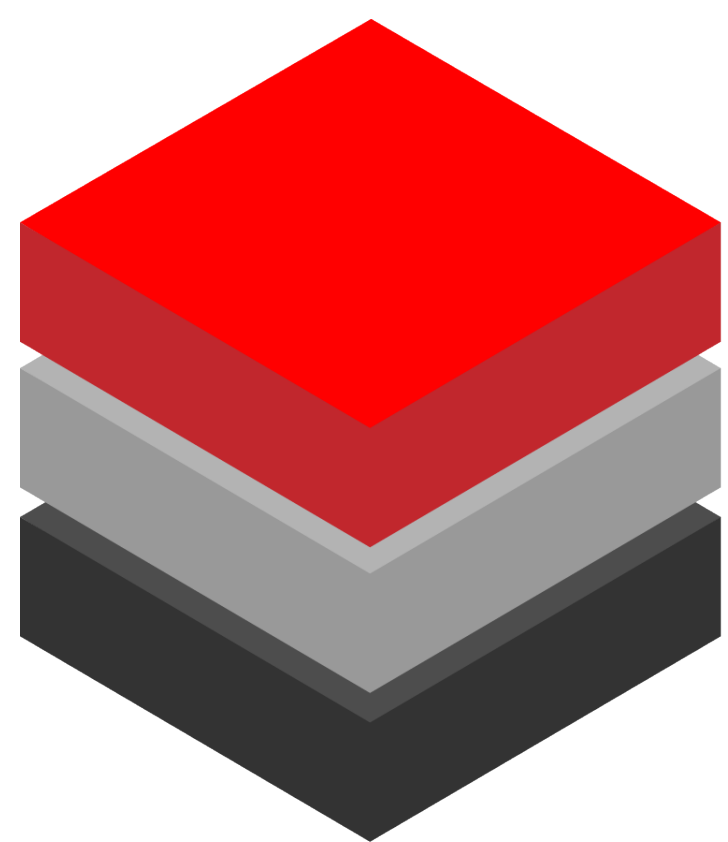
PROVISIONING



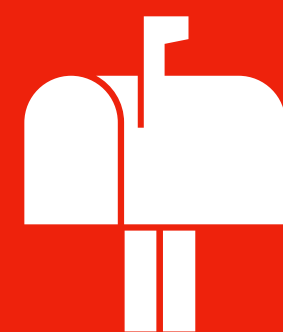
**USER
MANAGEMENT**



**COMMAND LINE
TOOLS**



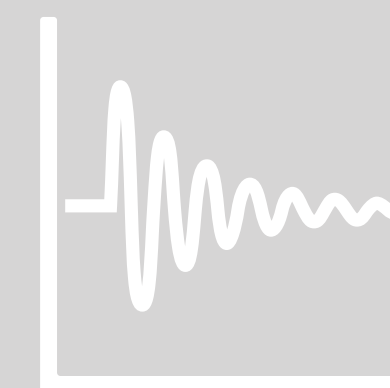
clustervisor



**FLEXIBLE
DELIVERY**



**CONFIGURATION
MANAGEMENT**



**STATISTICS &
MONITORING**



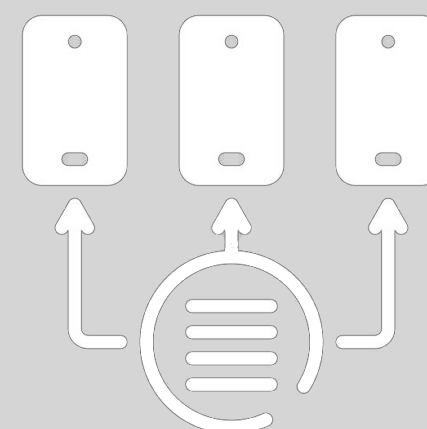
**RACK
DIAGRAMMING**



**INTEGRATION
WITH SLURM**



**CUSTOMIZABLE
DASHBOARDS**



PROVISIONING



**USER
MANAGEMENT**



**COMMAND LINE
TOOLS**

Delivery Options

Installed on head node

- Traditional approach
- Head node used for managing cluster and user logins

Dedicated ClusterVisor appliance node

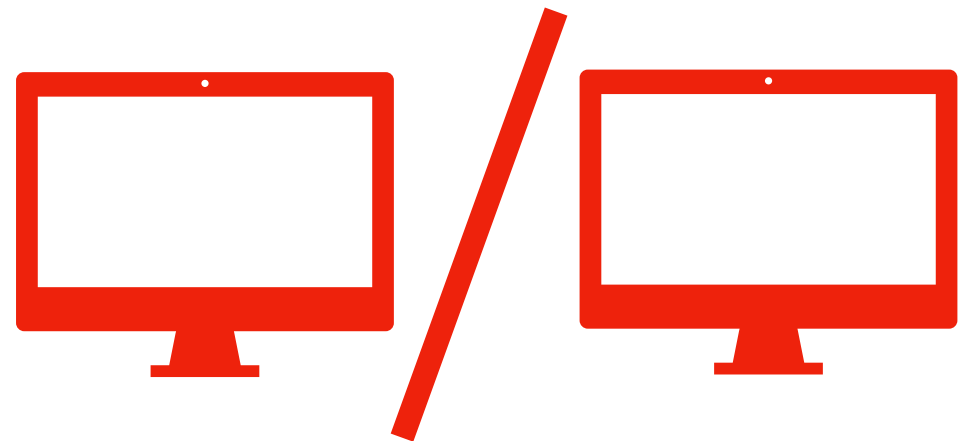
- Dedicated system for managing cluster
- Provides easy disaster recovery
- Simple to manage and update

Why an appliance?



DISASTER RECOVERY

The appliance holds images for all nodes including compute, login, storage, etc. Makes it easy to restore any node (including infrastructure ones) in your cluster when something goes wrong.



PROCESS ISOLATION

Separate the functions of cluster management, and user logins. Appliance runs limited shell and regular users don't have access. Limits issues created by users doing inappropriate things on the head node.



MAINTENANCE

Based on standardized hardware and software configuration. Makes upgrades, and hardware maintenance easy. If the appliance fails, remove hard drives and install them into a replacement system, and you are back up and running again.

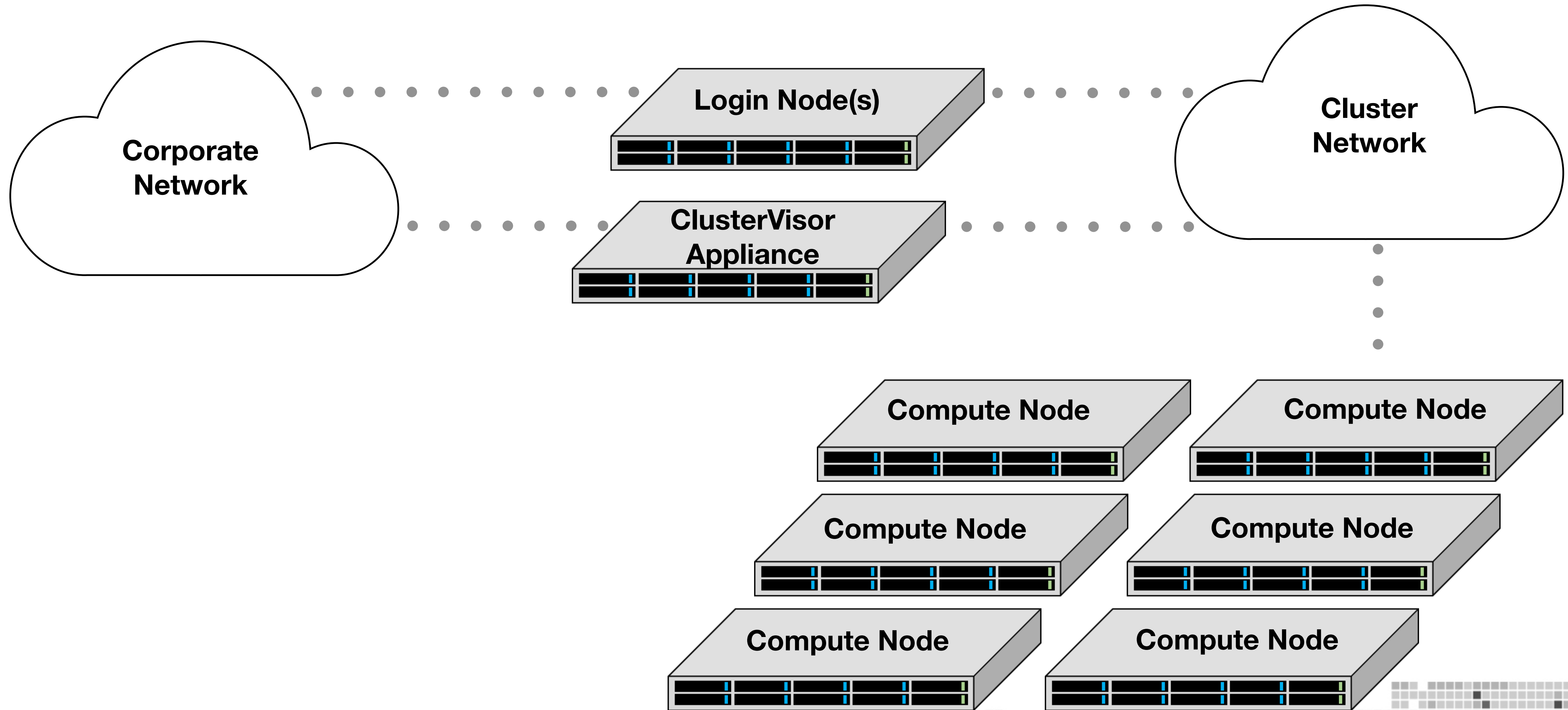
ClusterVisor Appliance

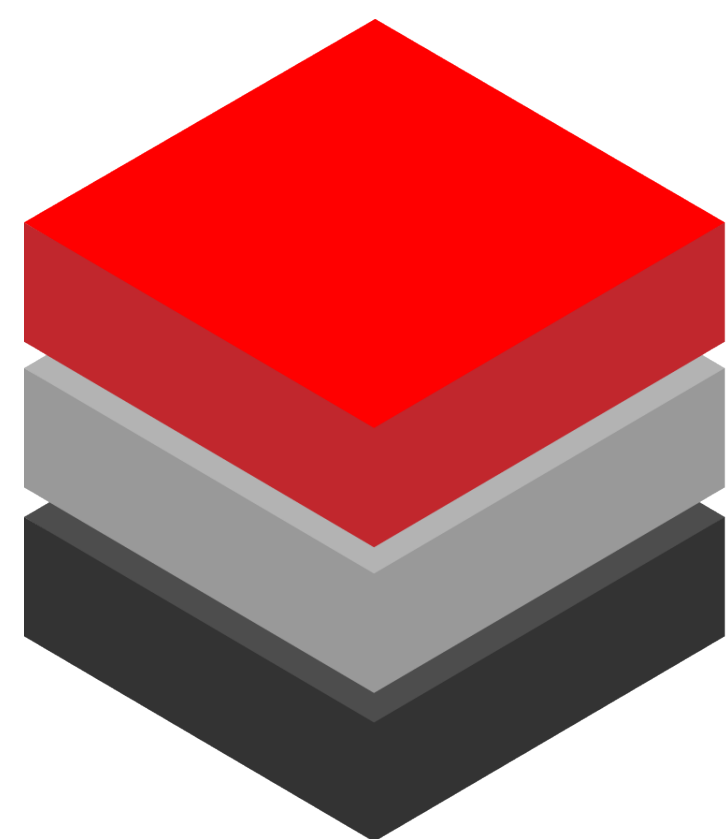
- Dedicated 1U server with redundant power, optimized for running ClusterVisor on systems with up to 100s of nodes
- Runs a minimal locked down version Rocky Linux 9.x with a dedicated management Web GUI.
- All user customer specific data stored on removable drives in the front of the chassis
- Much lower cost compared to standard head nodes

The ClusterVisor appliance runs all the essential services needed to make your cluster function: DHCP, DNS, PXE, LDAP, Node Provision, Monitoring and Alerting



Network design

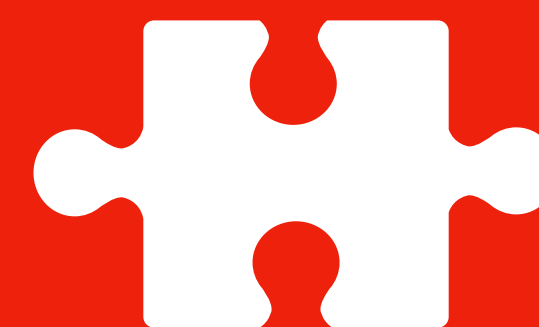




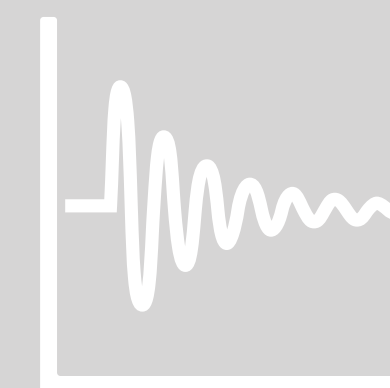
clustervisor



FLEXIBLE
DELIVERY



CONFIGURATION
MANAGEMENT



STATISTICS &
MONITORING



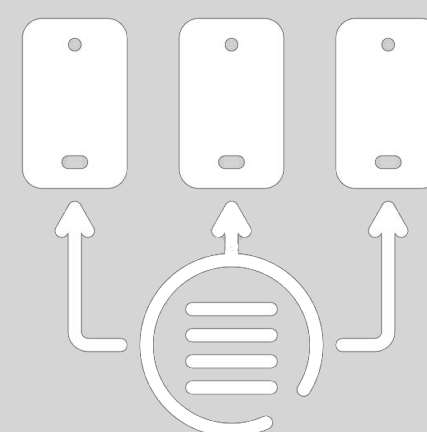
RACK
DIAGRAMMING



INTEGRATION
WITH SLURM



CUSTOMIZABLE
DASHBOARDS



PROVISIONING



USER
MANAGEMENT



COMMAND LINE
TOOLS

Configuring Nodes, etc.

ClusterVisor's powerful **Configuration** tool stores your entire cluster's configuration, making it possible to manage every part of the cluster from within this interface.

Configuration

[Commit changes](#)[Change queue](#)[Reload](#)

Appliance Configuration **Node** Netboot Cloner Image Cloner Disk Cloner BIOS Diskless Image Switch Misc Device Chassis Synced File Group Group Stat
Computed Stat Monitoring Rule Monitoring Action

[Add Node](#)[Add Node from template](#) ▼

Node name	Firewall	Cloner Image	Actions
ks-slurm-head	false		Edit ⋮
node[01-04]		cloner_image.node	Edit ⋮
node[05-60]			Edit ⋮

Configuring Nodes, etc.

- Everything is organized as parts of a collection. At right is the full list of collection types and a description of each.
- Each collection has unique schema of fields that store data pertinent to that type
- Collections are not only hardware elements in the cluster but configuration elements as well

Collection	Description
config	Global configuration variables and settings
appliance	ClusterVisor appliance nodes
node	Nodes in your cluster (compute, login, head, etc.)
netboot	iPXE network boot options
cloner_image	A cloner image to deploy to nodes
cloner_disk_layout	Partition, format, RAID options for disks to be used during cloning
switch	Network switch devices
chassis	An enclosure for multi-node blade style systems
device	Any other device you'd like to include in your cluster
group	A collection of other devices (usually nodes)
monitoring_rule	A set of tests to apply to device stats
monitoring_action	Actions to perform when monitoring rules are true (email, scripts, etc.)
computed_stat	Take existing stats and perform additional operations on them
group_stat	Rollup stats for all members of a group



Configuration

- Configuration elements can be managed via the command line or the Web GUI.
- The Web GUI is easier for discovery to see all the possible fields and values available
- The command line can be faster to edit once you are familiar with the syntax and options
- When editing via the command line, the element's entire config will be present to you as YAML

Configuration Plugins

- Certain collections support plugins (currently node and appliance)
- Plugins are used to take data stored in ClusterVisor and write that configuration to nodes.
- For example, the networking plugin takes all the networking configuration for the nodes and configures those interfaces for you
- Plugins can be enabled or disabled depending on if you want ClusterVisor to manage that part of the configuration for you.
- For example, it may make sense to disable the chrony plugin if you have a very advanced time server setup and would prefer to manage that configuration on your own.
- Certain node types will have different plugins enabled (compute nodes will not need all the plugins as a head or login node)
- ***If no plugins are enabled on nodes, no configuration changes will ever be done.***

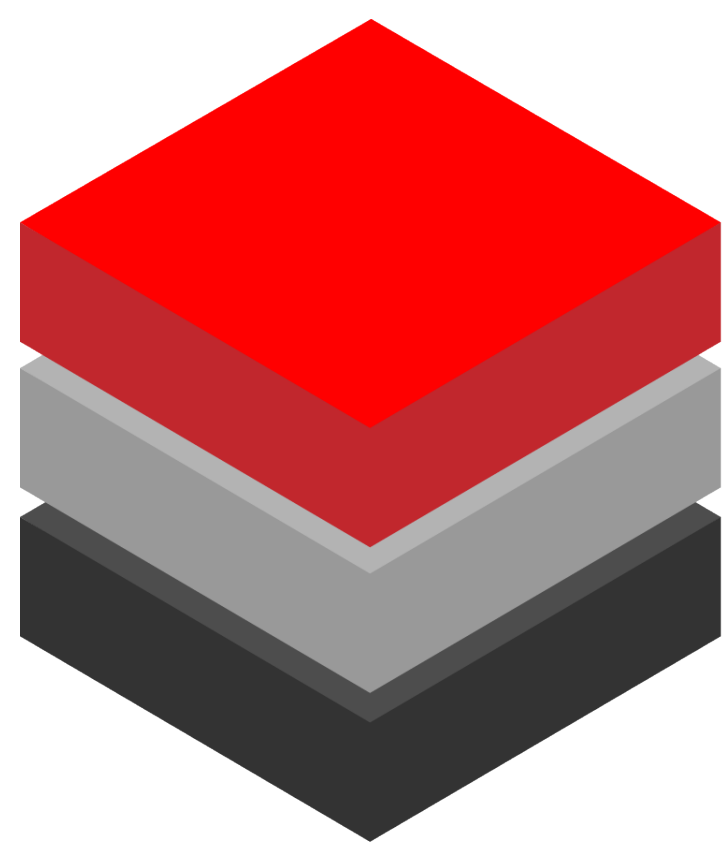
Plugins

Collection	Description
chronyd	Setup/configure the chrony daemon to synchronize a system's date/time
cloner_server	Manage the cloner server (rsyncd)
dhcp_dns_server	Manage the dnsmasq DHCP/DNS server
efiboot	Set the EFI boot order of the system
filesync	Synchronizes files to other nodes in the cluster
fstab	Write the extra mount entries into the fstab on a node
hosts	Create the /etc/hosts file on the node
ipmisettings	Configure the IPMI interface and/or authentication on the node
ldapauth	Configure LDAP authentication on the node
networking	Create all the networking configuration files on the node
nfsserver	Configure the NFS exports for this node
resolv_conf	Set the nameserver in the /etc/resolv.conf for a node
rsyslog	Configure rsyslog on the node
slurmclient	Manage the SLURM scheduling client on a node
slurmserver	Manage the SLURM scheduling server on a management node
ssh	Manage the /etc/ssh/ssh_known_hosts file on a node



Plugins

Collection	Description
sudoers	Handle the sudoers settings on the node
timezone	Set a node's timezone
yumrepo	Manage the yum repositories on a node
yumrepo_server	Run a local yum repository on the node
limits	Set memlock and stack limits needed for IB networks
lmod	Setup and configure LMod
selinux	Configure Security Enhanced Linux (SELinux)
serialconsole	Configure Linux's serial console redirection
bootoptions	Setup various boot configuration options
tuned	Set the node's tuned profile



clustervisor



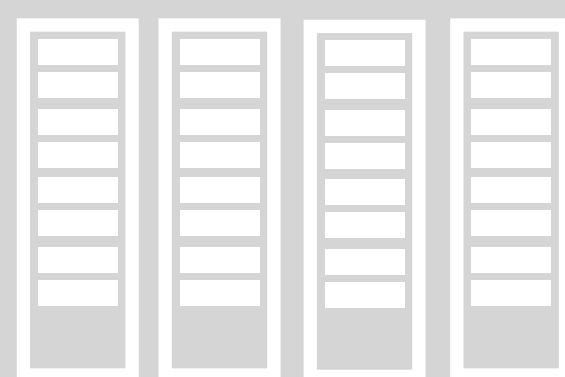
FLEXIBLE
DELIVERY



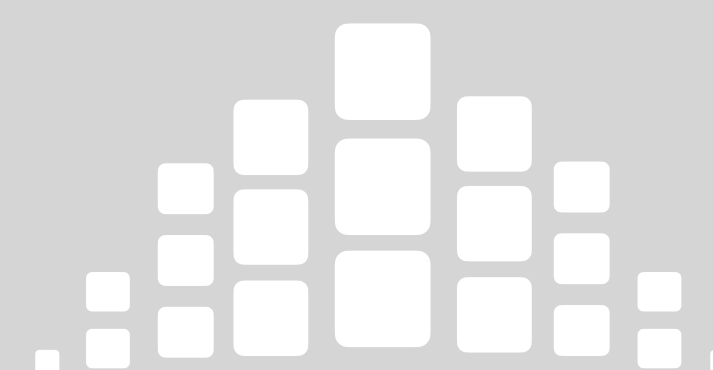
CONFIGURATION
MANAGEMENT



STATISTICS &
MONITORING



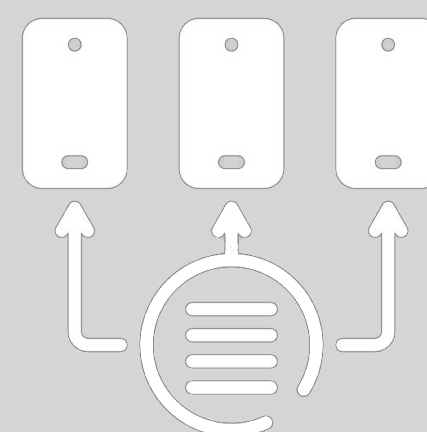
RACK
DIAGRAMMING



INTEGRATION
WITH SLURM



CUSTOMIZABLE
DASHBOARDS



PROVISIONING



USER
MANAGEMENT



COMMAND LINE
TOOLS



Stat/monitoring system

- The entire stats and monitoring subsystem was totally revamped in the 1.0 release
- Fixes many bugs and offers a lot more features
 - Better retention and rollover of older stats
 - Easy to view history and of stats per node, group, by job, etc
- Monitoring rules much more powerful with a full logic engine

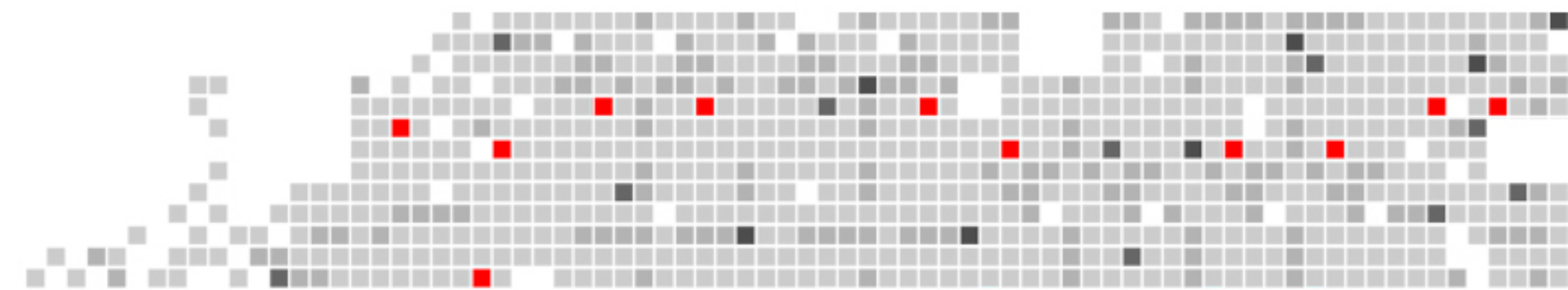


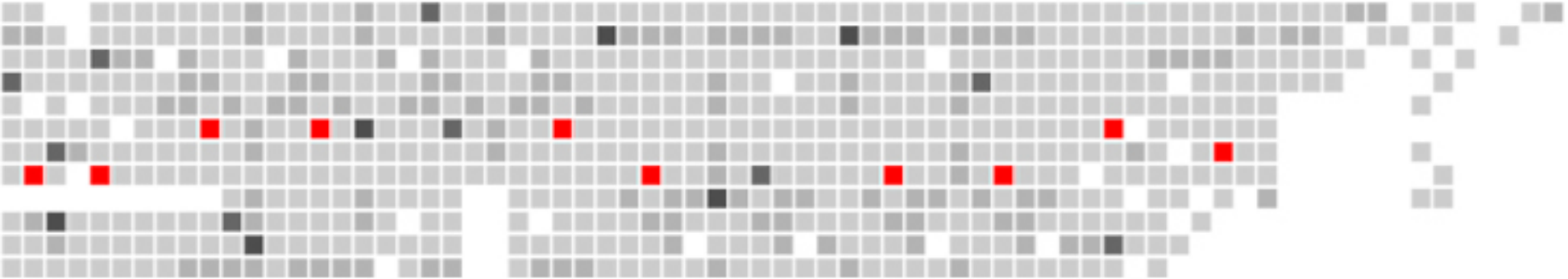
Stats and monitoring

- Stats are collected on the node via a plugin architecture
- Any stat collected can be used by the monitoring rules engine and perform actions
- Rule example: node is down, temperature is too high, RAID array degraded, etc.
- Action examples: send emails, run scripts
- You can view history
- You can compare stats among nodes

Stat history

interval	retention period
30 seconds	last 2 weeks
5 minutes	last 2 months
15 minutes	last 3 months
1 hour	last 5 years





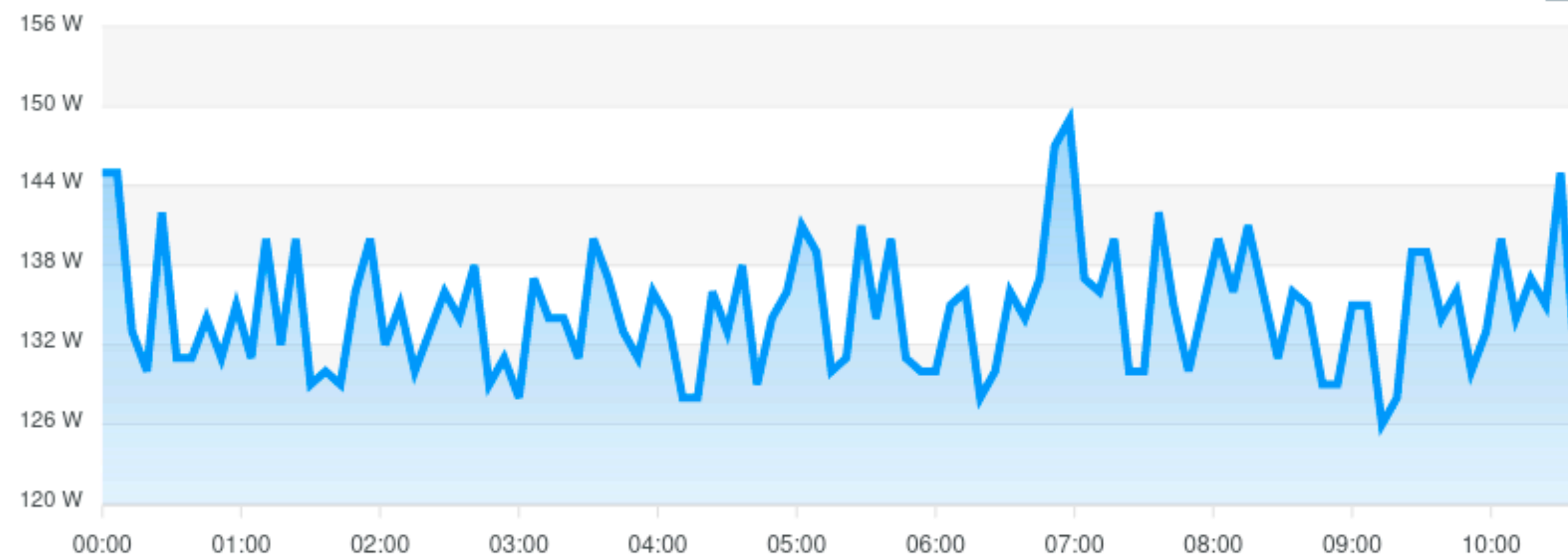
Stat plugins

Collection	Description
cpu	CPU usage and load information
cvclient	ClusterVisor version installed on the node
disks	Details about each disk and filesystem
firmware	BIOS and BMC firmware versions
infiniband	InfiniBand or Omni-Path fabric device performance and error counters
intelssd	Health info specific to intel branded datacenter SSDs
ipmi	Temperature, voltage, fans from IPMI
md	Software RAID status including if any arrays are degraded
megaraid	Hardware RAID status from LSI/Avago/Broadcom RAID controllers
mem	Memory utilization and ECC error counts
net	Network stats by Ethernet device
nfs	Status of NFS mounts and exports
ntp	Date / time info
nvidia	GPU information including power, ECC, utilization, memory
power	Power consumption on the node
system	Kernel version, uptime, etc.

Stats and Monitoring

ClusterVisor has a full statistics and monitoring engine.

CPU Power Consumption

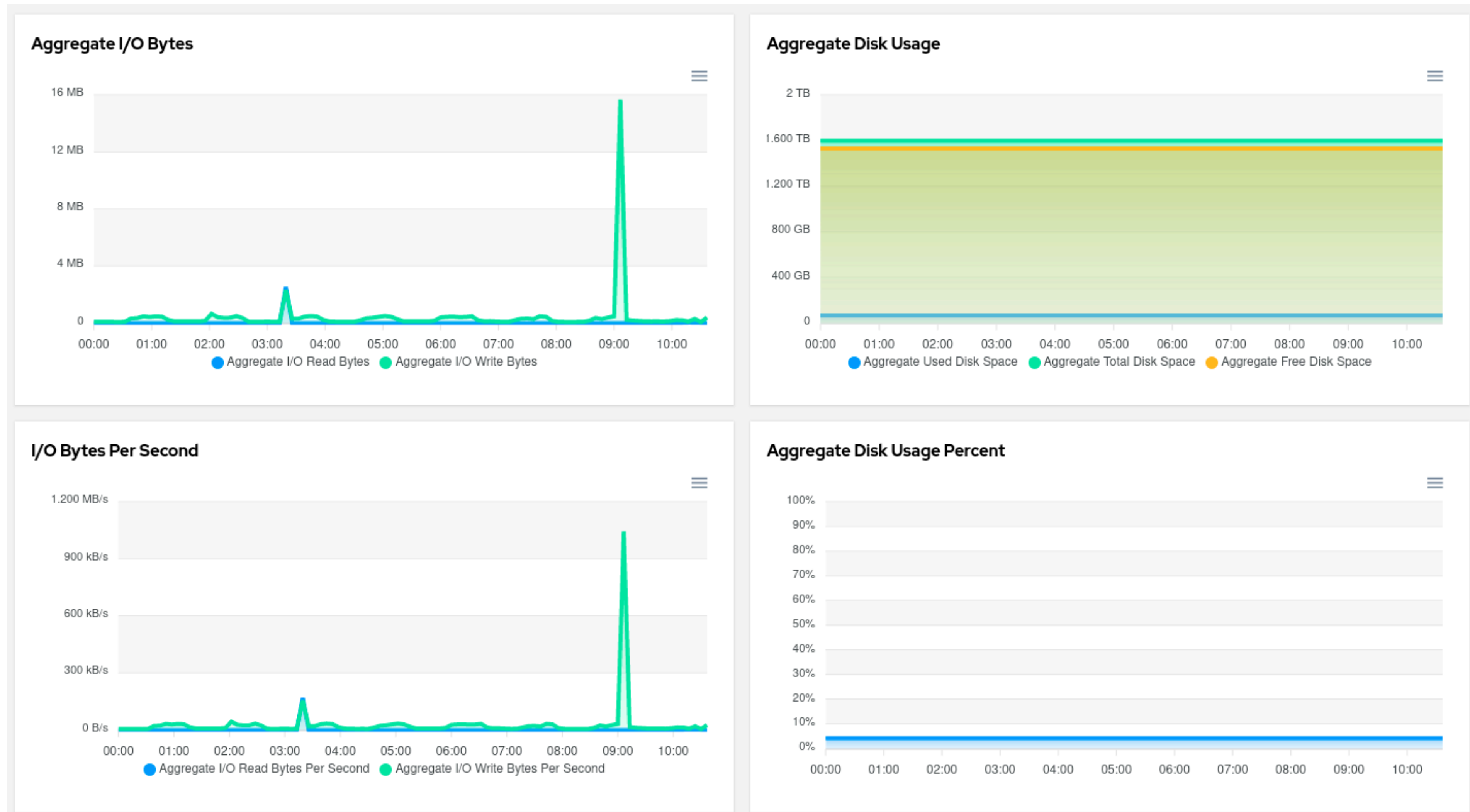


Node Power Consumption



Stats and Monitoring

ClusterVisor has a full statistics and monitoring engine.



Stats and Monitoring

ClusterVisor has a full statistics and monitoring engine.



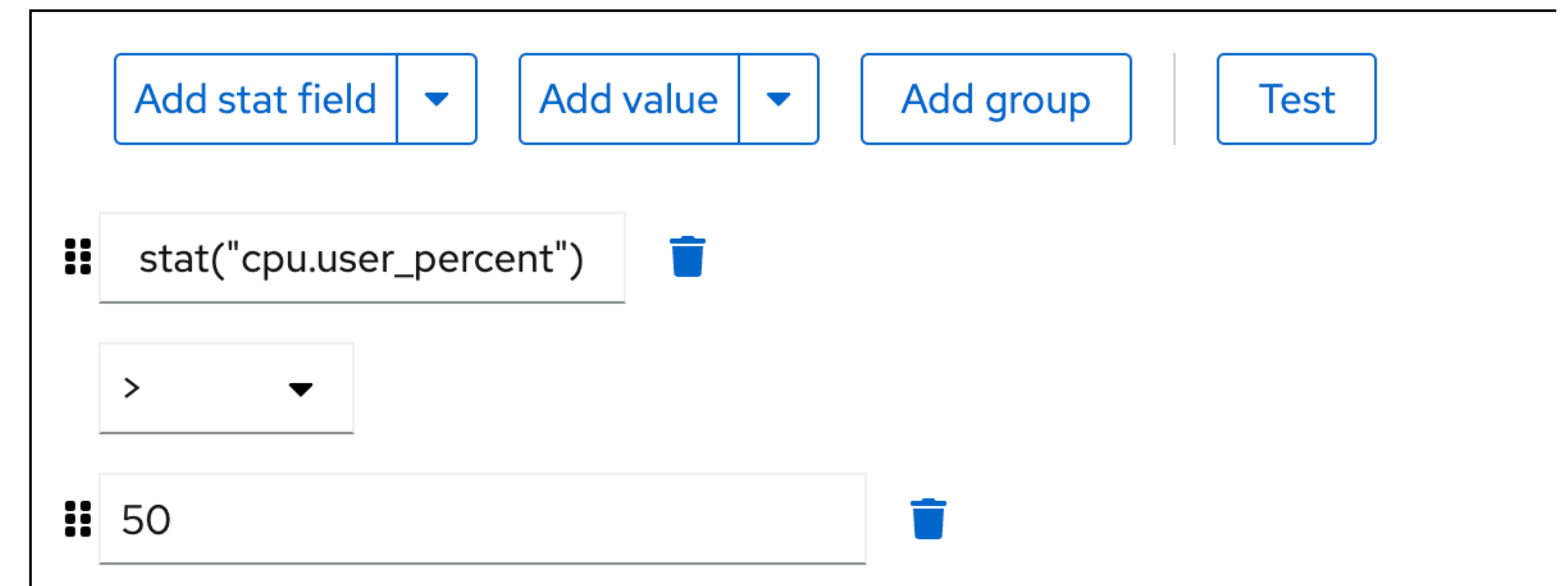
Cluster Health Checks

- One of ClusterVisor's most important features is its ability to act as a health monitor for all aspects of your cluster via our monitoring rule system
- Any stat collected can be turned into a monitoring rule. Examples:
 - Node not responding, InfiniBand down, RAID arrays degraded, ECC errors, InfiniBand down, Low disk space, etc.



Monitoring rules

- Templates for many common rules are available to make adding rules for your system easier
- Rules are run at 30 second intervals; rules can be defined to run once, always, or on clear
- The monitoring rule can be as simple or as complicated as you like. The Web GUI has a full GUI-based rule editor
- Customizable actions including emails, or running scripts on your system



The screenshot shows a web-based rule editor interface. At the top, there are four buttons: "Add stat field" with a dropdown arrow, "Add value" with a dropdown arrow, "Add group", and "Test". Below these buttons, the rule is defined in two parts. The first part consists of a list icon, a text box containing "stat('cpu.user_percent')", and a trash icon. Below this is a comparison operator dropdown menu showing ">". The second part consists of a list icon, a text box containing "50", and a trash icon.

Rule History

- Full queryable history of all monitoring rules
- Stat snapshot taking of all values when rule fails
- Helpful for diagnosing problems on your system
- Can acknowledge rules, so you won't get notified in the future

Monitoring rules: History

Triggered rules between 04/21/2023 00:00 and

Now

Search...

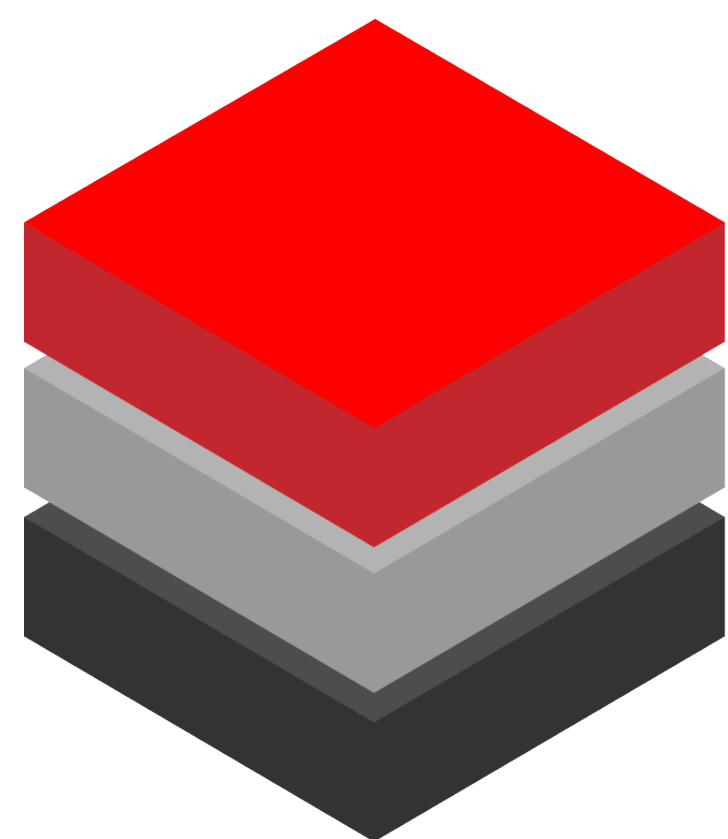
1 - 17 of 17

Start ↑	End ↓	Duration ↓	Start st... ↓	End state ↓	Device ↓	Rule ↓	Descrip... ↓
04/21/2023 14:10		3s	Failed	Still active	node.node02	node_down	
04/21/2023 14:10		3s	Failed	Still active	node.node03	node_down	
04/21/2023 14:10		3s	Failed	Still active	node.node06	node_down	
04/21/2023 14:10		3s	Failed	Still active	node.node10	node_down	
04/21/2023 14:10		3s	Failed	Still active	node.node14	node_down	
04/21/2023 14:10		3s	Failed	Still active	node.node18	node_down	
04/21/2023 14:10		3s	Failed	Still active	node.node22	node_down	
04/21/2023 14:10		3s	Failed	Still active	node.node26	node_down	
04/21/2023 14:10		3s	Failed	Still active	node.node30	node_down	
04/21/2023 14:10		3s	Failed	Still active	node.node33	node_down	



Coming soon

- Ability to query and collect stats for non-node devices
- Currently planned:
 - PDUs, UPSs, Switches, Storage arrays
- Command line tools and API so you can inject your own stats into ClusterVisor



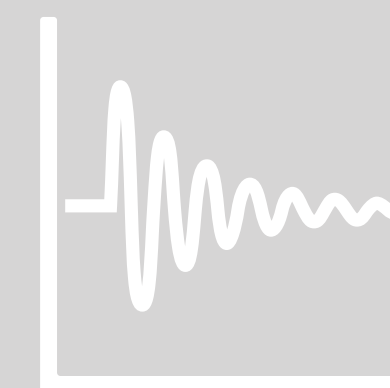
clustervisor



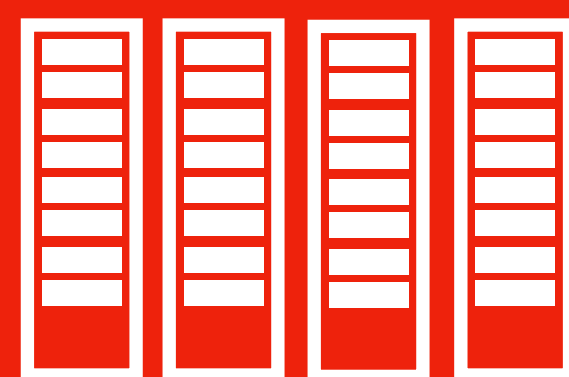
FLEXIBLE
DELIVERY



CONFIGURATION
MANAGEMENT



STATISTICS &
MONITORING



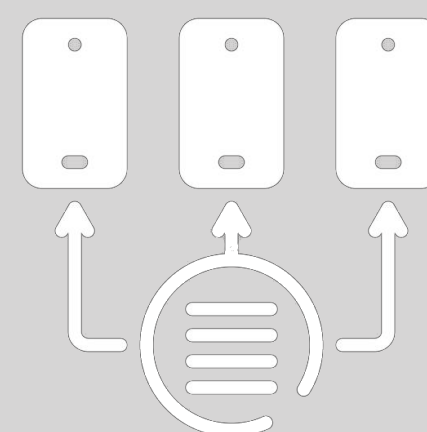
RACK
DIAGRAMMING



INTEGRATION
WITH SLURM



CUSTOMIZABLE
DASHBOARDS



PROVISIONING



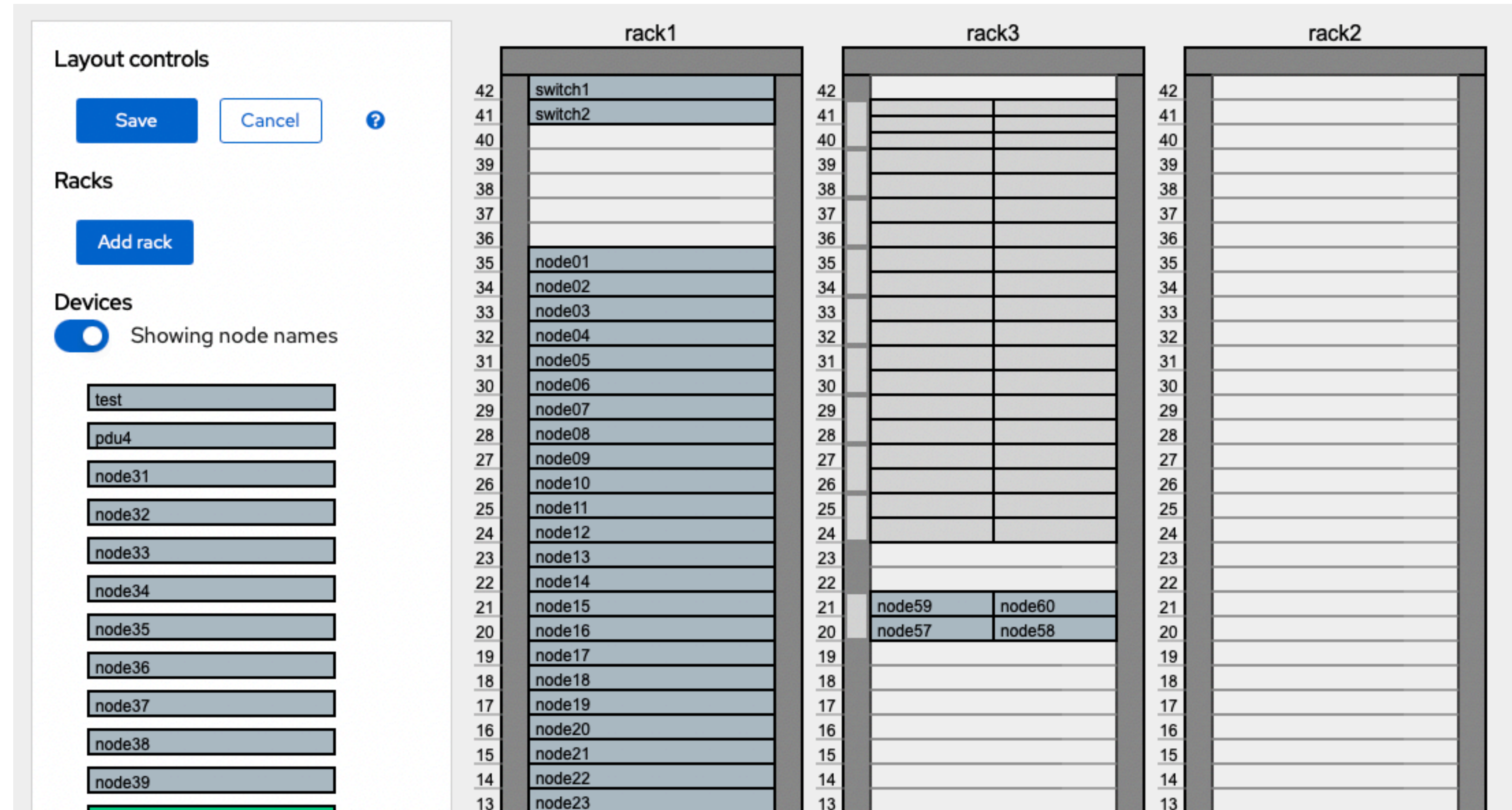
USER
MANAGEMENT



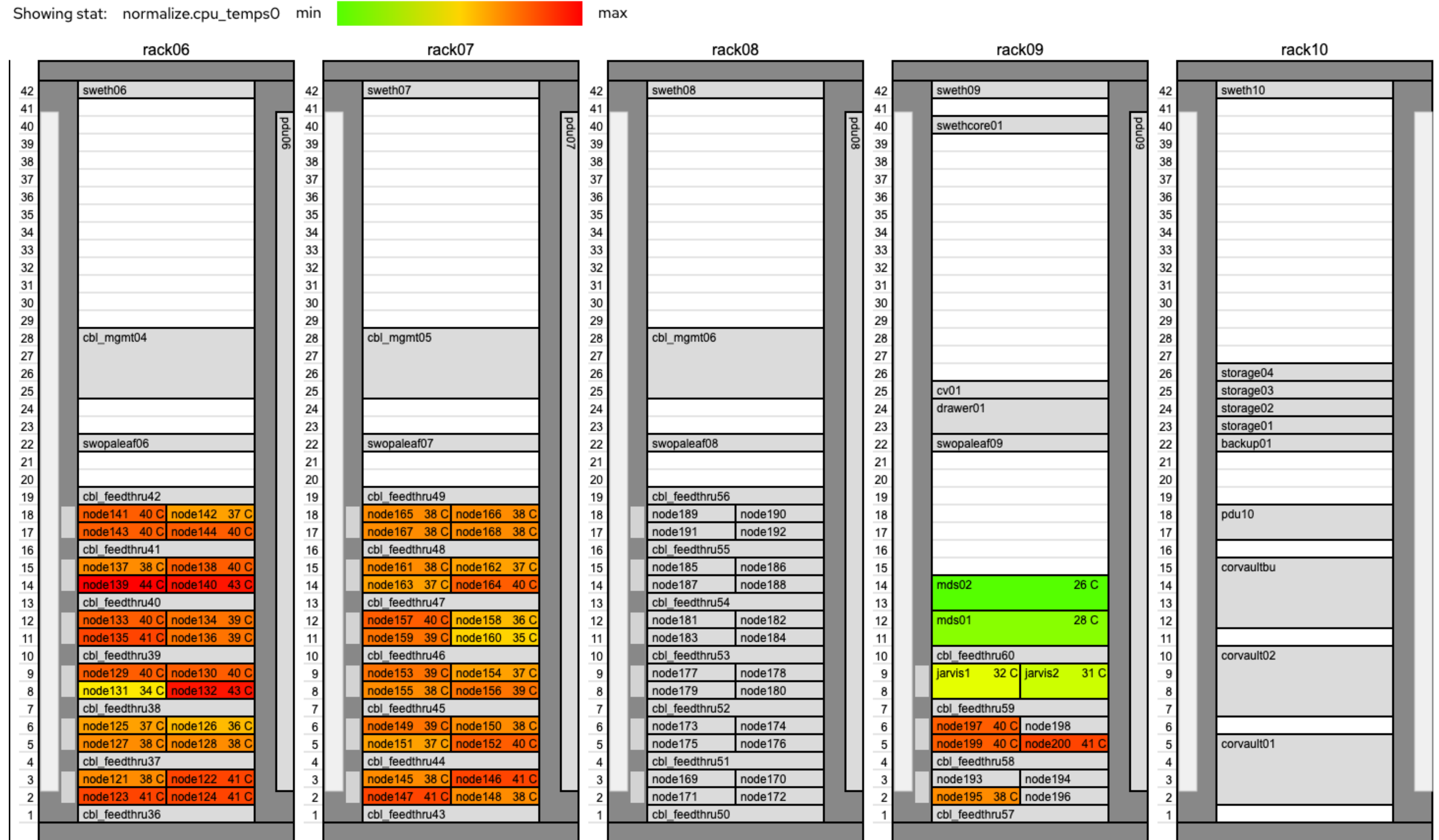
COMMAND LINE
TOOLS

Rack Diagrams

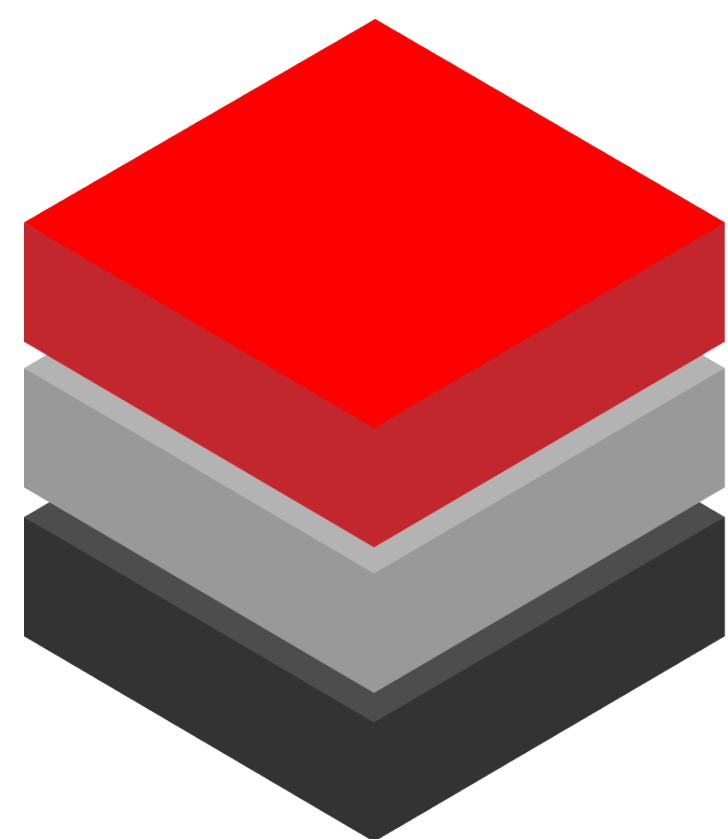
- Old versions of ClusterVisor required editing text values to create diagrams.
- With the 1.x release, rack editing is all drag and drop



Rack Diagrams



- Stats and monitoring rules can be applied to rack diagrams to create visual heatmaps
- Multiple racks and rows can be created to simulate the real world setup of your datacenter



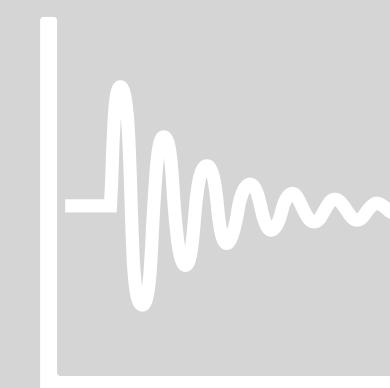
clustervisor



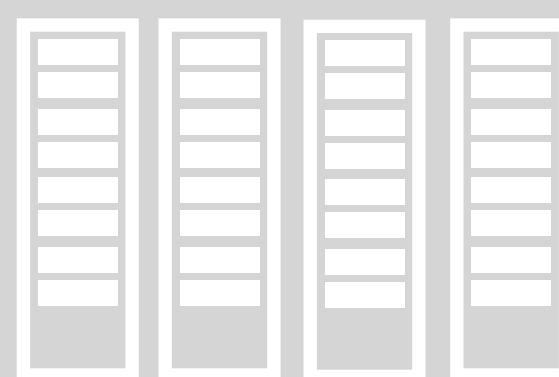
FLEXIBLE
DELIVERY



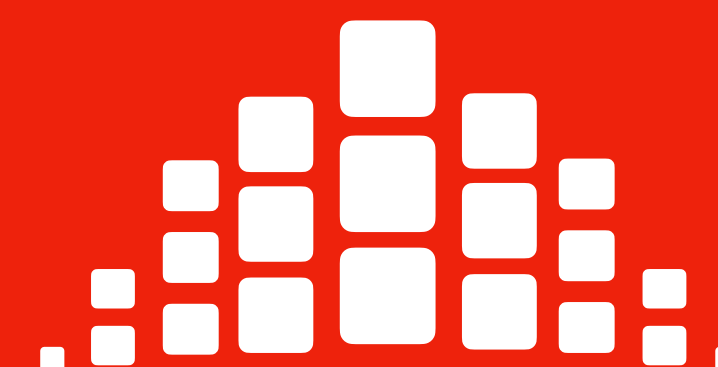
CONFIGURATION
MANAGEMENT



STATISTICS &
MONITORING



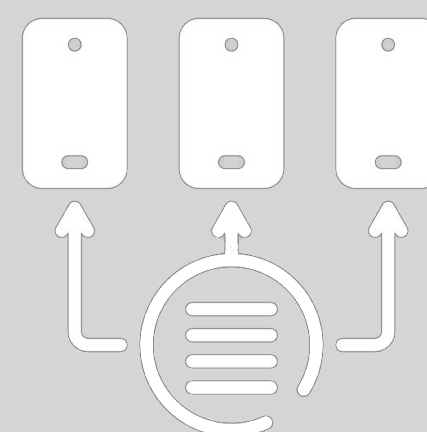
RACK
DIAGRAMMING



INTEGRATION
WITH SLURM



CUSTOMIZABLE
DASHBOARDS



PROVISIONING



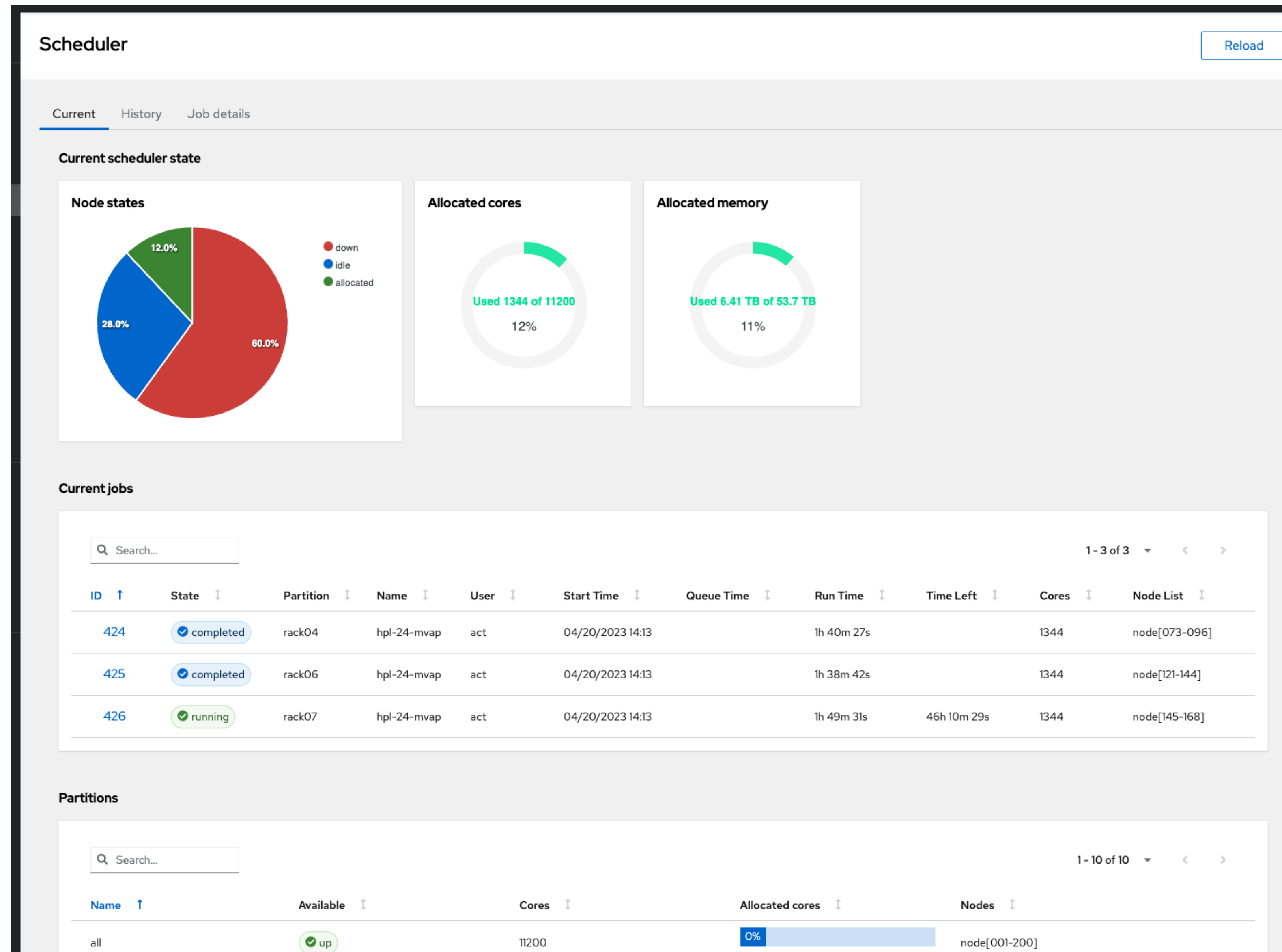
USER
MANAGEMENT



COMMAND LINE
TOOLS

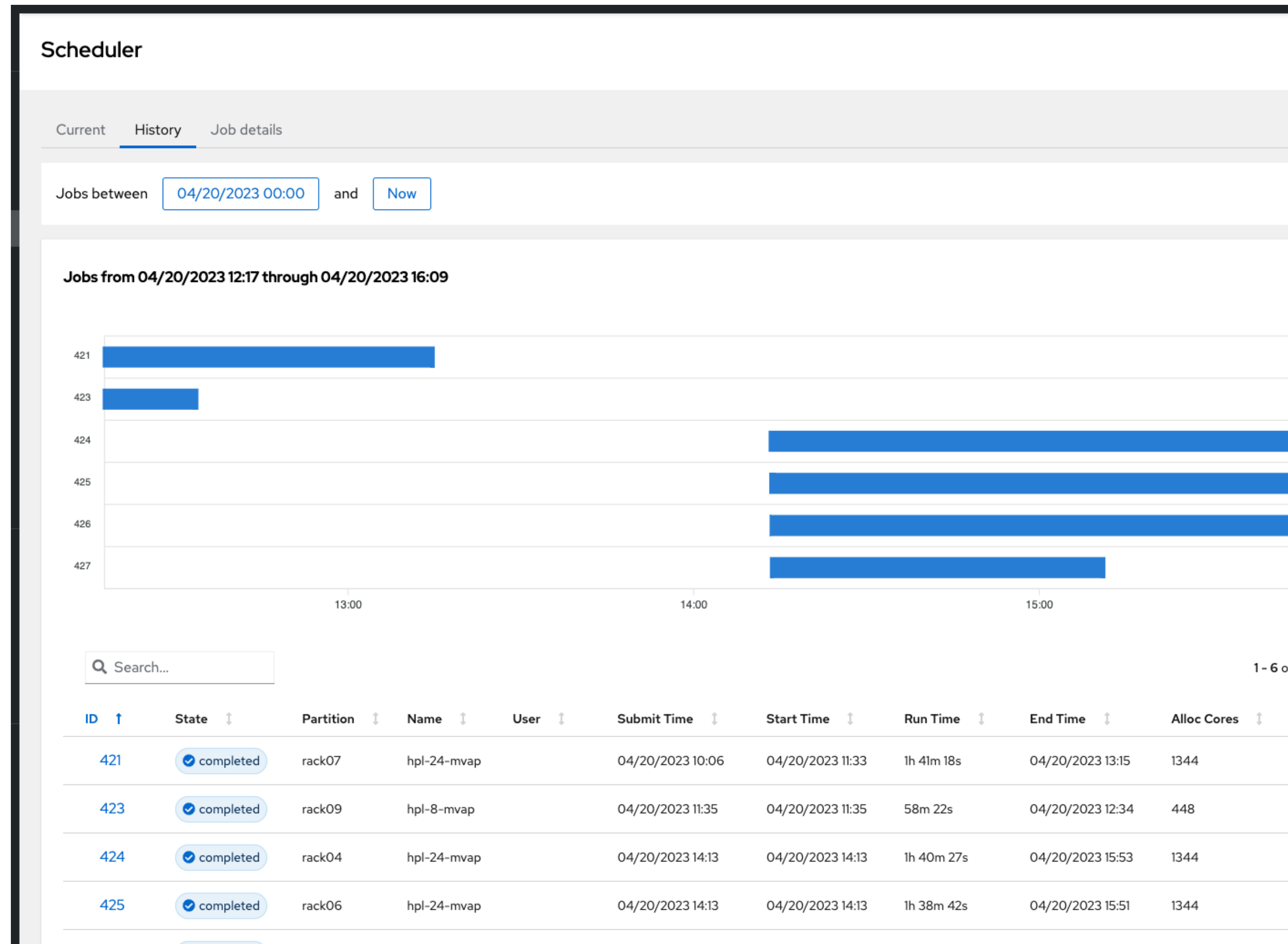
SLURM integration

- ClusterVisor now can pull data from your SLURM installation
- No data duplication, uses the existing SLURM REST API to pull data from slurmctld and slurmdbd



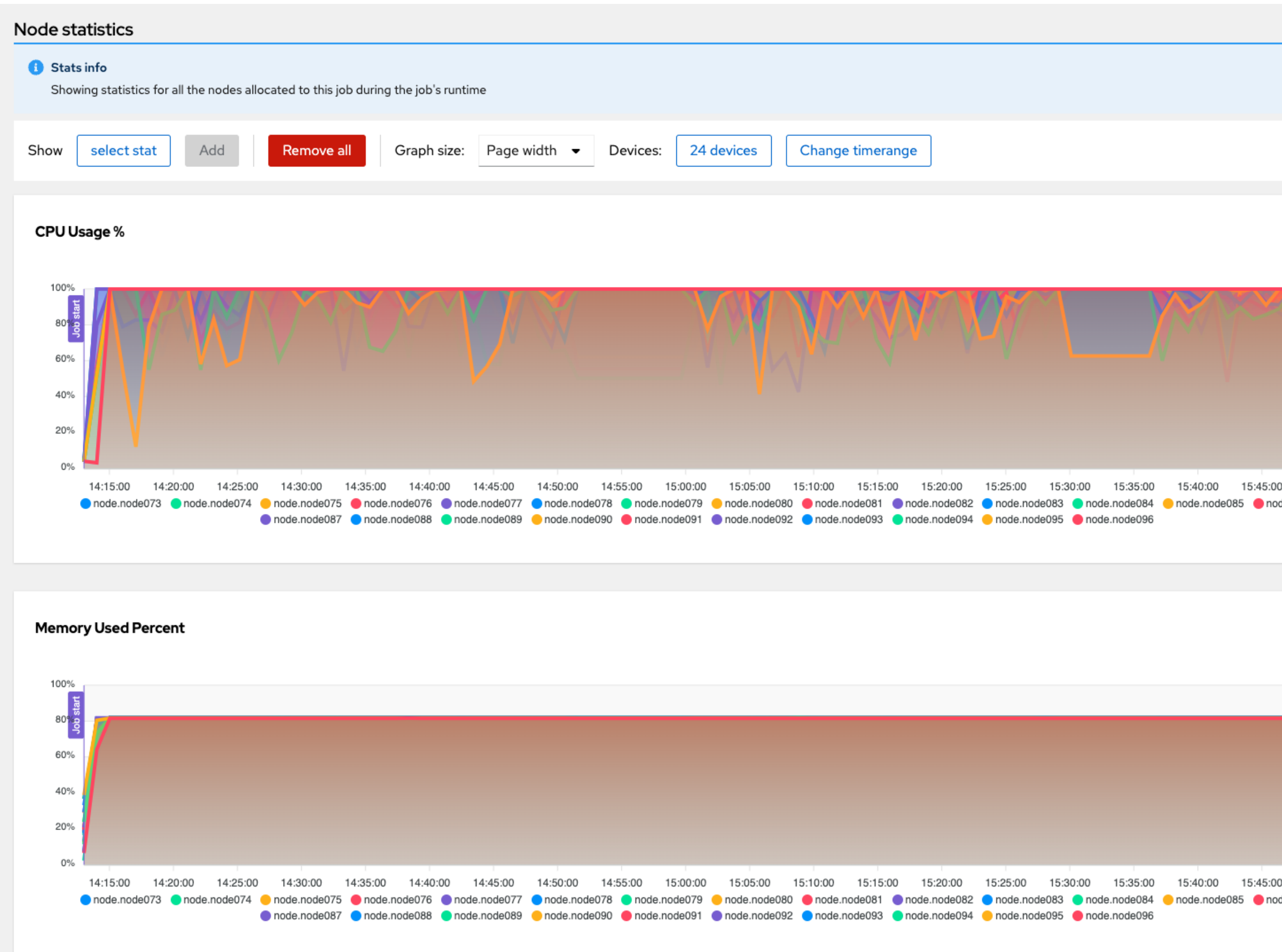
SLURM integration

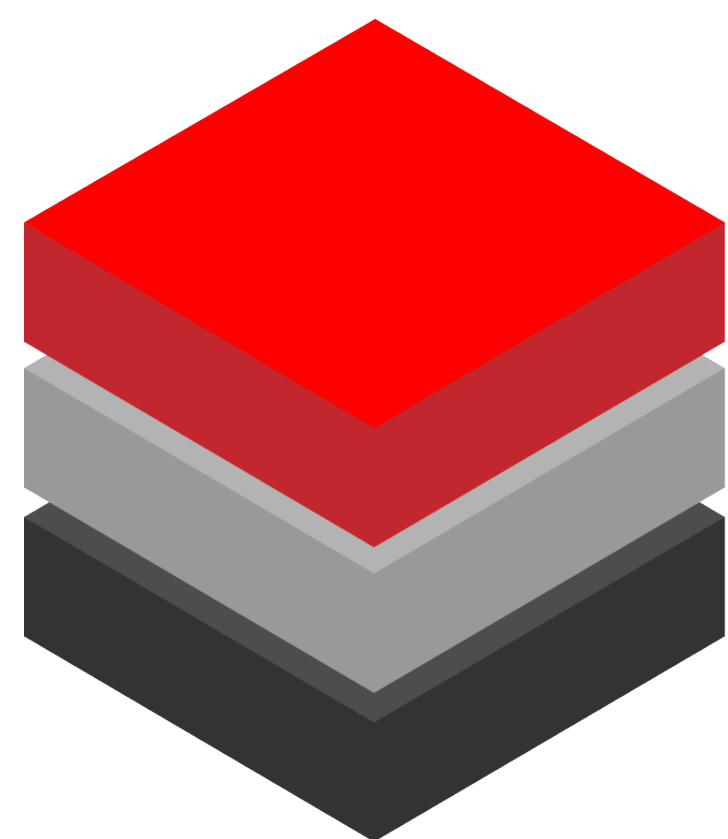
- Easily query job history information
- Query by time and see jobs on an interactive timescale graph to see how the system is operating



SLURM integration

- Connect job information with ClusterVisor stats
- Easily query any collected stat during a job run
- CPU utilization, memory usage, temperatures, power consumption, etc.
- Very helpful for diagnosing problems like “my job is slow”





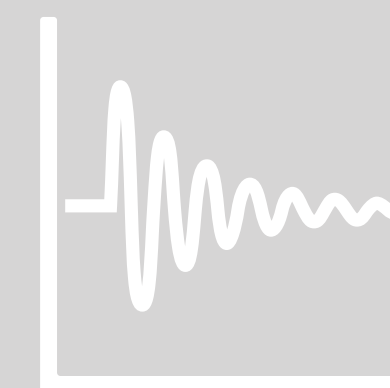
clustervisor



FLEXIBLE
DELIVERY



CONFIGURATION
MANAGEMENT



STATISTICS &
MONITORING



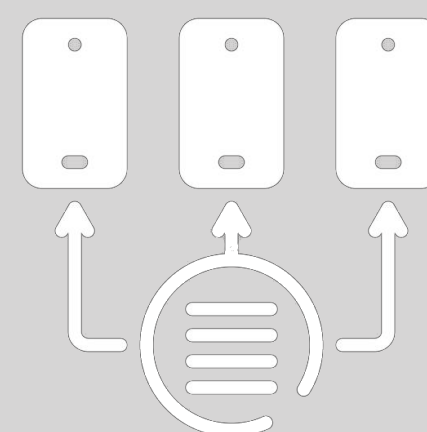
RACK
DIAGRAMMING



INTEGRATION
WITH SLURM



CUSTOMIZABLE
DASHBOARDS



PROVISIONING



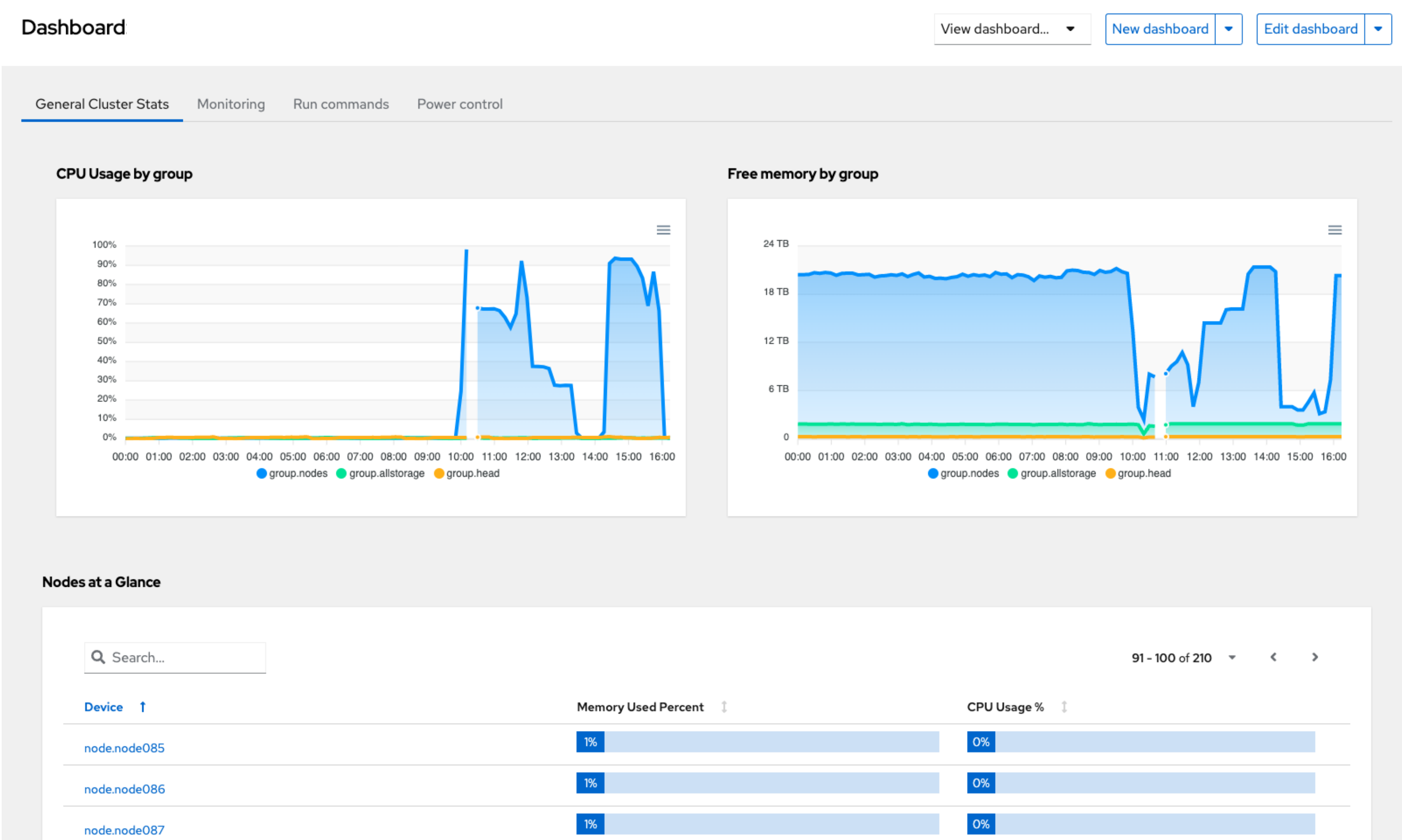
USER
MANAGEMENT



COMMAND LINE
TOOLS

Dashboards

Dashboards in ClusterVisor are customizable by the admin to show the information that they find most important on their cluster.

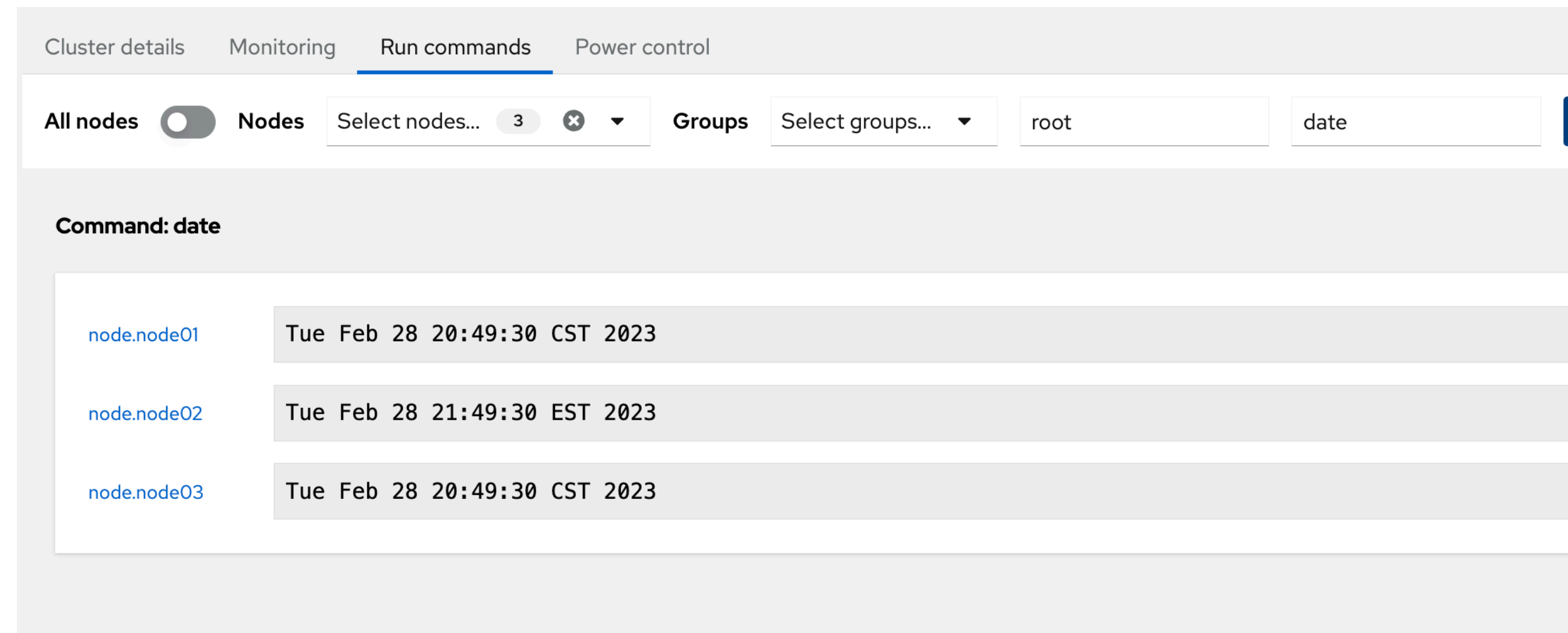


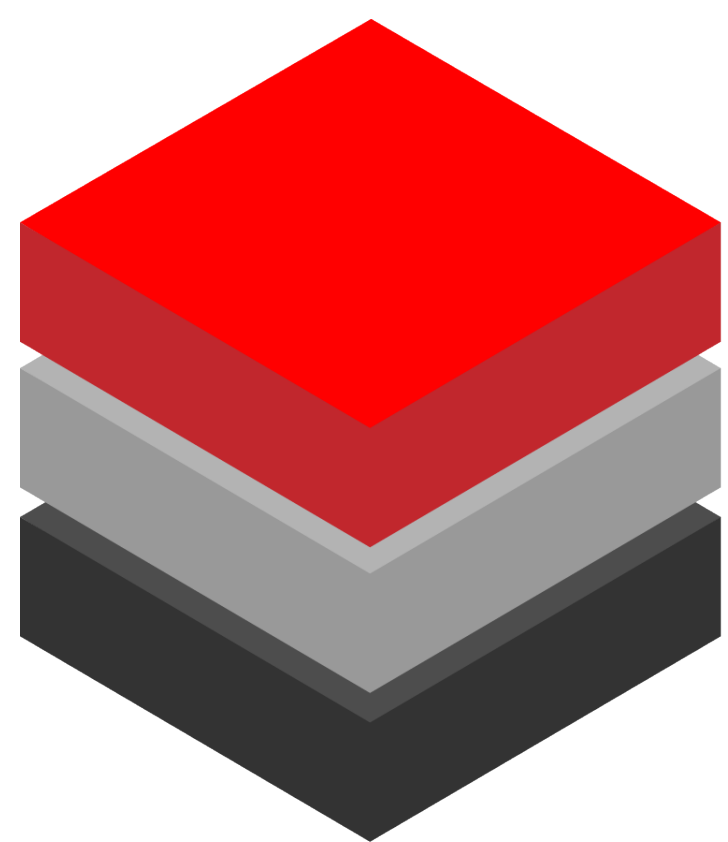
- The dashboards are useful to show you cluster-wide information and to compare multiple nodes or devices against each other
- You can create as many dashboards as you wish, with a different focus on each dashboard

Dashboards

- Dashboards contains multiple widgets that are configurable with the stats you care about:
 - Stat tables
 - Graphs
 - Gauges
 - Rack diagrams
 - Scheduler info
- You can adjust the dashboard to view different time windows

- The Dashboard section also lets you run commands across multiple nodes, and power cycle selected devices





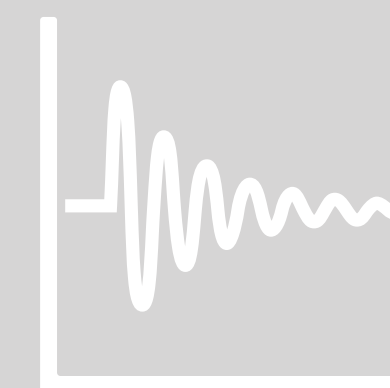
clustervisor



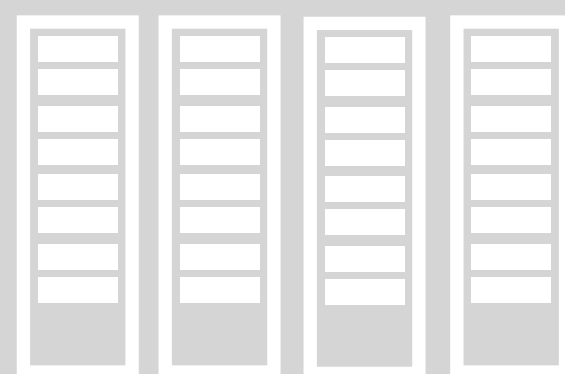
FLEXIBLE
DELIVERY



CONFIGURATION
MANAGEMENT



STATISTICS &
MONITORING



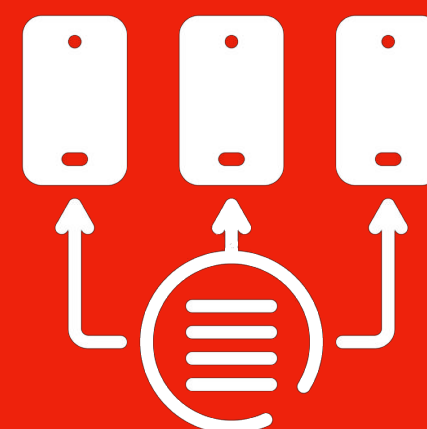
RACK
DIAGRAMMING



INTEGRATION
WITH SLURM



CUSTOMIZABLE
DASHBOARDS



PROVISIONING



USER
MANAGEMENT



COMMAND LINE
TOOLS

Provisioning

- ClusterVisor's provisioning subsystem is called "cloner"
- Full suite of tools to create images and take images from existing nodes
- Images can be updated from existing running systems
- Image data separate from disk configuration
- Image can be multicast to multiple nodes simultaneously
- From bare-metal to working node in just minutes

Cloner: Images

Images Disk Layout

New image from distro

Capture new image from node

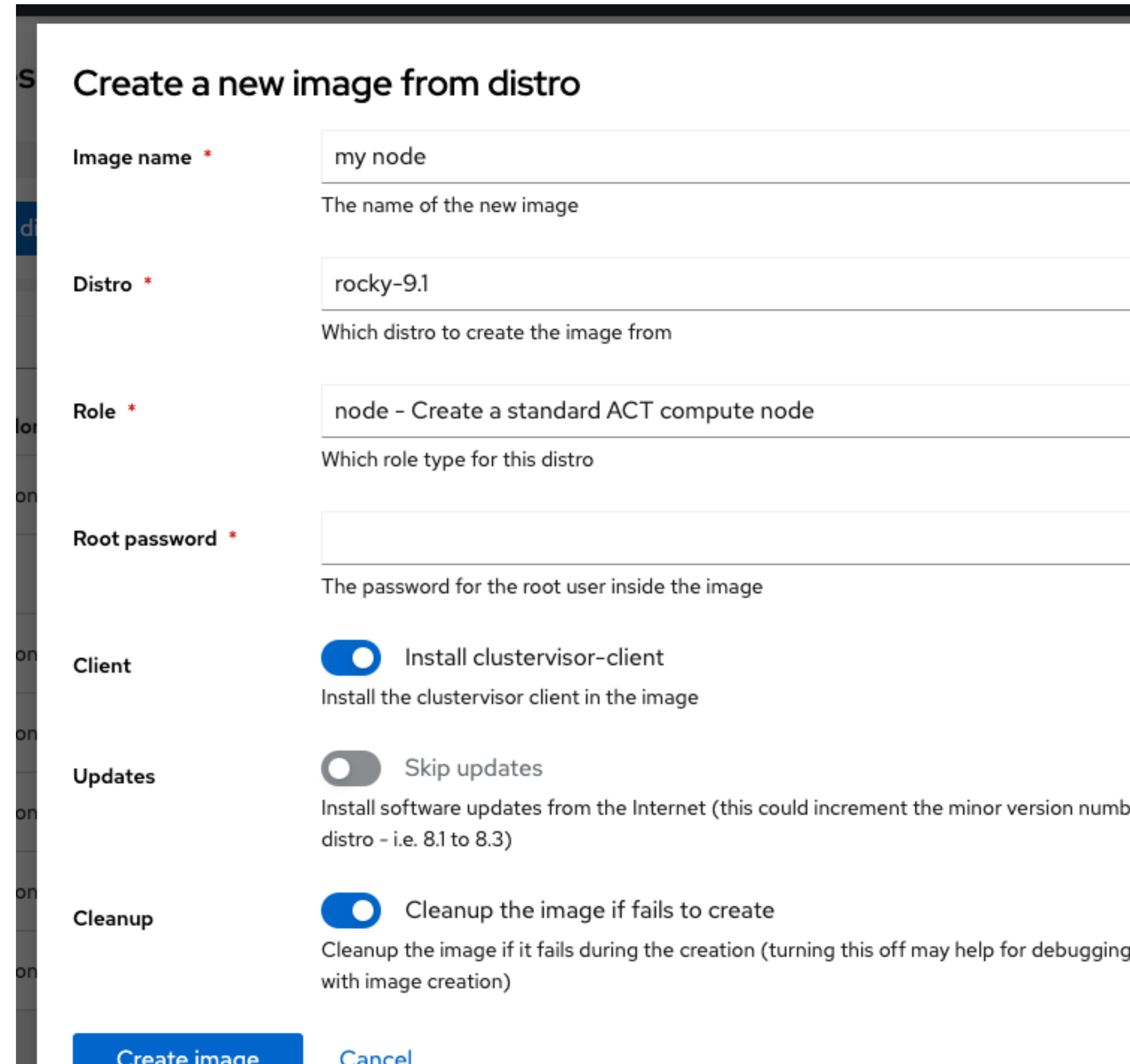
Search...

1 - 4 of 4

Image ↑	Cloner Image ↓	Path ↓	Detecte... ↓	Source ... ↓	Role ↓
node	cloner_image.node	/clustervisor/images/node	rocky 8.5		
node-rocky8	cloner_image.node-rocky8	/clustervisor/images/node-rocky8	rocky 8.6	rocky-8.6	node
node-rocky9	cloner_image.node-rocky9	/clustervisor/images/node-rocky9	rocky 9.1	rocky-9.1	node
test	cloner_image.test	/clustervisor/images/test	rocky 9.1	rocky-9.1	node

Provisioning

- New in the 1.x release:
- Updated to support EL8 and EL9 distros: Rocky, RedHat, Alma, etc
- Upload distro ISOs and then create new images from the uploaded distro
- No need to install a node first
- Templates for: node, login, storage, etc



The screenshot shows a web form titled "Create a new image from distro". It contains several input fields and toggle switches. The "Image name" field is filled with "my node". The "Distro" field is filled with "rocky-9.1". The "Role" field is filled with "node - Create a standard ACT compute node". The "Root password" field is empty. There are three toggle switches: "Client" (checked), "Updates" (unchecked), and "Cleanup" (checked). At the bottom, there are two buttons: "Create image" and "Cancel".

Create a new image from distro

Image name * my node
The name of the new image

Distro * rocky-9.1
Which distro to create the image from

Role * node - Create a standard ACT compute node
Which role type for this distro

Root password *
The password for the root user inside the image

Client ☒ Install clustervisor-client
Install the clustervisor client in the image

Updates ☐ Skip updates
Install software updates from the Internet (this could increment the minor version number of the distro - i.e. 8.1 to 8.3)

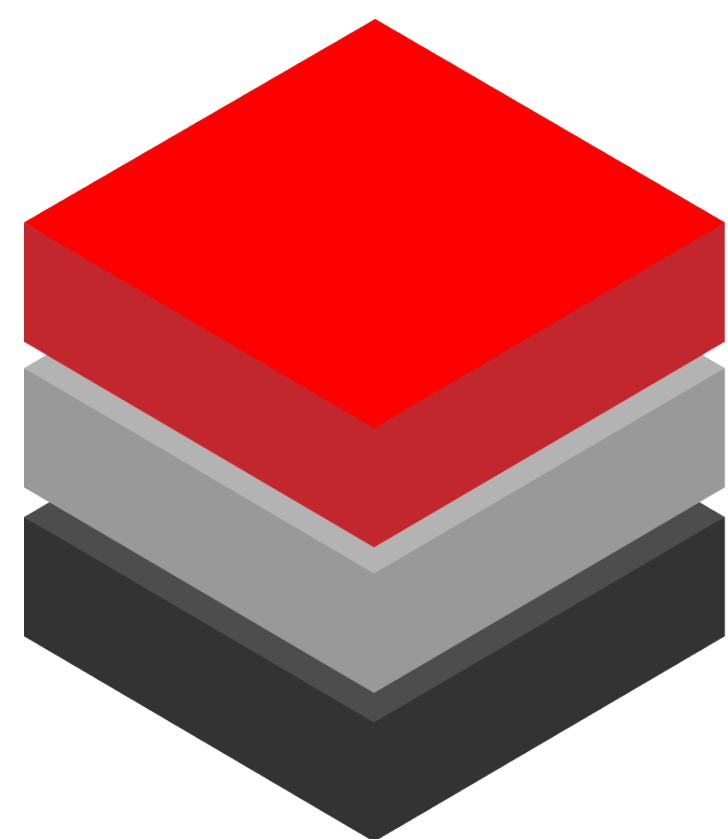
Cleanup ☒ Cleanup the image if fails to create
Cleanup the image if it fails during the creation (turning this off may help for debugging with image creation)

[Create image](#) [Cancel](#)



Provisioning

- Cloner combined with a ClusterVisor Appliance node:
 - Install your login node, and storage nodes directly from the appliance
 - Special supported added for non-internet connected systems for highly secure environments
- Take backup images of all your systems in case of a hardware failure or software issue
- Easily restore not only compute nodes but login and storage nodes too



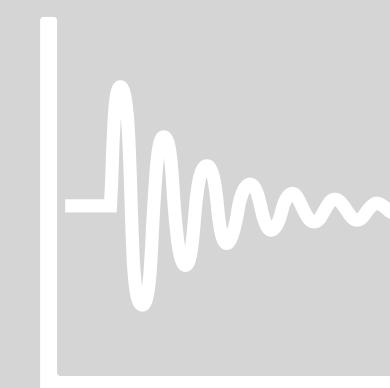
clustervisor



FLEXIBLE
DELIVERY



CONFIGURATION
MANAGEMENT



STATISTICS &
MONITORING



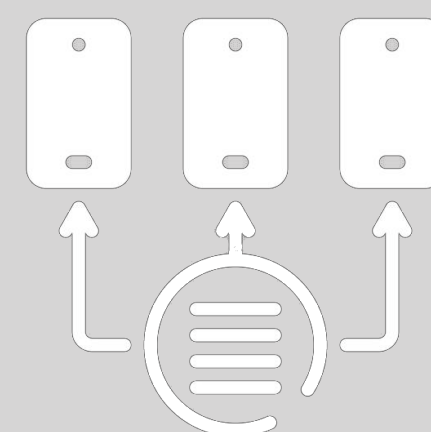
RACK
DIAGRAMMING



INTEGRATION
WITH SLURM



CUSTOMIZABLE
DASHBOARDS



PROVISIONING



USER
MANAGEMENT



COMMAND LINE
TOOLS

Users and groups

- ClusterVisor can run and manage an LDAP server for use inside your cluster
- Web and command line tools for making users, home directories, SSH keys, and slurm accounting users
- Internal LDAP support is optional, and can use external authentication if desired (Active Directory, etc)

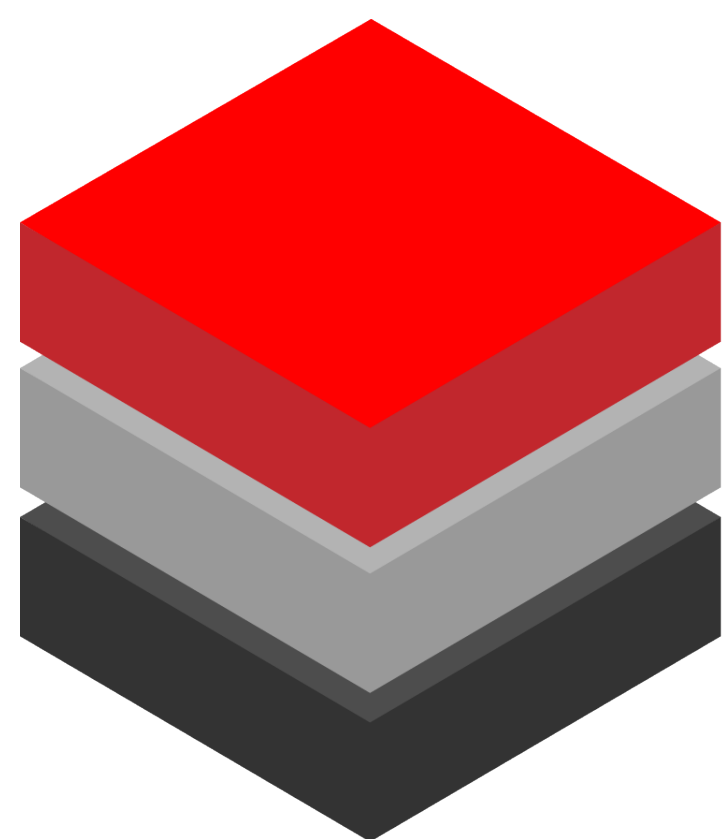
LDAP Users/Groups

User Group			
New name: <input type="text"/>		<button>Add User</button>	
User name	Full name	UID number	Primary
demo_user	Demo User	99998	99998
test_user	Test User	199998	1000
cv-admin	cv-admin	199999	100000



Access Control

- ClusterVisor supports setting users as “admins” so they have control to make changes on your cluster.
- With “admin” privileges they can login to ClusterVisor and edit configuration, dashboards, make users, etc.
- ClusterVisor also has the ability manage your sudo setup for privilege based access on login and/or compute nodes



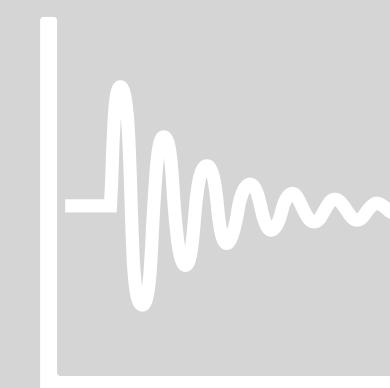
clustervisor



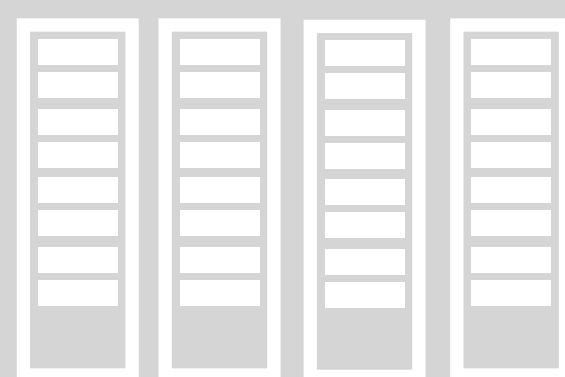
FLEXIBLE
DELIVERY



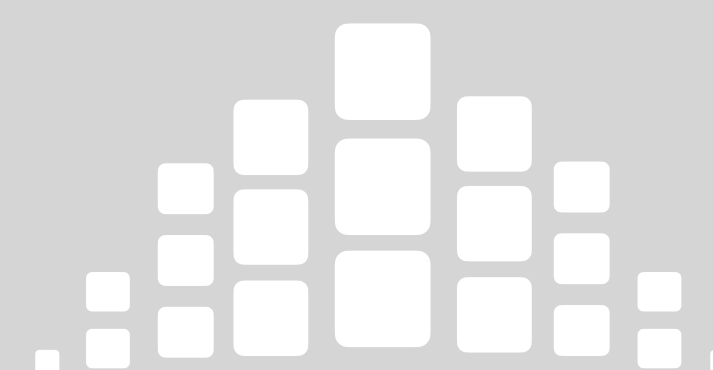
CONFIGURATION
MANAGEMENT



STATISTICS &
MONITORING



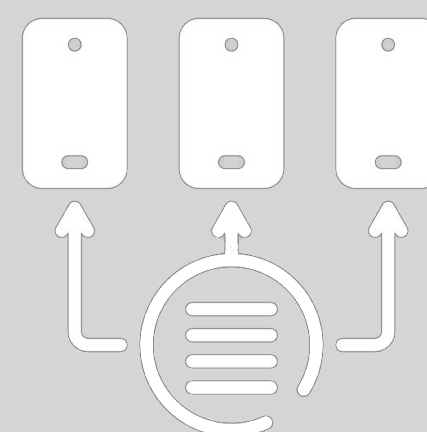
RACK
DIAGRAMMING



INTEGRATION
WITH SLURM



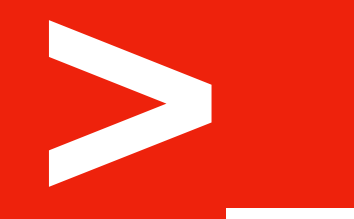
CUSTOMIZABLE
DASHBOARDS



PROVISIONING



USER
MANAGEMENT



COMMAND LINE
TOOLS



advanced clustering
technologies, inc.

cli command overview

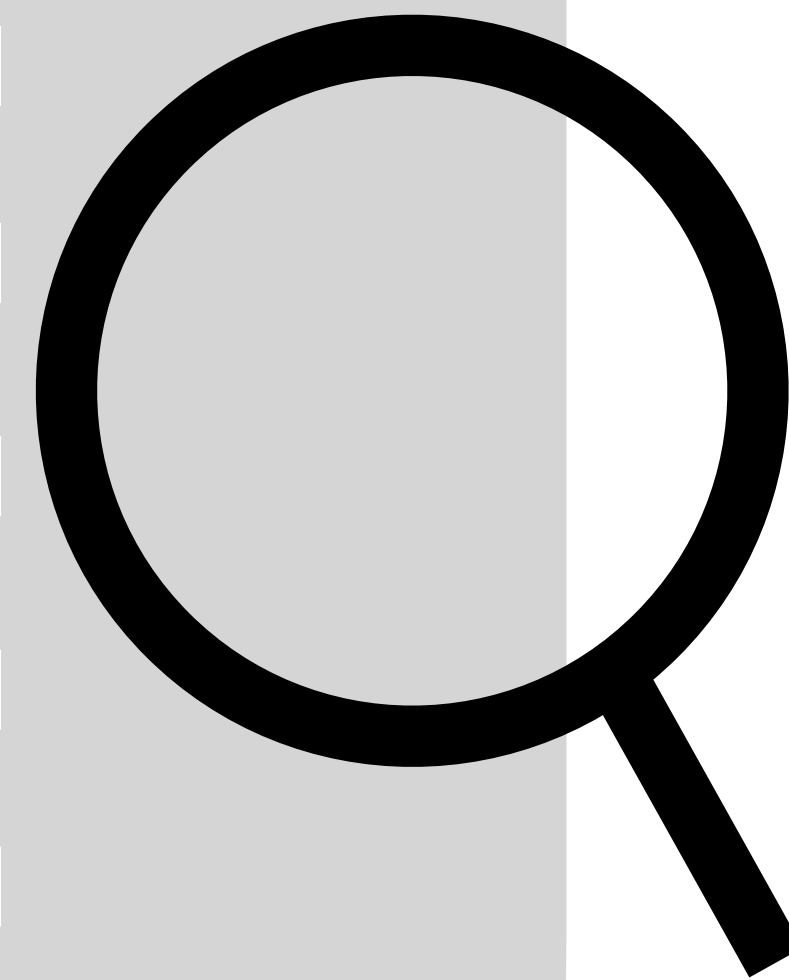
command	Description
cv-authsync	Sync password / group files across nodes (deprecated: use LDAP system instead)
cv-cloner	Create / update cloner images and disk layouts
cv-commit	Commit changes in the ClusterVisor database to the nodes
cv-conf	Edit ClusterVisor configuration settings
cv-console	Connect to a node's IPMI serial console for remote debugging
cv-cp	Copy files in parallel to multiple nodes in the cluster
cv-exec	Execute commands in parallel across the cluster
cv-identify	Turn on the identifier LED to help find systems in the rack / datacenter
cv-ipmitool	Issue arbitrary IPMI commands to devices on the cluster
cv-netboot	Change how the node will start up on it's next boot
cv-nodenames	Output node names (useful for scripting)
cv-power	Power control a node via IPMI (power on, off, reboot, etc.)
cv-reconfigure	Take all ClusterVisor configuration for this node and reapply it to the node
cv-sel	View the nodes IPMI event log
cv-sensor	Query the nodes IPMI sensors
cv-sshkey	Help create user authorized keys



cli command overview

command	Description
cv-stats	Query ClusterVisor collected stats
cv-useradm	Add / edit LDAP users
cv-db-image	Backup and restore ClusterVisor configuration
cv-distro	Upload Linux distribution to ClusterVisor
cv-image	Create / edit cloner images from distros
cv-cp	Copy files in parallel to multiple nodes in the cluster
cv-statsadm	Manage the stats database

CASE STUDY





Case Study

Caltech

A research team at Caltech is among the first customers to utilize ClusterVisor 1.0 in their HPC cluster.

The team starting using features in February such as:

- Creating custom dashboards
- Relying on user management to add new accounts and recover user passwords
- Using rack layout to see how each job is distributed across the cluster

Case Study



Ivan Maliyov

Postdoctoral Research

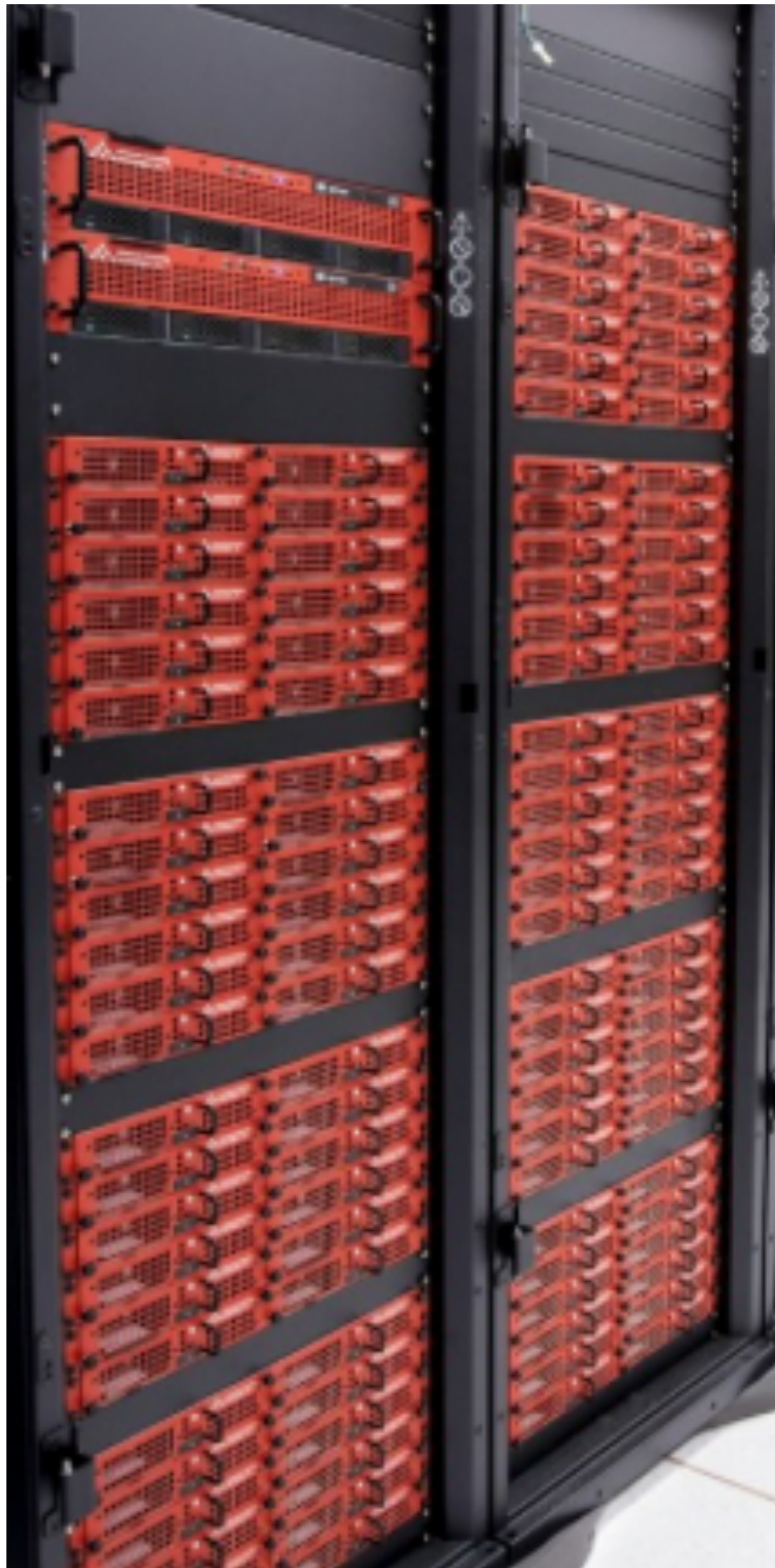
Caltech

“The one feature I really wanted to mention and say how it’s really important to us is for a particular job that we are running on the cluster, **we are able to track the RAM memory as a function of time.** This is wonderful, not only to see how our software runs on the cluster, but also to test our software. We are talking to our peers in the U.S. and Europe, and nobody among them has these tools. We share our excitement with other people and say we have this feature. This is great.”

Next Steps

Visit us at <https://www.advancedclustering.com/products/software/clustervisor/> to:

- Stay tuned for our next webinar that features a deep dive into the technical aspects of ClusterVisor 1.0.
- Download our PDF on How to Upgrade from ClusterVisor to ClusterVisor 1.0.
- Contact us to ask questions or talk about adding ClusterVisor to your HPC cluster.



Stay Up to Date

Join our **clustervisor-announce mailing list** to receive updates and announcements about ClusterVisor by Advanced Clustering.

<https://lists.advancedclustering.com/mailman/listinfo/clustervisor-announce>

You can also look back at the archives to see what updates have already been shared.



Contact Us



866-802-8222



info@advancedclustering.com



advancedclustering.com