

Federator.ai User Guide

Federator.ai version 4.5.1 User Guide

ProphetStor Data Services, Inc.
830 Hillview Court, Suite 100
Milpitas, CA 95035 USA
Phone: 1.408.508.6255
Website: www.prophetstor.com

Copyright © 2020-2021 ProphetStor Data Services, Inc. All Rights Reserved.

Federator.ai® is a registered trademark of ProphetStor Data Services, Inc in the United States and other countries.

Kubernetes is a registered trademark of the Linux Foundation, and OpenShift is a trademark of Red Hat, Inc.

All other brand and product names are trademarks or registered trademarks of their respective owners.

4.13.2021

Contents

Overview	4
Terminology	4
Getting Started	6
Access the Federator.ai Portal.....	6
Setup Wizard	8
Kubernetes Cluster	8
VM Cluster	12
Federator.ai Administration Portal	14
Portal Sections	14
Common Administration Portal Functions	14
Refresh Statistics	15
License Status	15
User Functions	15
Filter Panel.....	15
Specify Time Range.....	16
Search/Sort Information in Tables.....	16
Show/Hide Metrics in Charts	17
Dashboard	18
Cluster Workload Prediction	18
Cluster Charts	18
Node/Virtual Machine Chart	19
Namespace Chart (Kubernetes only)	20
Application Workload Prediction (Kubernetes only)	22
Application Charts	22
Controllers/Consumer Groups Chart	23
Insight - Cluster Health	25
Insight - Node Health (Kubernetes only)	27
Planning – Kubernetes or VM Workload Prediction.....	28
Managed Nodes Table (Kubernetes only)	28
Managed VMs Table (VM only)	28
Managed Containers Table (Kubernetes only)	29
Workload Prediction Table and Workload Observation and Prediction Charts	29

Workload Prediction Table.....	29
Workload Observation and Prediction Charts	30
Recommendation - HPA Recommendation (Kubernetes only)	32
CPU and memory Charts	32
Optimization – Kafka Consumer (Kubernetes only)	33
Number of Replicas Chart.....	33
Production Rate and Consumption Rate Chart	33
Consumer Lag Chart	34
Consumer Queue Latency Chart.....	34
CPU and Memory Observation Charts.....	35
Cost – Multi-cloud Cost Analysis (Kubernetes and VM Clusters)	36
CPU and Memory Utilization Charts.....	38
Recommended Cluster Configuration	38
Cost – VM Cost Analysis (VM Clusters and VMs)	40
VM Cost	40
Daily Spend Chart	41
Cumulative Cost Savings Projection Chart.....	41
Cost – Application Cost Analysis (Kubernetes Applications).....	43
Application Cost Savings Analysis	43
Application Costs	43
Cost Trends.....	44
Cost - Cost Allocation (Kubernetes Namespaces).....	45
Current Cluster Configuration	45
Namespace Charts.....	45
Configuration – Kubernetes Applications	48
Add an Application	48
Generic:	49
Kafka Consumer:	50
Manage Applications	51
Configuration - Kubernetes Clusters.....	52
Add a Cluster	53
Manage Clusters.....	54
Configuration - VM Clusters	55
Add a Cluster	56

Manage Clusters	57
Configuration - System	58
Admin Password	58
Metrics Data Source	58
Licenses	59
Add a License.....	59
Manage Licenses	60
Price Books	61
Public Cloud Price Books	61
Custom Price Book	61
Events.....	62

Overview

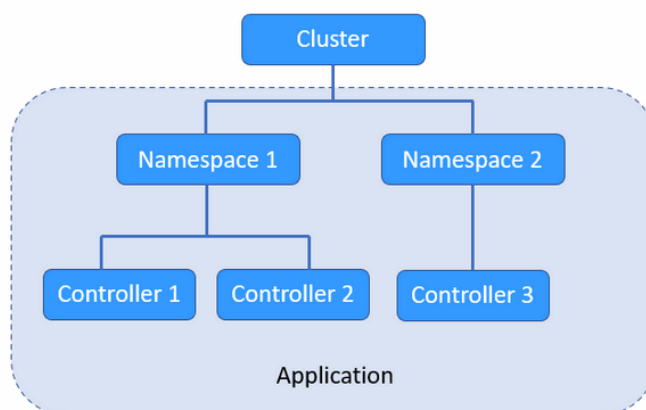
ProphetStor Federator.ai is an AI-based solution that helps enterprises manage and optimize resources for applications on Kubernetes and virtual machines (VMs) in VMware clusters. Using advanced machine learning algorithms to predict application workloads, Federator.ai offers:

- AI-based workload prediction for containerized applications in Kubernetes clusters and VMs in VMware clusters
- Resource recommendations based on workload prediction, application, Kubernetes, and other related metrics
- Automatic scaling of application containers
- Multicloud cost analysis and recommendations based on workload predictions for Kubernetes clusters and VM clusters
- Actual cost and potential savings based on recommendations for clusters, Kubernetes applications, VMs, and Kubernetes namespaces

If you have not installed Federator.ai yet, refer to your *Federator.ai Installation Guide* for information.

Terminology

Application – Defined by Federator.ai as a group of Kubernetes controllers that work together to serve tasks from the view of the end user. For example, an e-commerce web application consists of controllers for frontend and backend and can be considered as an application. An application is not a Kubernetes object.



Autoscaling – In Kubernetes, Horizontal Pod Autoscaling (HPA) provides the ability for the system to automatically increase or decrease containers/pods or CPU/memory.

Container – An object that contains a software module with everything needed to run an application.

Controller – In Kubernetes, controllers are control loops that watch the state of your cluster, then make or request changes where needed. Each controller tries to move the current cluster state closer to the desired state. The types of controllers supported by Federator.ai are *Deployment* and *StatefulSet*. Additionally, Federator.ai supports *DeploymentConfig* controllers for OpenShift.

Cluster – A Kubernetes cluster with one or more nodes or a VM cluster with one or more VM's

Deployment – A Deployment provides declarative updates for Pods and ReplicaSets. The user describes a desired state in a deployment and the deployment controller changes the actual state to the desired state at a controlled rate.

HPA – Horizontal Pod Autoscaling – In Kubernetes, the system automatically increases or decreases the number of containers/pods based on the CPU/memory usage workload.

Namespace – Kubernetes supports multiple virtual clusters backed by the same physical cluster. These virtual clusters are called namespaces.

Node – In Kubernetes, nodes are server-like machines, such as a virtual machine running complete systems and multiple applications. There can be master nodes and worker nodes.

Pod – A group of one or more containers with shared storage/network resources and a specification for how to run the containers. Typically, one container runs in each pod.

Replica – A copy of a pod running for an application.

StatefulSet – A Kubernetes object that manages stateful applications. Unlike a Deployment, a StatefulSet maintains a sticky identity for each of its pods that remains the same across any rescheduling.

VM – A VMware virtual machine running on a physical *host* machine.

VM Cluster – A cluster with one or more VM's.

Related topic:

[Federator.ai Administration Portal](#)

[Configure Applications](#)

[Configure Kubernetes Clusters](#)

[Configure VMWare Clusters](#)

[System Configuration](#)

Getting Started

After installation of Federator.ai, you must access the Federator.ai portal in order to configure your system.

Access the Federator.ai Portal

To access the Federator.ai administration portal, use the URL that is displayed at the end of the installation process.

You can also find the URL for the Federator.ai administration portal via the following methods:

Kubernetes

In a Kubernetes environment, use the `kubectl` command to find the administration portal service port number and node IP address.

```
# kubectl get svc -n federatorai |grep federatorai-dashboard-frontend-node-port
```

The output will look something like this:

```
federatorai-dashboard-frontend-node-port NodePort 10.103.181.133 <none> 9001:31012/TCP
```

Get the node's IP to access (INTERNAL-IP).

```
$kubectl get nodes -o wide
```

For example:

```
# kubectl get nodes -o wide
NAME          STATUS    ROLES    AGE   VERSION   INTERNAL-IP   EXTERNAL-IP   OS-IMAGE
KERNEL-VERSION   CONTAINER-RUNTIME
h7-130        Ready     master   35d   v1.18.5   172.31.7.130  <none>        CentOS Linux 7 (Core)
3.10.0-957.el7.x86_64 docker://19.3.13
h7-131        Ready     <none>    35d   v1.18.5   172.31.7.131  <none>        CentOS Linux 7 (Core)
3.10.0-957.el7.x86_64 docker://19.3.13
h7-132        Ready     <none>    35d   v1.18.5   172.31.7.132  <none>        CentOS Linux 7 (Core)
3.10.0-957.el7.x86_64 docker://19.3.13
h7-133        Ready     <none>    35d   v1.18.5   172.31.7.133  <none>        CentOS Linux 7 (Core)
3.10.0-957.el7.x86_64 docker://19.3.13
```

The URL will be `https://172.31.7.130:31012`.

OpenShift

In an OpenShift environment, use the `oc get route` command to find the URL.

```
# oc get route -n federatorai | grep federatorai-dashboard-frontend
```

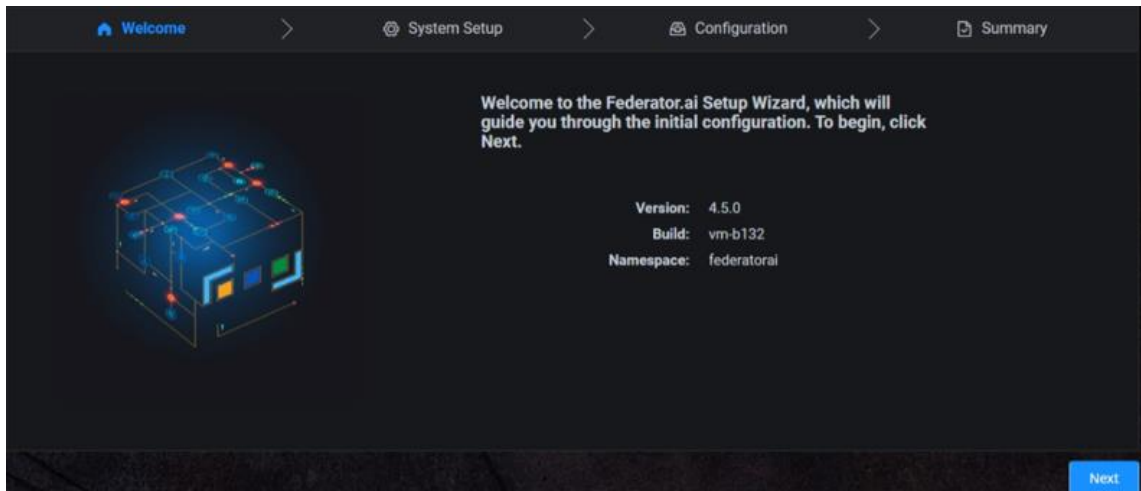
The output will look something like this:

```
federatorai-dashboard-frontend "federatorai-dashboard-frontend-federatorai.apps.ocp4.172-31-11-30.nip.io"
```

The URL will be `https://federatorai-dashboard-frontend-federatorai.apps.ocp4.172-31-11-30.nip.io`

Setup Wizard

The first time you log in after installation of Federator.ai, a configuration wizard launches that allows you to configure a cluster that should be monitored by Federator.ai. You can add more clusters from the *Configuration* section of the portal after completing the setup wizard. Click *Next* to begin configuration.



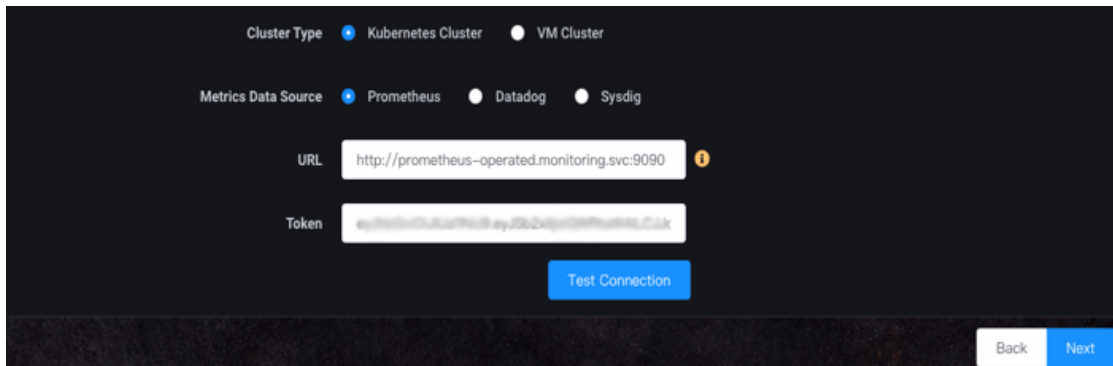
Kubernetes Cluster

1. Set the administrator password and the name of a cluster to be monitored by Federator.ai, select *Kubernetes Cluster* and the source of metrics for this cluster, and specify authentication information. Then, click *Test Connection* to confirm that all information is correct. Click *Next* when the system can connect to the cluster.

The screenshot shows the 'System Setup' step of the Federator.ai Setup Wizard. The navigation bar at the top is the same as the previous screen, but the 'System Setup' tab is now active. The main content area is divided into two sections. The first section, titled 'Administrator account', has a user icon and asks to 'Set the administrator password.' It contains two password input fields: 'New Password' and 'Confirm Password'. The second section, titled 'Monitoring & Metrics Data Source', has a database icon and asks to 'Specify the name of a cluster to be monitored by Federator.ai and the source of metrics for this cluster.' It contains several fields: 'Cluster Name' (with the value 'my-k8s-1'), 'Cluster Type' (with radio buttons for 'Kubernetes Cluster' and 'VM Cluster', where 'Kubernetes Cluster' is selected), 'Metrics Data Source' (with radio buttons for 'Prometheus', 'Datadog', and 'Sysdig', where 'Prometheus' is selected), 'URL' (with the value 'http://prometheus-operated.monitoring.svc:9090'), and 'Token' (with a long alphanumeric string). A blue 'Test Connection' button is located below the 'Token' field. At the bottom right, there are 'Back' and 'Next' buttons.

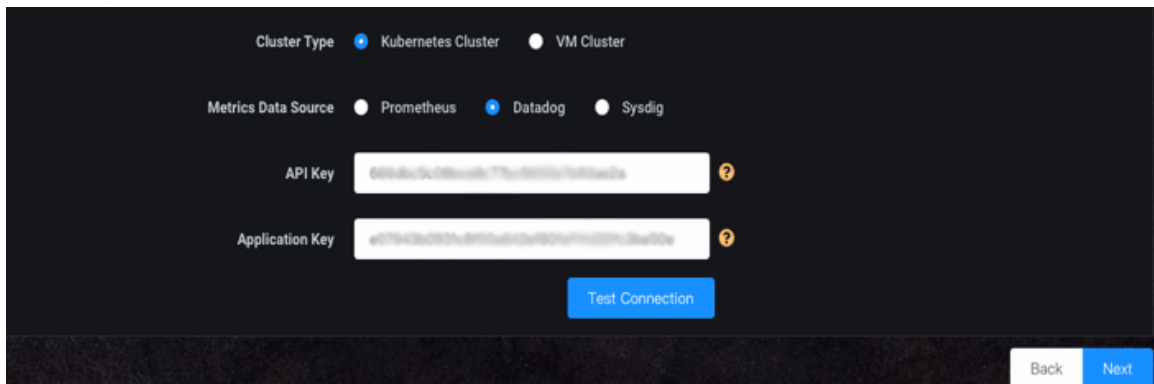
You must set the administrator password in order to continue but your cluster can be configured later from the *Configuration* section of the portal after completing the setup wizard.

For the Prometheus open-source monitoring system, the URL is required but the token is optional for authentication. If needed, you can click on the link to the Prometheus website, which is included in the popup help text.



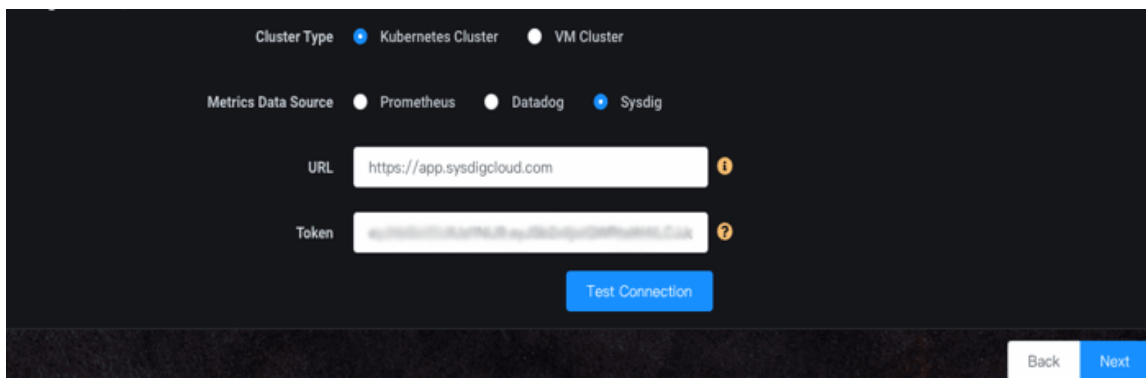
The screenshot shows the Prometheus configuration interface. At the top, 'Cluster Type' has 'Kubernetes Cluster' selected. Below, 'Metrics Data Source' has 'Prometheus' selected. The 'URL' field contains 'http://prometheus-operated.monitoring.svc:9090' with a help icon. The 'Token' field is empty. A 'Test Connection' button is at the bottom center. 'Back' and 'Next' buttons are at the bottom right.

For Datadog, the API key and application key are required for authentication. If needed, you can click on the link to the Datadog website, which is included in the popup help text.



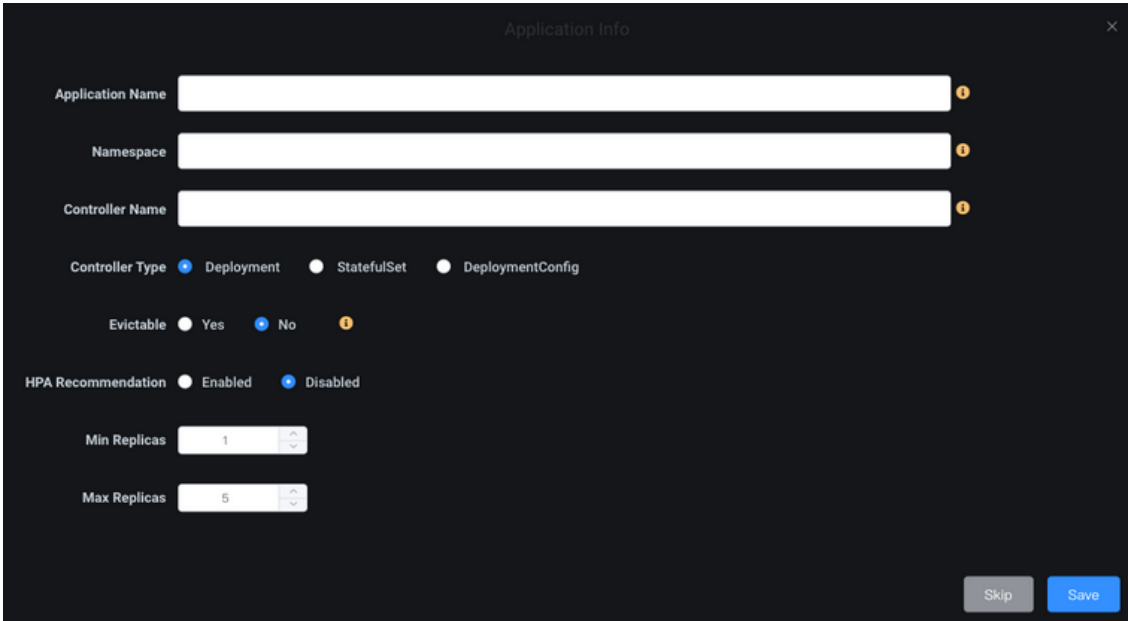
The screenshot shows the Datadog configuration interface. 'Cluster Type' is 'Kubernetes Cluster'. 'Metrics Data Source' has 'Datadog' selected. The 'API Key' and 'Application Key' fields both contain placeholder text and have help icons. A 'Test Connection' button is at the bottom center. 'Back' and 'Next' buttons are at the bottom right.

For Sysdig, a URL and token are required for authentication. If needed, you can click on the link to the Sysdig website, which is included in the popup help text.



The screenshot shows the Sysdig configuration interface. 'Cluster Type' is 'Kubernetes Cluster'. 'Metrics Data Source' has 'Sysdig' selected. The 'URL' field contains 'https://app.sysdigcloud.com' with a help icon. The 'Token' field contains placeholder text with a help icon. A 'Test Connection' button is at the bottom center. 'Back' and 'Next' buttons are at the bottom right.

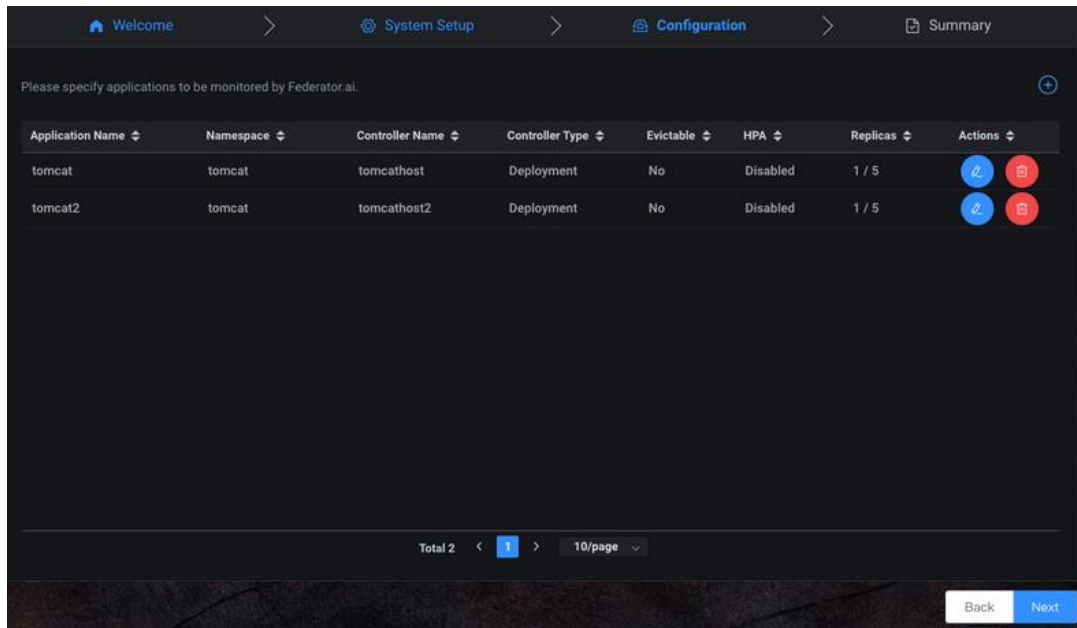
2. Enter information about the first application you want to monitor. You can skip this step and configure applications to be monitored from the *Configuration* section of the portal.



The screenshot shows a dark-themed 'Application Info' dialog box. It contains several input fields and radio button options. The fields are: 'Application Name', 'Namespace', and 'Controller Name', each with a yellow information icon to its right. Below these are three radio buttons for 'Controller Type': 'Deployment' (selected), 'StatefulSet', and 'DeploymentConfig'. Then are two radio buttons for 'Evictable': 'Yes' and 'No' (selected), with a yellow information icon to its right. Below that are two radio buttons for 'HPA Recommendation': 'Enabled' and 'Disabled' (selected). At the bottom are two numeric input fields: 'Min Replicas' (set to 1) and 'Max Replicas' (set to 5). In the bottom right corner are 'Skip' and 'Save' buttons.

You can add applications now or from the *Configuration* section of the portal after completing the setup wizard.

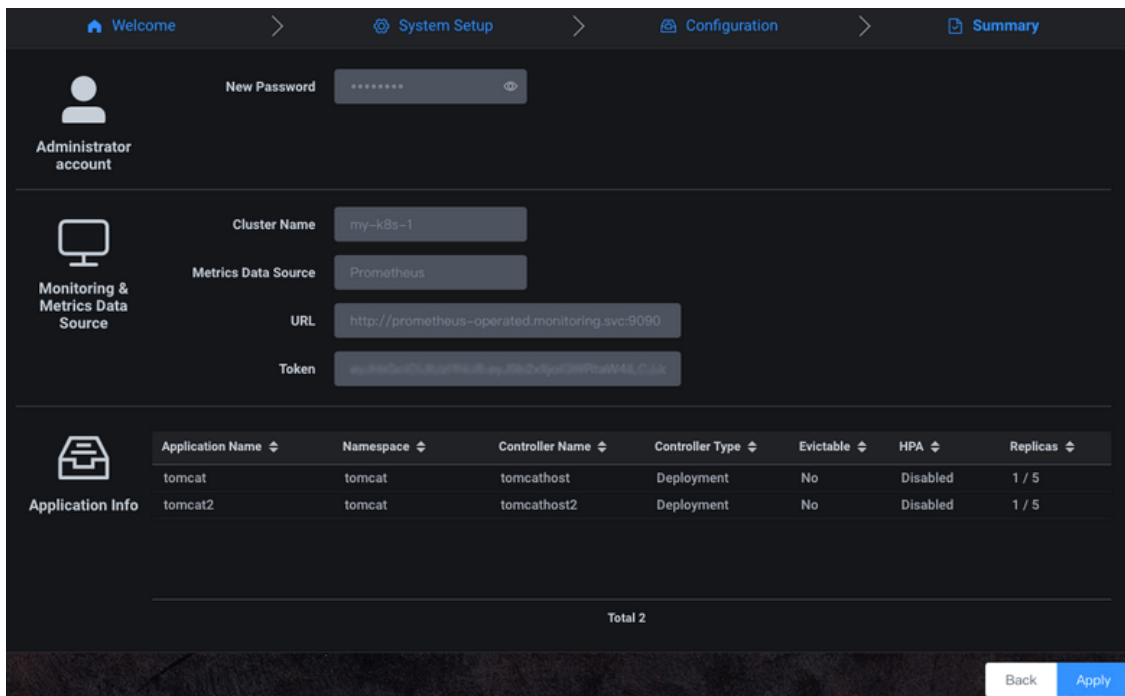
- *Application Name* - The name of your application to be monitored by Federator.ai. An application is a group of one or more Kubernetes controllers that work together to serve tasks from the view of the end user; an application is not a Kubernetes object.
 - *Namespace* – The Kubernetes namespace where the controller is deployed.
 - *Controller Name* - The name of controller to be monitored.
 - *Controller Type* – Supported controller types are *Deployment*, *StatefulSet*, and *DeploymentConfig* (OpenShift only).
 - *Evictable* – Indicate if the controller can be interrupted if the node is shut down. Evictable controllers are good candidates to be deployed in Spot instances.
 - *HPA Recommendation* – Indicate if you want to enable Horizontal Pod Autoscaling (HPA). When enabled, CPU and memory usage is monitored, and the number of pods is automatically increased/decreased based on the CPU/memory usage workload.
 - *Min/Max Replicas* – Specify the minimum and maximum number of pods when HPA is enabled.
3. Click *Save* to view a list of applications that will be monitored by Federator.ai.



Click the + key to add an application.

Click the blue *Edit* icon to update application information or click the red *Delete* icon to remove an application.

4. Click *Next* to view the *Summary* screen.



5. Click *Apply* to apply all changes or click *Back* to edit information.

VM Cluster

1. Set the administrator password and the name of a cluster to be monitored by Federator.ai, select *VM Cluster* and *vCenter* as the source of metrics for this cluster, and specify connection information.

Set the administrator password.

Administrator account

New Password

Confirm Password

Specify the name of a cluster to be monitored by Federator.ai and the source of metrics for this cluster.

Cluster Name

Cluster Type

Metrics Data Source

vCenter

Login ID

Password

Cluster Path

Test Connection

Back Next

You must set the administrator password in order to continue but your cluster can be configured later from the *Configuration* section of the portal after completing the setup wizard.

The cluster name must have a maximum of 253 lowercase characters, "-", or "." allowed. The name must start and end with an alphanumeric character.

Specify the vCenter IP address, login ID, password, and the path to the cluster, within vCenter. If needed, you can click on the link to the vCenter website, which is included in the popup help text.

2. Click *Test Connection* to confirm that all information is correct.
3. Click *Next* to view the *Summary* screen when the system can connect to the cluster.

Administrator account

New Password

Monitoring & Metrics Data Source

Cluster Name

Cluster Type

Metrics Data Source

vCenter

Login ID

Password

Cluster Path

Back Apply

The system will discover the Virtual Machines (VMs) in this cluster, which may take a few minutes to complete. Use the *Edit Cluster* function on the *Configuration / Clusters* page to select which VMs that you want to monitor.

Related topics:

[Terminology](#)

[Federator.ai Administration Portal](#)

[Configure Applications](#)

[Configure Kubernetes Clusters](#)

[Configure VM Clusters](#)

[System Configuration](#)

Federator.ai Administration Portal

The Federator.ai administration portal displays the overall health of each cluster, as well as application workload and resource recommendations. Information is presented in tables and charts.

Portal Sections

The portal is separated into the following sections:

- Dashboard – Overall system information, including the number of monitored resources, as well as cluster and Kubernetes application workload predictions and recommendations.
- Insight - Cluster and Kubernetes node health information, including CPU utilization, memory utilization, disk capacity, and VM network throughput.
- Planning – Forecasting tools, including actual CPU and memory usage observations, predicted workload usage, and recommendations.
- Recommendation - (Kubernetes only) HPA recommendations for controllers enabled with autoscaling.
- Optimization - (Kubernetes only) Kafka consumer replica autoscaling predictions for each individual topic/consumer group.
- Cost – Actual cost and potential savings based on recommendations for clusters, Kubernetes applications, VMs, and Kubernetes namespaces.
- Configuration – Configuration of clusters, Kubernetes applications and controllers/consumer groups, as well as system configuration, including resetting the admin password, metrics data source, licensing, and price books.
- Events – System events that have occurred.

Common Administration Portal Functions

The administration portal presents information in tables and charts. At the top right of each portal page, you can do the following:

- Refresh statistics
- Check license status
- Get technical support contact information
- View Federator.ai product documentation
- Display the product software version
- Log out



Refresh Statistics

By default, Federator.ai information is refreshed every five minutes. To change the interval, click the drop-down at the top right and select a 1, 5, 15, or 30-minute interval.

To force a refresh, click *Refresh Now* where the current interval is displayed.

License Status

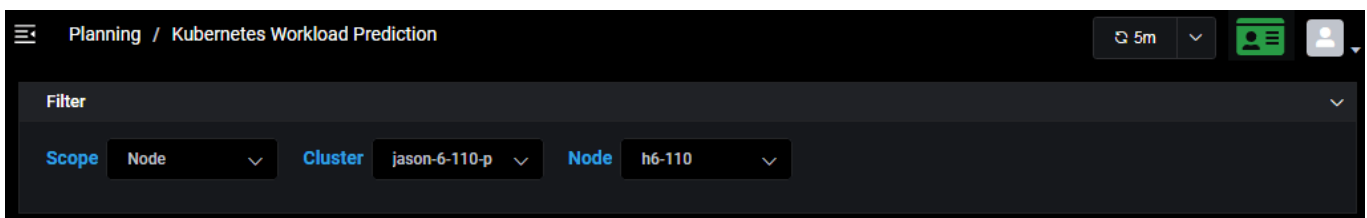
Click the *License* icon at the top right of the dashboard to see Federator.ai license information, including key code, type, CPU cores being monitored, as well as license expiration and status.

User Functions

Click the *User* icon to contact technical support, view the Federator.ai product documentation, display the product software version, or log out from the system.

Filter Panel

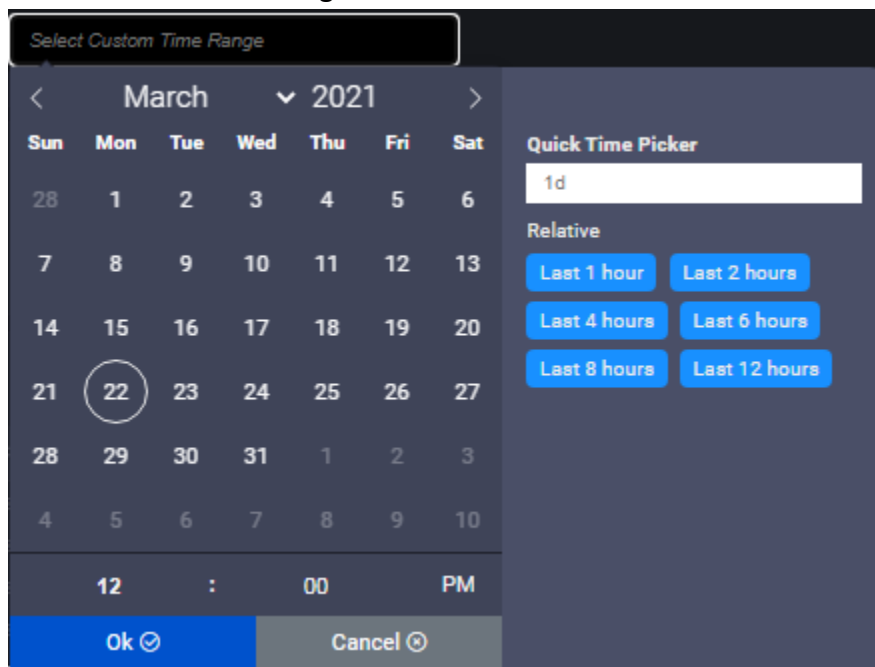
The *Filter* panel appears on most pages and allows you to filter the display of information. Depending upon the page and cluster type, you may be able to select a scope/resource type, cluster, application, namespace, controller, time range, etc.



Specify Time Range

When a chart allows you to specify a time range to display data, you can select a predefined time frame (e.g., last 1 hour, last 24 hours) from the drop-down box or you can specify a custom time range via one of the following methods:

- Use the calendar to select the start and end dates.
- Specify the number of hours (e.g., 5h), days (e.g., 5d), weeks (e.g., 3w), or months (e.g., 6m) in the *Quick Time Picker* box.
- Click a predefined relative time range.



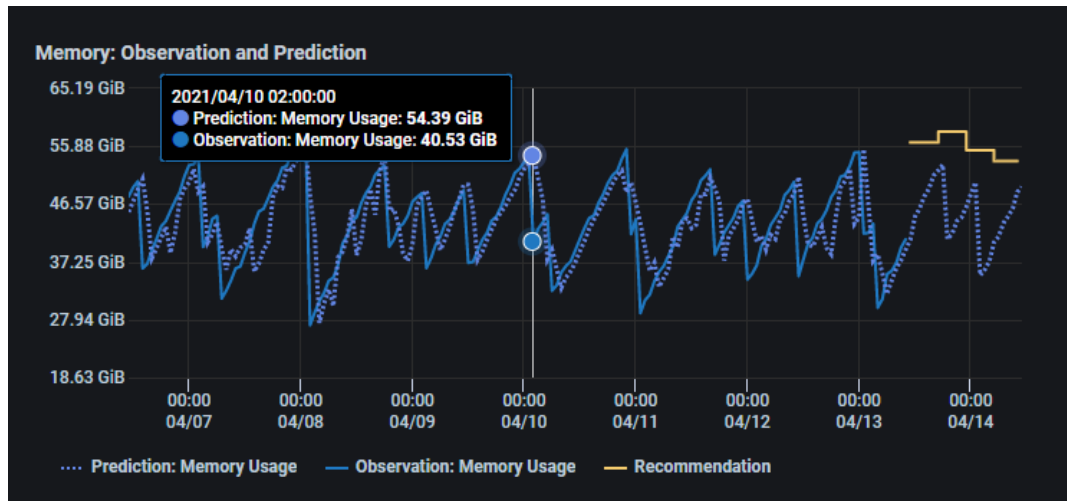
Search/Sort Information in Tables

Click a column heading to sort the entire list based upon values in that column. The blue highlighted triangle or inverted triangle on the column indicates the direction of the sort. To search, type a name or value in the *Search* box; clear the search field to return to the full view of the list. As soon as you start typing, only those items that have matching text are displayed. You can also determine how many rows to show per page (5, 10, or 20).

Managed Containers					Q Search
Container ↕	Project (Names...	Application ↕	Pod ↕	Node ↕	
my-nginx	nginx2	nginx2-jason-6-...	my-nginx-6f97b...	h6-182	
my-nginx	nginx1	alamedascaler-...	my-nginx-6c99d...	h6-182	
my-nginx	nginx1	alamedascaler-...	my-nginx-6c99d...	h6-182	
my-nginx	nginx1	alamedascaler-...	my-nginx-6c99d...	h6-182	
my-nginx	nginx1	alamedascaler-...	my-nginx-6c99d...	h6-182	
Total 6					< 1 2 > 5/page

Show/Hide Metrics in Charts

Click anywhere on a chart to see values for a specific point in time. Highlight or click on the key at the bottom of the chart to show/hide individual metrics.



Related topics:

[Common Administration Portal Functions](#)

[Refresh Statistics](#)

[License Status](#)

[User Functions](#)

[Search/Sort Information in Tables](#)

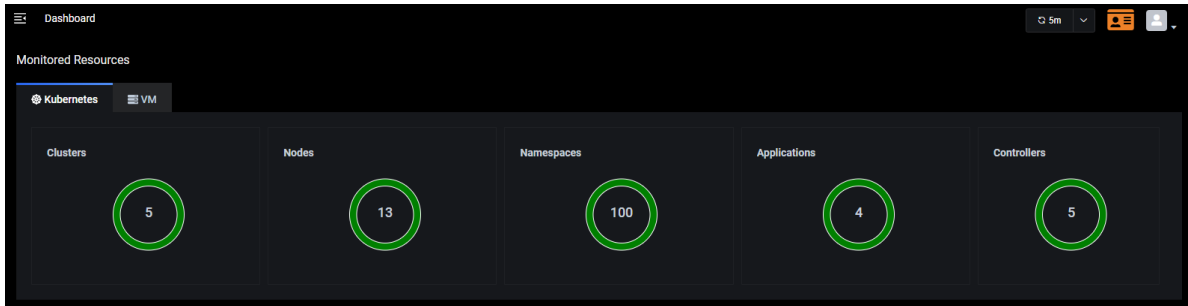
[Show/ Hide Information in Charts](#)

[Filter Panel](#)

[Dashboard](#)

Dashboard

The *Dashboard* displays the number of monitored resources in Kubernetes clusters or VM clusters, as well as cluster and Kubernetes application workload predictions and recommendations. Select Kubernetes or VM tab to view workload predictions and recommendations for resources in each type of clusters.



Cluster Workload Prediction

Select a cluster from the drop-down list and select the timeframe (daily, weekly, or monthly) to display CPU and memory observations, predictions, and recommendations.

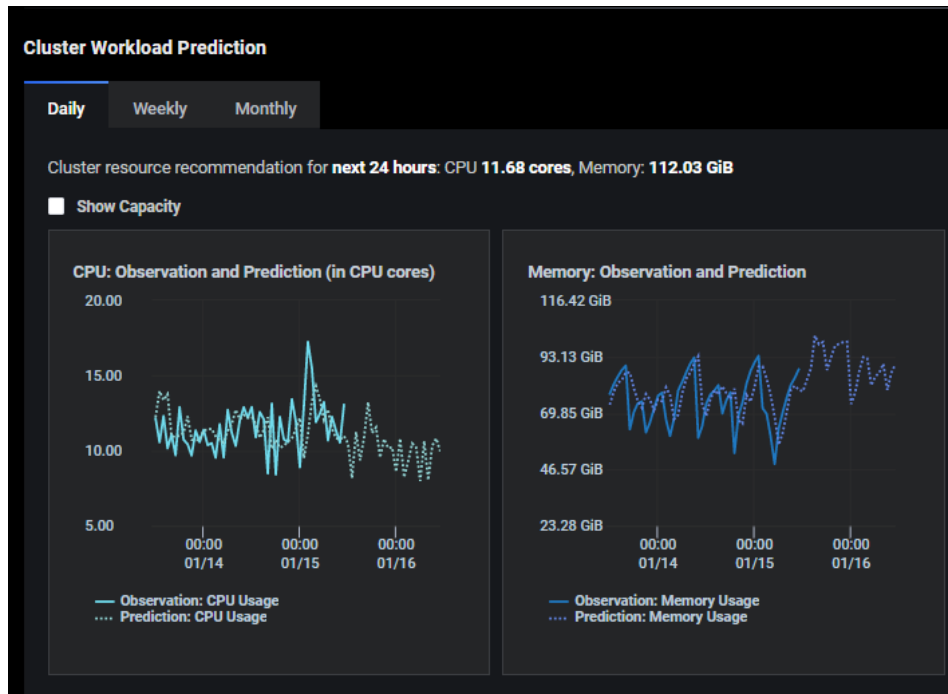
The first two charts display information for the selected cluster; the chart on the right toggles the display of information for nodes and namespaces of the selected Kubernetes cluster. For a VM cluster, the chart on the right display information for the VM's in the selected cluster.

The text above the charts summarizes the CPU and memory recommendations for the next 24 hours (daily), 7 days (weekly), or 30 days (monthly).

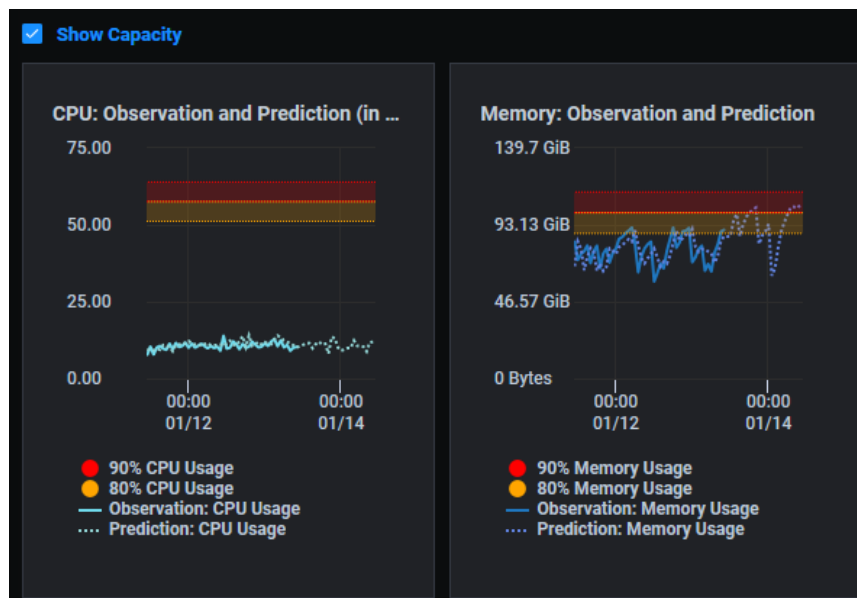
Cluster Charts

The cluster charts display CPU and memory observations and predictions for the cluster.

The solid lines represent the observed actual usage while the dotted lines show the historical and future predicted usage. Click anywhere on the charts to see values for a specific point in time. This will adjust the slider in the Node/Namespace chart accordingly.



Check *Show Capacity* to see the maximum CPU and memory usage limits for the cluster. Orange represents 80-90% and red represents 90-100%. This is a useful way to see if the utilization of resources is approaching the overall cluster capacity.

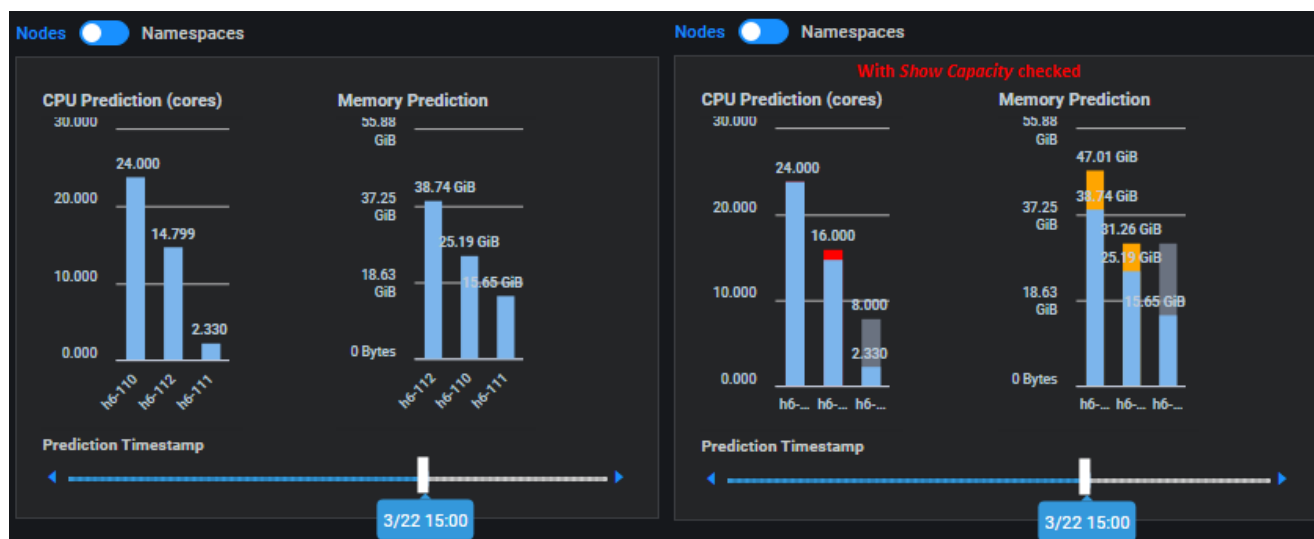


Node/Virtual Machine Chart

This chart displays CPU and memory observations and predictions for each member of the cluster. For Kubernetes, toggle to *Nodes* to display this chart.

The slider always starts at the current time but allows you to select any day/hour. Slide to the left of *Now* for historical usage; slide right for future predictions.

If *Show Capacity* was selected for the cluster, the chart will show if the node's utilization of resources is approaching the maximum CPU and memory usage limits. Orange represents less than 20% availability and red represents less than 10% availability.

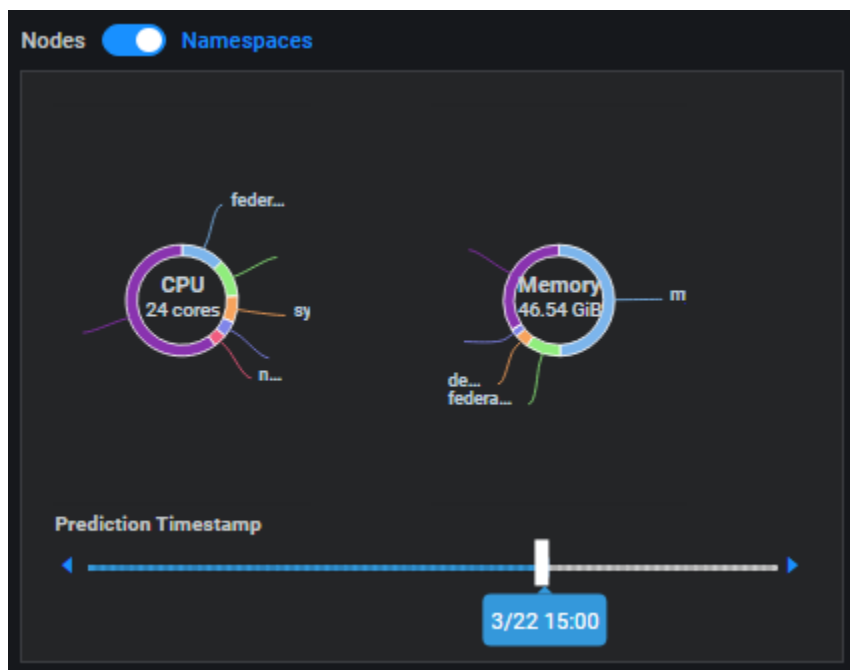


Namespace Chart (Kubernetes only)

Toggle to *Namespaces* to display CPU and memory predictions for each namespace and for when the system is idle. Move your cursor over each section to show the usage or predicted usage by a specific namespace and the overall percentage used or predicted to use by the namespace or when the resource is idle.

Highlighting the amount of predicted idle (unused) resources provides a useful way to determine where your cluster is over-provisioned and can help you balance the resource allocation within the cluster.

The slider always starts at the current time but allows you to select any day/hour. Slide to the left of *Now* for historical predictions; slide right for future predictions.



Application Workload Prediction (Kubernetes only)

Select an application from the drop-down list and select the timeframe (daily, weekly, or monthly) to display CPU and memory observations and predictions.

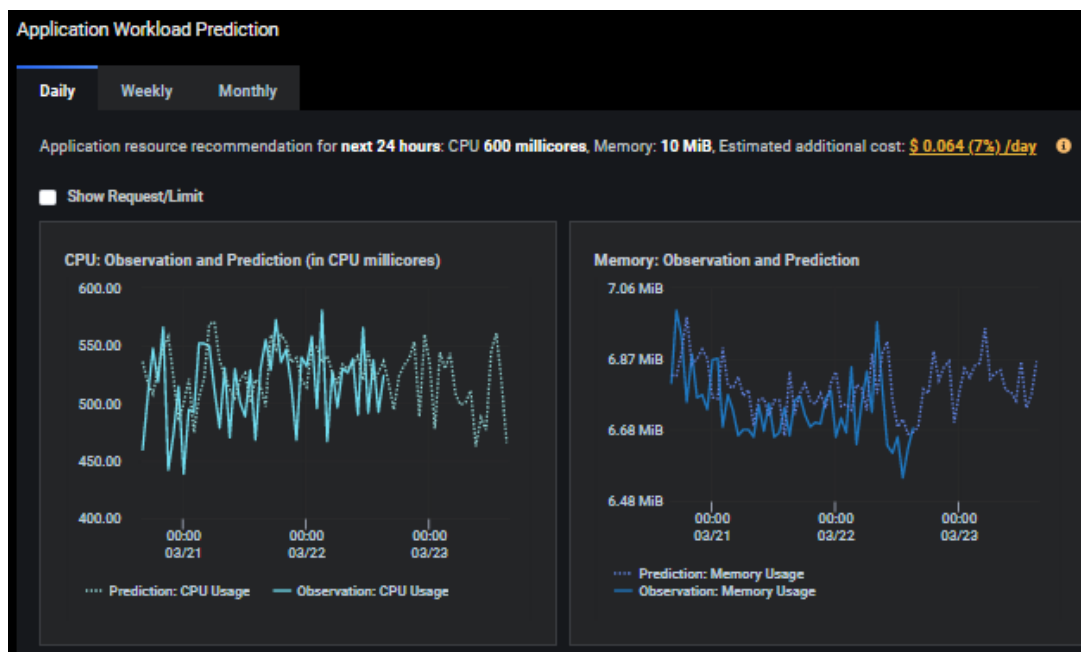
The first two charts display information for the selected application; the chart on the right displays information for controllers (generic application) or consumer groups (Kafka).

The text above the charts summarizes the CPU and memory recommendations for the next 24 hours (daily), 7 days (weekly), or 30 days (monthly). It may also show estimated savings based on the system recommendations. You can link to the *Application Cost Analysis* page for more details.

Application Charts

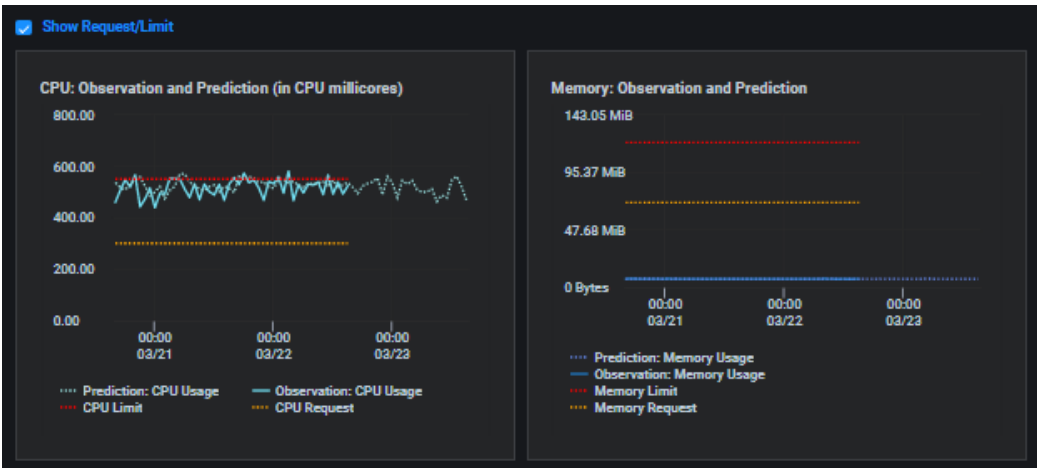
The application charts display CPU and memory observations, predictions, and recommendations for the application.

The solid lines represent the observed actual usage while the dotted lines show the historical and future predicted usage. Click anywhere on the charts to see values for a specific point in time. This will adjust the slider in the Controllers chart accordingly.



Select *Show Request/Limit* to see your application’s CPU and memory usage limits in Kubernetes. Orange represents your resource request and red represents the resource limit.

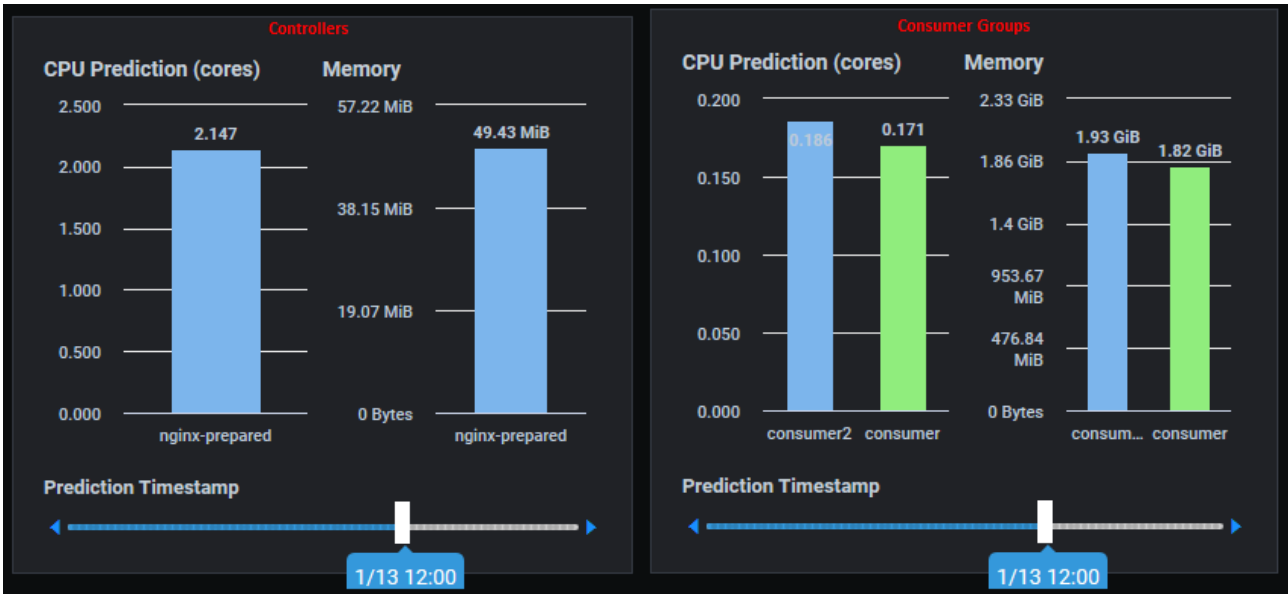
Generally, memory is a hard limit, but CPU is more *stretchable*. This is a useful way to see if you are over-provisioned or under-provisioned for your application.



Controllers/Consumer Groups Chart

The chart displays CPU and memory observations and predictions for each controller (generic application) or consumer group (Kafka) under the application.

The slider always starts at the current time but allows you to select any day/hour. Slide to the left of *Now* for historical predictions; slide right for future predictions.



Related topics:

[Common Administration Portal Functions](#)

[Refresh Statistics](#)

[License Status](#)

[User Functions](#)

[Search/Sort Information in Tables](#)

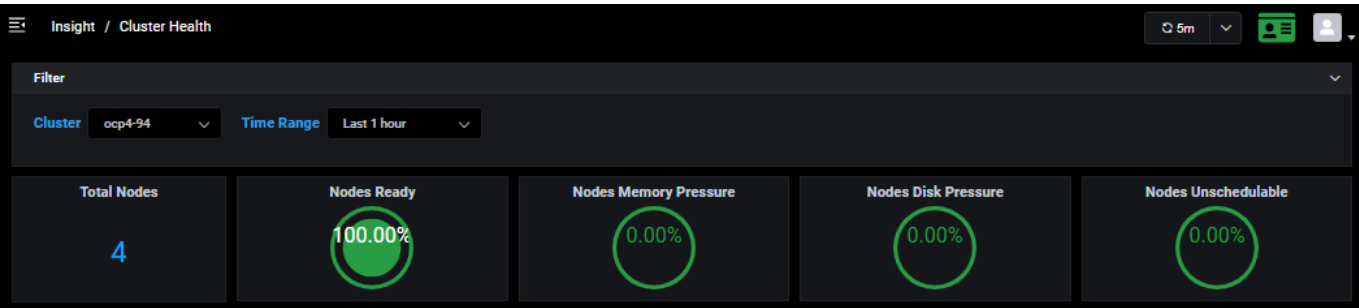
[Show/ Hide Information in Charts](#)

[Filter Panel](#)

Insight - Cluster Health

The *Cluster Health* page displays actual usage observations about the nodes/VMs in a cluster. Usage for the last 1, 2, 4, 6, or 12 hours can be displayed and can further be filtered by selecting a range of dates.

For the selected cluster, you will see the total number of nodes/VMs and the percentage that are ready. For Kubernetes, you will also see the percentage of nodes that are under memory or disk pressure, as well as the percentage of nodes that are not schedulable. Memory pressure and disk pressure are defined by Kubernetes. Refer to <https://kubernetes.io/docs/tasks/administer-cluster/out-of-resource/#node-conditions> for more information.



The table below displays the current configuration for each cluster node, including Kubernetes role (master, worker), instance types being used at your cloud provider or operating system (OS) for your local cluster, cloud provider region, number of CPUs, memory size, storage size, and node status.

Managed VMs							
Search							
Name	Role	Instance Type	Region	vCPU	Memory Size	Storage Size	Status
ocp4-qd7hn-master-0	master	m5.4xlarge	us-west-1a	16	62.91 GiB	115.83 GiB	Ready
ocp4-qd7hn-worker-0-wwnc2	worker	m5.4xlarge	us-west-1a	16	62.91 GiB	116.32 GiB	Ready
ocp4-qd7hn-worker-0-v9pbg	worker	m5.4xlarge	us-west-1a	16	62.91 GiB	116.32 GiB	Ready
ocp4-qd7hn-worker-0-2p4nn	worker	m5.4xlarge	us-west-1a	16	62.91 GiB	116.32 GiB	Ready
Total 4							

Kubernetes cluster

Managed VMs

Q Search

Name	OS Type	vCPU	Memory Size	Storage Size	Status
h4-137	CentOS 7 (64-bit)	32	64 GiB	630 GiB	Ready
h4-146	CentOS 7 (64-bit)	8	16 GiB	60 GiB	Ready
h4-145	CentOS 7 (64-bit)	8	16 GiB	60 GiB	Ready
h4-144	CentOS 7 (64-bit)	8	16 GiB	60 GiB	Ready
h4-143	CentOS 7 (64-bit)	8	16 GiB	60 GiB	Ready

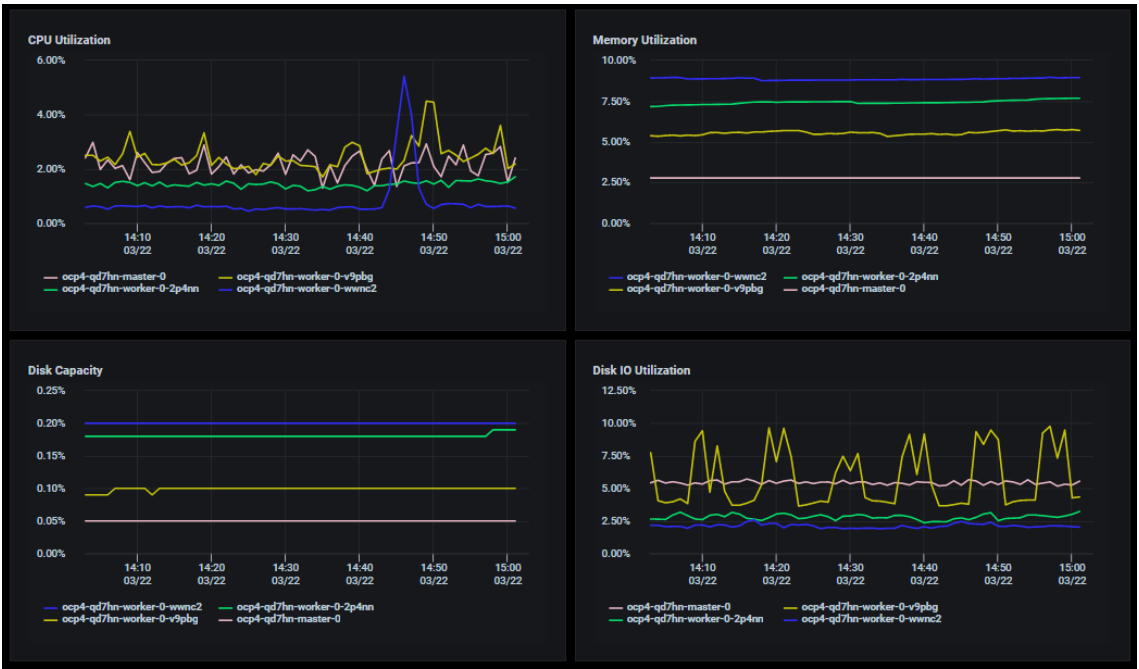
Total 9

12

5/page

VM cluster

CPU utilization, memory utilization, disk capacity, and disk IO utilization charts are displayed for each Kubernetes node.



CPU utilization, memory utilization, and network throughput charts are displayed for each VM.



Related topics:

[Search/Sort Information in Tables](#)

[Terminology](#)

Insight - Node Health (Kubernetes only)

The *Node Health* page displays actual usage observations about each node in a Kubernetes cluster. Select which cluster and node to display.

For the selected node, you will see the total CPU capacity and usage as well as memory capacity and usage.

The *Top 5* charts below display the CPU utilization of the top five pods on each node and the memory usage of the top five pods.

The *Pods Running Status Count* chart displays the number of pods running, the minimum and maximum number of pods that can run, along with the status of each pod.



Related topics:

[Search/Sort Information in Tables](#)

[Show/ Hide Information in Charts](#)

[Terminology](#)

Planning – Kubernetes or VM Workload Prediction

The *Workload Prediction* page displays actual CPU and memory usage observations, predicted usage, and recommendations.

For Kubernetes, Federator.ai monitors resource usage for monitored clusters, nodes, namespaces, user-defined applications, and controllers and provides workload prediction and recommendations at each level.

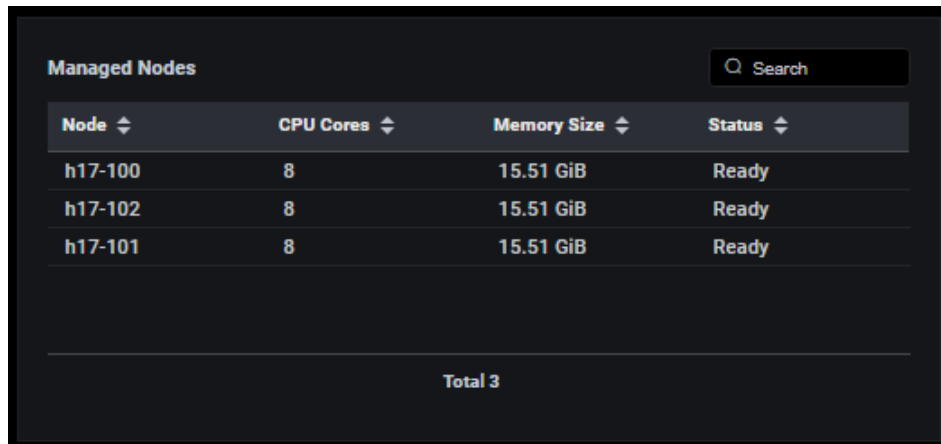
For VM, Federator.ai monitors resource usage for monitored clusters and VMs and provides workload prediction and recommendations at each level.

With the recommendations, you can decide if a resource is over-provisioned (wasting resources), or if it is under-provisioned and will not sustain an increased workload.

In the *Filter* panel, select the level of information you want to display. When you select a Kubernetes namespace, the namespace status will be displayed.

Managed Nodes Table (Kubernetes only)

This table shows the list of nodes in the selected cluster with the number of CPU cores, the size of memory, and the status for each node.



Node	CPU Cores	Memory Size	Status
h17-100	8	15.51 GiB	Ready
h17-102	8	15.51 GiB	Ready
h17-101	8	15.51 GiB	Ready
Total 3			

Managed VMs Table (VM only)

This table shows the list of VMs in the selected cluster with OS type, the number of CPU cores, the size of memory and the status for each VM.

Node	OS Type	CPU Cores	Memory Size	Status
h4-137	CentOS 7 (64-bit)	32	64 GiB	Ready
h4-148	CentOS 7 (64-bit)	8	16 GiB	Ready
h4-147	CentOS 7 (64-bit)	8	16 GiB	Ready
h4-146	CentOS 7 (64-bit)	8	16 GiB	Ready
h4-145	CentOS 7 (64-bit)	8	16 GiB	Ready

Total 11 < 1 2 3 > 5/page

Managed Containers Table (Kubernetes only)

Based on the selected scope (cluster, node, namespace, application, or controller), this table lists the containers for the selected scope. Each container is listed along with its project (namespace), application, Kubernetes pod name, and the node where this container runs.

Container	Project	App	Pod	Node
consumer	myproject	kafka	consumer-...	h17-102
nginx	nginx-prel...	nginx	my-nginx-...	h17-101
nginx-prepared	nginx-prel...	nginx	nginx-prep...	h17-100
nginx-prepared	nginx-prel...	nginx	nginx-prep...	h17-102
nginx	nginx-dem...	nginx-demo	web-2	h17-100

Total 14 < 1 2 3 > 5/page

Workload Prediction Table and Workload Observation and Prediction Charts

The *Workload Prediction* table displays daily, weekly, and monthly predictive CPU and memory data.

The *Workload Observation and Prediction* charts display observed actual usage for the selected time as well as predictive CPU and memory data:

- Daily – Predicts CPU and memory usage every hour for the next 24 hours.
- Weekly – Predicts CPU and memory usage every 6 hours for the next 7 days.
- Monthly – Predicts CPU and memory usage every day for the next 30 days.

Use the *Time Range* field to set a custom time period for observed CPU and memory usage.

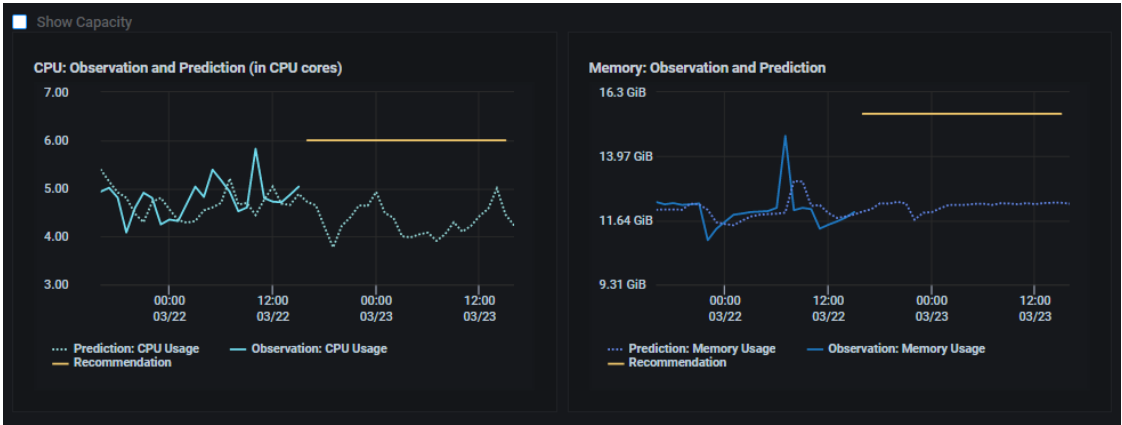
Workload Prediction Table

Average CPU	Minimum CPU	Maximum CPU	Recommended CPU	Average Memory	Minimum Memory	Maximum Memory	Recommended M...
4.34 cores	3.77 cores	5.00 cores	6.00 cores	12.18 GiB	11.67 GiB	12.31 GiB	15.51 GiB

This table displays average/minimum/maximum CPU and memory usage and recommendations for the upcoming time selected - 24 hours (daily), 7 days (weekly), 30 days (monthly).

If you are displaying information for a Kubernetes namespace and the status is anything but *Monitoring*, this section will provide more information. For example, you will see a message for newly added namespaces that do not have enough information for predictions.

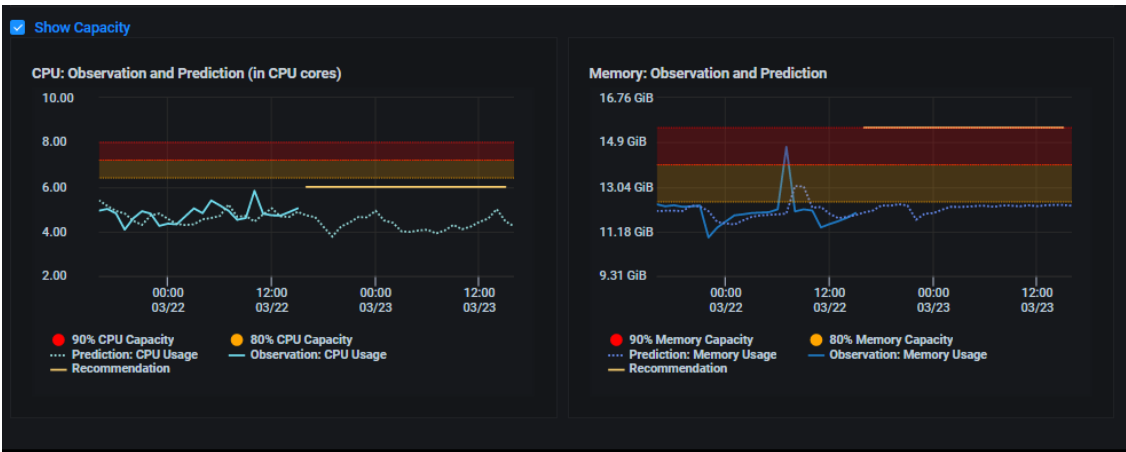
Workload Observation and Prediction Charts



These charts display CPU and memory observations and predictions for all resources specified in the *Filter* panel.

- The solid line represents the observed actual usage.
- The dotted green line represents the past and future predicted usage.
- The solid yellow line represents the recommended usage, which can help you from over-provisioning resources.
- If the selected scope is a Kubernetes node, the solid line represents the node’s total CPU and memory. A big difference between total resources and actual and predicted usage can indicate that you are over-provisioned. A small difference between total resources and actual and predicted usage can indicate that you might be under-provisioned.

Check *Show Capacity* to see the maximum CPU and memory usage limits for the cluster, node, or VM. Orange represents 80-90% and red represents 90-100%. This is a useful way to see if the utilization of resources is approaching the overall capacity.



Related topics:[Terminology](#)[Search/Sort Information in Tables](#)[Show/Hide Information in Charts](#)[Cluster Configuration](#)

Recommendation - HPA Recommendation (Kubernetes only)

The *HPA Recommendation* page displays usage and recommendation information about replicas for selected controllers. These controllers must be enabled with autoscaling during configuration. Refer to the Configuration section for information about how to enable autoscaling for a controller. When autoscaling is enabled, CPU and memory usage is monitored, and the number of pods is increased/decreased based on the workload. An auto-scaled pod is called a *replica*.

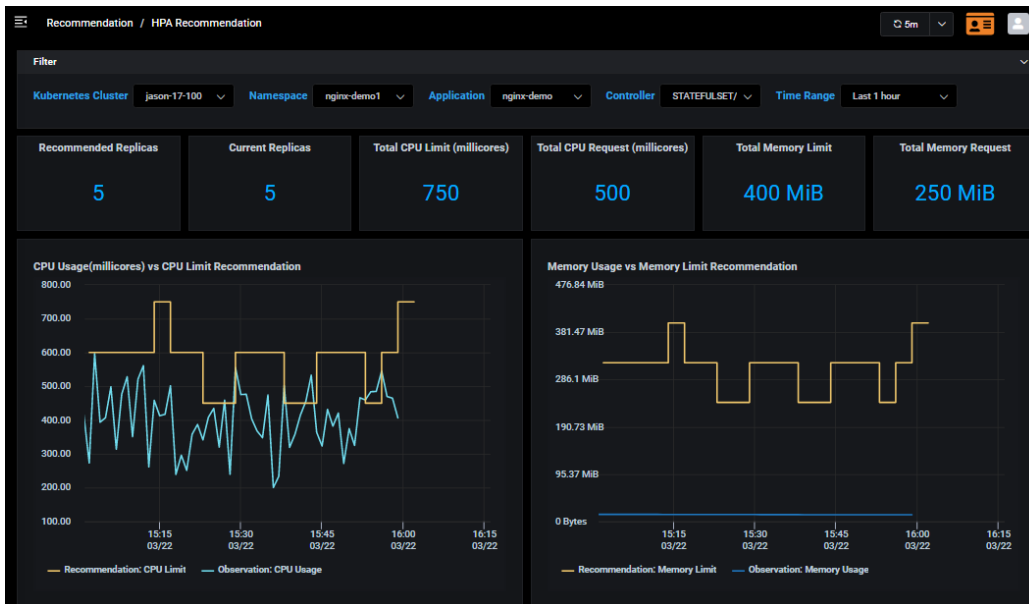
Use the correct namespace, application, and controller name to see the history of autoscaling for a specific controller. The historical number of replicas and CPU/memory usage for the last 1, 2, 4, 6, or 12 hours can be displayed and can further be filtered by selecting a range of dates.

When the system was configured, the requested (minimum) CPU/memory and the limit (maximum) CPU/memory were set. This page shows the number of recommended and current replicas along with the total CPU/memory limit/request being used by all current replicas.

CPU and memory Charts

These charts display CPU and memory usage and recommendations.

- The blue line represents the observed actual usage.
- The yellow line represents the limits after autoscaling.



Related topics:

[Common Administration Portal Functions](#)

[Terminology](#)

[Configuration](#)

[Search/Sort Information in Tables](#)

[Show/Hide Information in Charts](#)

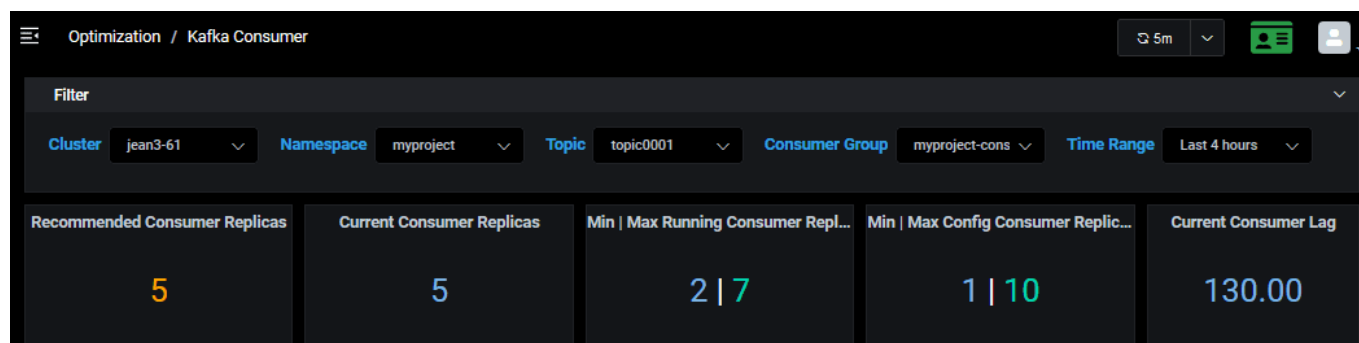
[Configure Applications](#)

Optimization – Kafka Consumer (Kubernetes only)

The traditional method of autoscaling, based on Kafka consumer CPU/memory usage, is not enough to achieve desired performance goals. Kafka production and consumption of messages (including message arrival rate and consumer lag) offers a better indicator of workload and performance, and is monitored by Federator.ai. Using message production rate predictions, Federator.ai auto-scales the number of Kafka consumer pods to fit the workload and optimize performance.

If you are using Kafka software and you have configured Federator.ai to monitor and auto-scale Kafka consumers, the *Kafka* page displays predictions for the message production rate and scales Kafka consumer replicas to satisfy the workload. You can configure multiple Kafka topics and consumer groups. Federator.ai will predict and auto-scale consumers for each individual topic/consumer group. Refer to the [Add an Application](#) section for information about how to configure Kafka consumer monitoring and auto-scaling.

In the *Filter* panel, select the correct cluster, namespace, topic, and consumer group for the Kafka consumer being monitored. The number of replicas and Kafka message production/consumption rate and the prediction of message production rate for the last 1, 2, 4, 6, or 12 hours can be displayed and can further be filtered by selecting a range of dates.



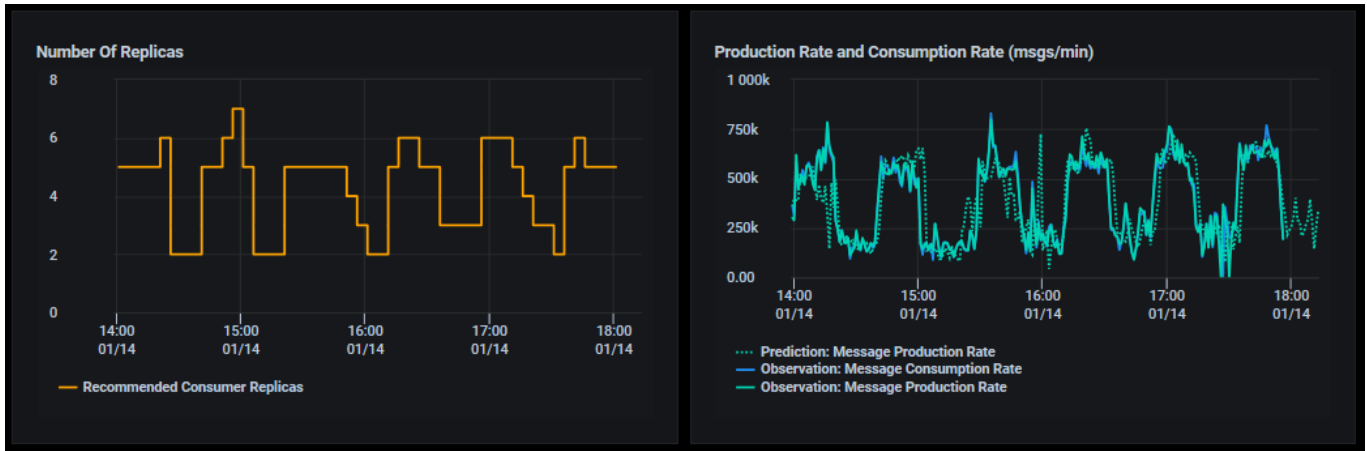
The number of recommended and current consumer replicas is displayed along with the minimum and maximum number of running and configured replicas and the current consumer lag.

Number of Replicas Chart

The *Number of Replicas* chart displays the Kafka consumer replicas for the specified time range as a result of recommendations from Federator.ai.

Production Rate and Consumption Rate Chart

The *Production Rate and Consumption Rate* chart displays the actual observed message production rate (solid green line) and consumption/processed rate (solid blue line) for the specified time range and the historical and future predicted rate (dotted green line).

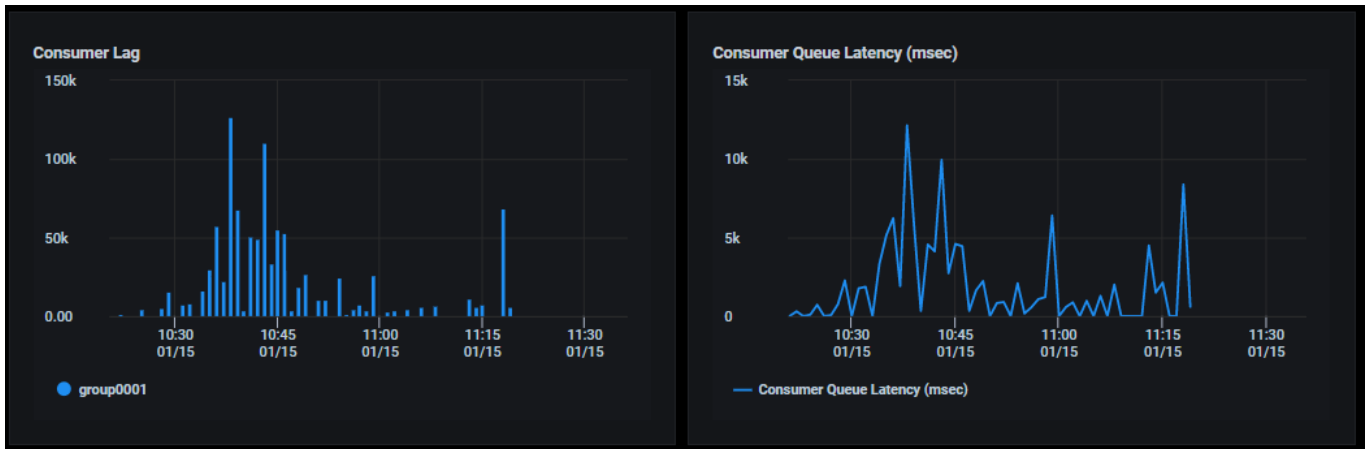


Consumer Lag Chart

The *Consumer Lag* chart displays the number of messages in the Kafka brokers that are yet to be processed by the Kafka consumers for the selected topic and consumer group and time range. This represents messages in the queue waiting to be processed.

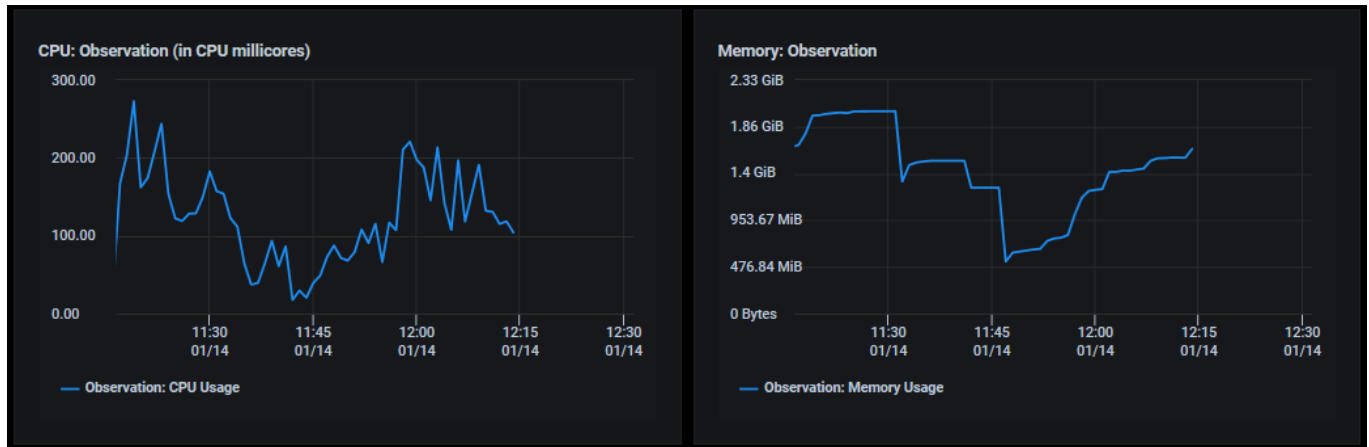
Consumer Queue Latency Chart

The *Consumer Queue Latency* chart displays how long it takes for each message to be processed for the selected time range.



CPU and Memory Observation Charts

The *CPU Observation* and *Memory Observation* charts display observed actual usage for the selected time range. This reflects the resources used as a result of autoscaling consumer pods.



Related topics:

[Common Administration Portal Functions](#)

[Terminology](#)

[Show/Hide Information in Charts](#)

[Setup Wizard](#)

[Configure Applications](#)

Cost – Multi-cloud Cost Analysis (Kubernetes and VM Clusters)

The *Multi-cloud Cost Analysis* page calculates the amount you are spending with your current cloud provider (if applicable) or calculates what you are spending for your local cluster (based on your custom price book) and outlines your estimated cost for doing business with a variety of cloud providers based on fees charged by each provider and recommended changes to your current resource configuration. Federator.ai monitors and analyzes the overall CPU/memory usage of a cluster and predicts the usage for the next 30 days. Based on the past usage and the forecast, Federator.ai recommends the right instance types of cluster nodes for the workload with the lowest cost from each cloud provider.

The screenshot shows the 'Multicloud Cost Analysis' interface for cluster 'ocp4-94'. The filter bar at the top shows 'Cluster: ocp4-94', 'Provider: Each', 'Country: All', and 'Pricing Option: On-Demand'. Below the filter bar, there are four cards representing different cloud providers: AWS (Current), Google Cloud (Recommended), Azure (Recommended), and AWS (Recommended). Each card displays the provider logo, region, current cost, recommended cost, and savings. Below these cards is a table titled 'AWS : Current Cluster Configuration' showing the details of the current cluster nodes.

Name	Role	Instance Type	Region	vCPU	Memory Size	Storage Size	Status
ocp4-qd7hn-master-0	master	m5.4xlarge	us-west-1a	16	62.91 GiB	115.83 GiB	Ready
ocp4-qd7hn-worker-0-wwnc2	worker	m5.4xlarge	us-west-1a	16	62.91 GiB	116.32 GiB	Ready
ocp4-qd7hn-worker-0-v9pbq	worker	m5.4xlarge	us-west-1a	16	62.91 GiB	116.32 GiB	Ready
ocp4-qd7hn-worker-0-2p4nn	worker	m5.4xlarge	us-west-1a	16	62.91 GiB	116.32 GiB	Ready

Total 4

Spending with current cloud provider

The screenshot shows the 'Multicloud Cost Analysis' interface for cluster 'cluster1'. The filter bar at the top shows 'Cluster: cluster1', 'Provider: Each', 'Country: All', and 'Pricing Option: On-Demand'. Below the filter bar, there are four cards representing different cloud providers: Local Cluster (Current), Google Cloud (Recommended), Azure (Recommended), and AWS (Recommended). Each card displays the provider logo, region, current cost, recommended cost, and savings. Below these cards is a table titled 'ONPREMISES : Current Cluster Configuration' showing the details of the current cluster nodes.

Name	OS Type	vCPU	Memory Size	Storage Size	Status
JC-K8s-h17131	CentOS 7 (64-bit)	8	16 GiB	400 GiB	Ready
win7-client	Microsoft Windows 7 (32-bit)	8	16 GiB	100 GiB	Ready
wanem-1	Ubuntu Linux (32-bit)	1	512 MiB	1 GiB	Ready
t1	Ubuntu Linux (64-bit)	2	2 GiB	32 GiB	Ready
r1	Ubuntu Linux (64-bit)	2	4 GiB	32 GiB	Ready

Total 10 < 1 2 > 5/page

Spending for local cluster

You can filter the data that appears on the whole page by selecting:

- Cluster
- Provider – Display data for the:
 - Three lowest costs among all providers.
 - Lowest cost from each provider.
 - Three lowest costs from a selected provider.
- Country – All or a specific country.
- Pricing option - The following pricing options are available:
 - Reserved - Stationary workloads will be served with Reserved instances and temporary, increased workloads will be served with On-Demand instances.
 - On-Demand - All workloads will be served with On-Demand instances.
 - Spot – (Kubernetes only) Evictable workloads will be served with Spot instances while the other workloads will be served with On-Demand instances.
 - Spot + Reserved - (Kubernetes only) Evictable workloads will be served with Spot instances while the other workloads will be served with Reserved and On-Demand instances.

The *Current* box displays what you are currently paying (locally or to your cloud provider) with your existing resource configuration. The current configuration for each cluster node, including Kubernetes role (master, worker), instance types being used at your cloud provider or operating system (OS) for your local cluster, cloud provider region, number of CPUs, memory size, and storage size is displayed in the table below.

The *Recommended* boxes display potential savings with different cloud providers or different regions for your current provider. Click the *Detail* button to see how these savings are calculated. The savings typically come from lower provider fees, as well as from reducing idle resources and using more cost-effective instance types with the provider. The information continually refreshes itself as new data becomes available.

While it is difficult to move between cloud providers, this information can help you reduce costs by changing instance types, reducing idle resources, and selecting a different region from your current provider.

Even if you are not currently using a cloud provider, this information can show you what your options are if you move to the cloud.

CPU and Memory Utilization Charts

The *CPU Utilization* and *Memory Utilization* charts display utilization information for all nodes/VMs in the selected cluster based on your daily, weekly, or monthly workload.

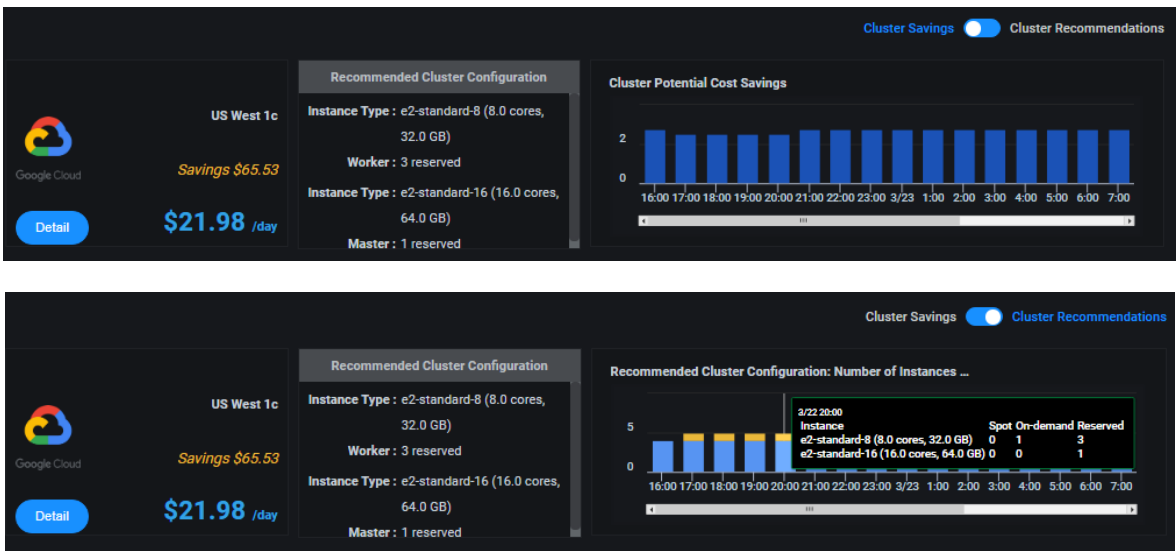


The solid lines represent the observed actual usage while the dotted lines show the historical and future predicted usage. The predicted usage is used in the analysis for cluster configuration recommendations.

If there are evictable Kubernetes applications in this cluster, you will also see their actual and predicted usage.

Recommended Cluster Configuration

The recommended configuration with each provider is displayed based on your daily, weekly, or monthly workload prediction (e.g., if *Daily* is selected for the *CPU Utilization* and *Memory Utilization* charts, the recommendation and saving are calculated based on the workload prediction for the next 24 hours). This includes the recommended instance types (scroll down in the table to see all instance types), as well as the number of Kubernetes master and worker servers for the selected pricing option.



Use the toggle to display *Cluster Savings* or *Cluster Recommendations*.

Cluster Savings shows how much you can save by following the recommended configuration.

Cluster Recommendations displays a *Time Series* chart that shows the recommended instances (number and type) for future workloads at specific times and can help you determine how many resources to reserve. Place your cursor over one of the bars to see the recommended number of instances. The recommendations for each data point include the number of on-demand, reserved, or spot instances (for Kubernetes), depending on the pricing option selected.

For daily predictions, data is displayed in the charts for each hour. For weekly predictions, data is displayed for every six hours. For monthly predictions, data is displayed for each day.

Related topics:

[VM Cost Analysis](#)

[Application Cost Analysis](#)

[Cost Allocation Kubernetes Namespaces](#)

[Common Administration Portal Functions](#)

[Terminology](#)

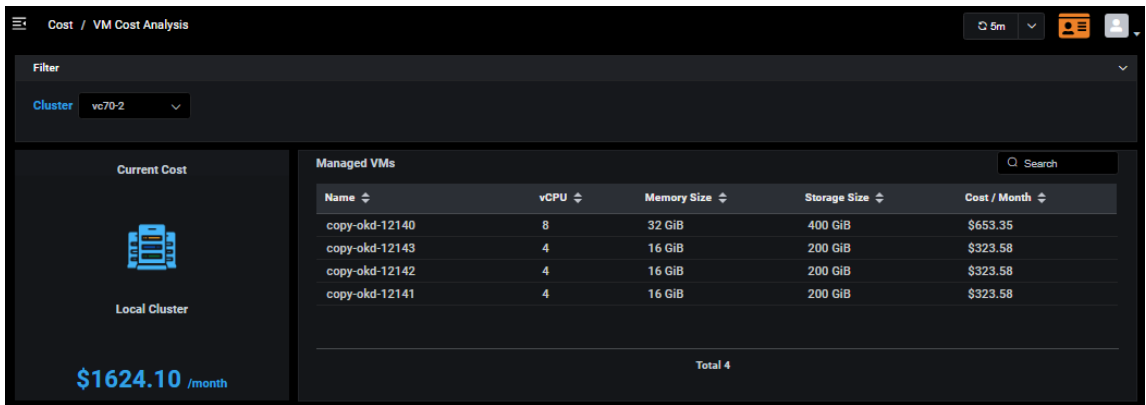
[Search/Sort Information in Tables](#)

[Show/Hide Information in Charts](#)

[Price Books](#)

Cost – VM Cost Analysis (VM Clusters and VMs)

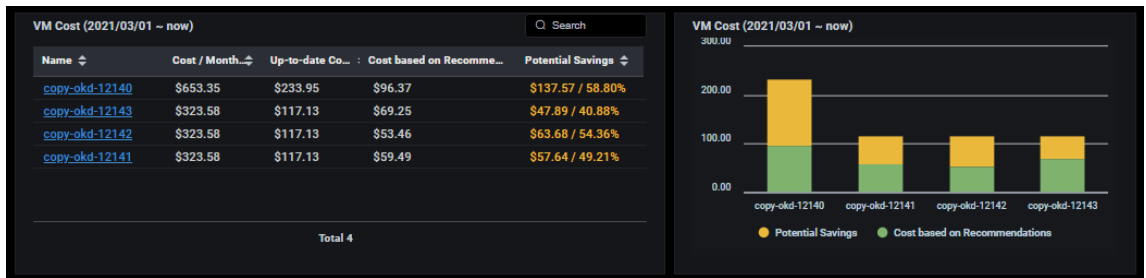
The *VM Cost Analysis* page displays your cost for an individual cluster and for each managed VM for the current month and shows your potential cost/savings based on usage recommendations for the next 30 days.



The *Current Cost* box displays the current monthly cost for the cluster and the *Managed VMs* table displays a list of managed VM in the cluster with the number of CPU cores, memory size, storage size, and monthly cost for each VM.

VM Cost

This section shows the usage and cost for each VM along with potential savings based on recommendations. The recommendations used in the calculations can be viewed on the *Planning / VM Workload Prediction* page. Be sure to check *Show Capacity* to see the maximum CPU and memory usage limits for the VM, which shows if resources are over-provisioned.



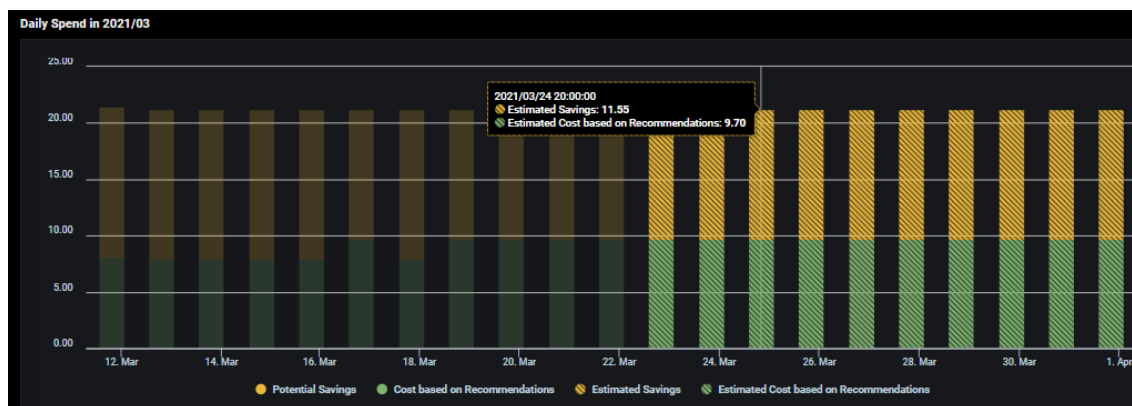
The table on the left shows the full monthly cost, the cost from the beginning of the current month to now, what that cost would be based on the recommendations, and the potential savings for each VM. Click on a VM to update the charts below.

The bar chart on the right presents a graphical view of the cost based on the recommendations and the potential savings from the beginning of the current month to now. The green area of each bar represents your monthly cost based on the recommendations and the yellow represents your estimated monthly potential savings based on the recommendations.

Daily Spend Chart

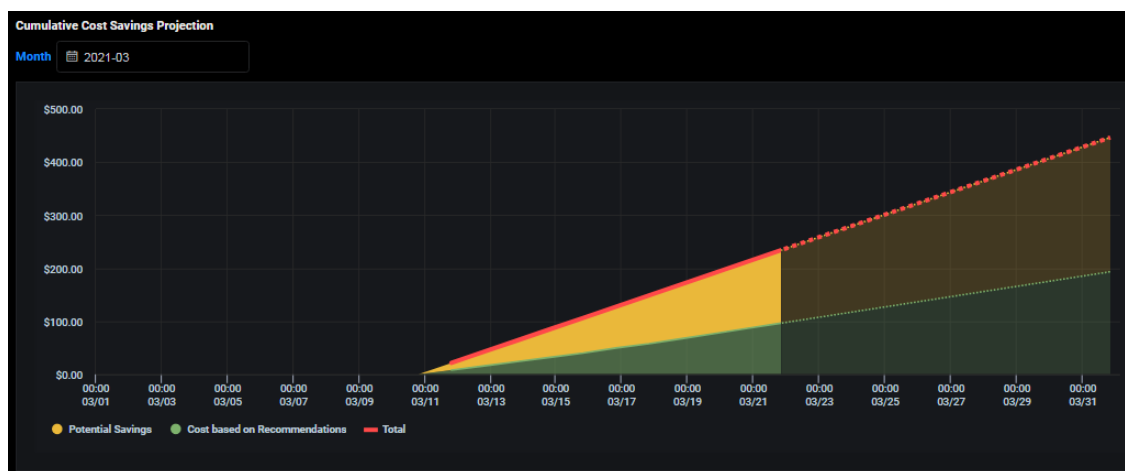
This chart displays your daily spending for the selected VM from the start of the month to end of the month. The green area of each bar represents your daily cost based on the recommendations (actual or estimated for the future) and the yellow represents your daily potential savings based on the recommendations.

Move your cursor to a specific day to see detailed information.



Cumulative Cost Savings Projection Chart

This chart displays the cumulative savings for the VM selected from the start of the month and going forward. The default is the current month.



The green section represents the cost based on the recommendations, the yellow section represents estimated potential savings, and the red dotted line represents the total.

Related topics:

[Multicloud Cost Analysis](#)

[Workload Prediction Page](#)

[Workload Observations and Predictions](#)

[Common Administration Portal Functions](#)

[Terminology](#)

[Search/Sort Information in Tables](#)

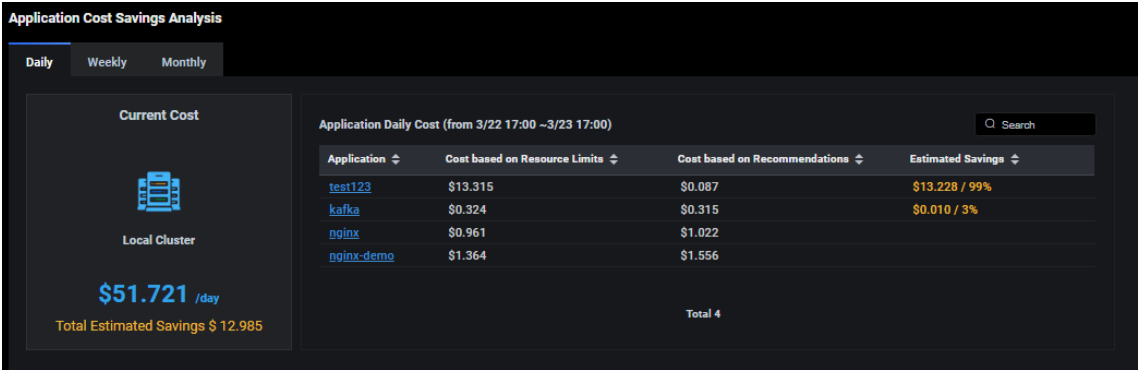
[Show/Hide Information in Charts](#)

[Price Books](#)

Cost – Application Cost Analysis (Kubernetes Applications)

The *Application Cost Analysis* page displays your cost for each monitored Kubernetes application in a selected cluster and shows your potential cost/savings based on usage recommendations. The recommendations used in the calculations can be viewed on the *Planning / Kubernetes Workload Prediction* page.

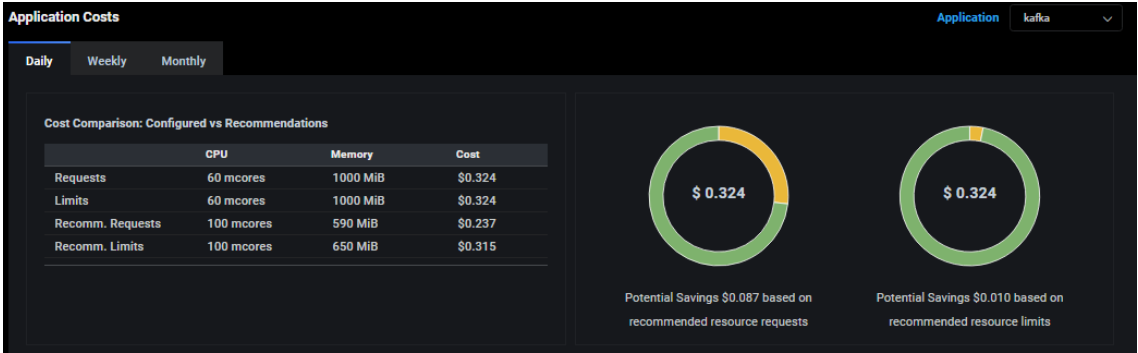
Application Cost Savings Analysis



The *Current Cost* box displays the current monthly cost for the cluster along with the savings for all applications. The *Application Daily Cost* table displays the cost based on current resource limits, what that cost would be based on the recommendations, and the potential savings for each application in the cluster. Click on an application to update the charts below.

Information can be viewed daily, weekly, or monthly.

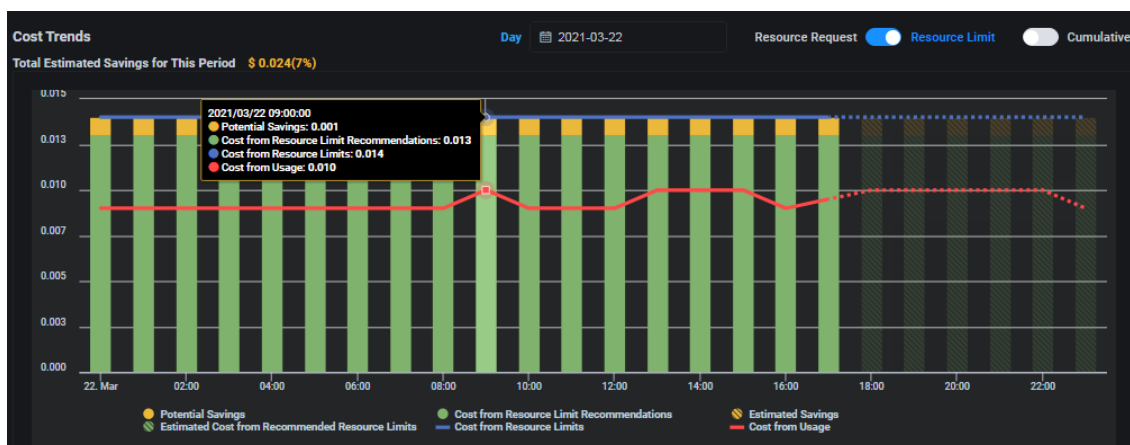
Application Costs



Select an application and day/week/month to display a cost comparison between your current Kubernetes application CPU and memory requests and limits and the recommended requests and limits.

The charts on the right present a graphical view of the difference between the current and recommended requests and limits as well as the estimated savings for the selected application. Move your cursor over the chart to see detailed information.

Cost Trends



Use the sliders to see requests or limits for the current selected time period (day/week/month) or cumulatively through the end of the selected time period. The total savings for the period are shown above the chart.

The green area of each bar represents your cost based on the recommendations (actual or estimated for the future) and the yellow represents your potential savings based on the recommendations.

Move your cursor over the chart to see detailed information for that data point.

Related topics:

[Multicloud Cost Analysis](#)

[Cost Allocation Kubernetes Namespaces](#)

[Workload Prediction Page](#)

[Workload Observations and Predictions](#)

[Common Administration Portal Functions](#)

[Terminology](#)

[Search/Sort Information in Tables](#)

[Show/Hide Information in Charts](#)

[Price Books](#)

Cost - Cost Allocation (Kubernetes Namespaces)

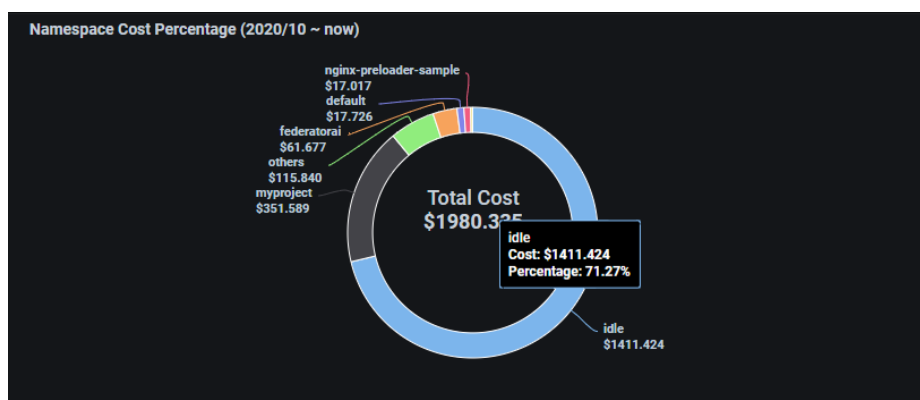
The *Cost Allocation* page provides an in-depth cost analysis for individual Kubernetes namespaces of a cluster hosted by the current cloud service provider. With the information displayed on this page, you can compare the cost of each namespace and determine the estimated cost of each namespace based on the workload prediction. You can also see the percentage of cost of the system spent while the system is idle. A high percentage of cost while the system is idle indicates that the cluster is over-provisioned.

Current Cluster Configuration

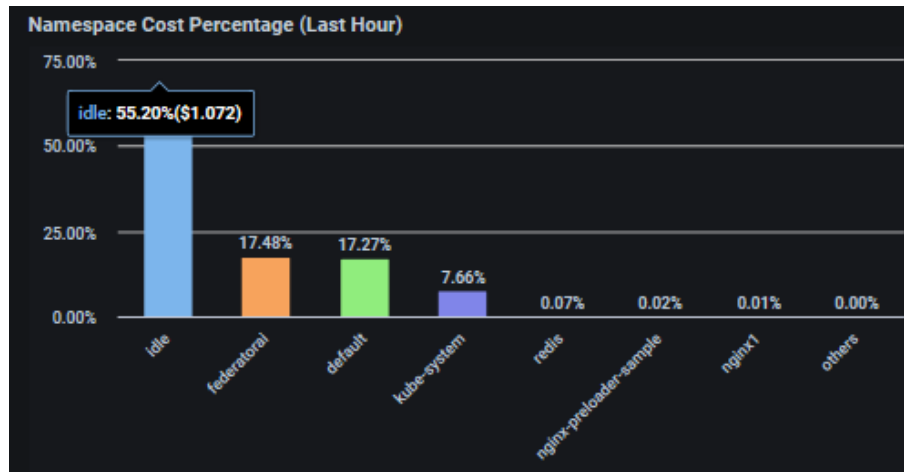
The existing configuration with your provider is displayed. This includes the role (master, worker), instance types being used, region, number of CPUs, memory size, and storage space.

Namespace Charts

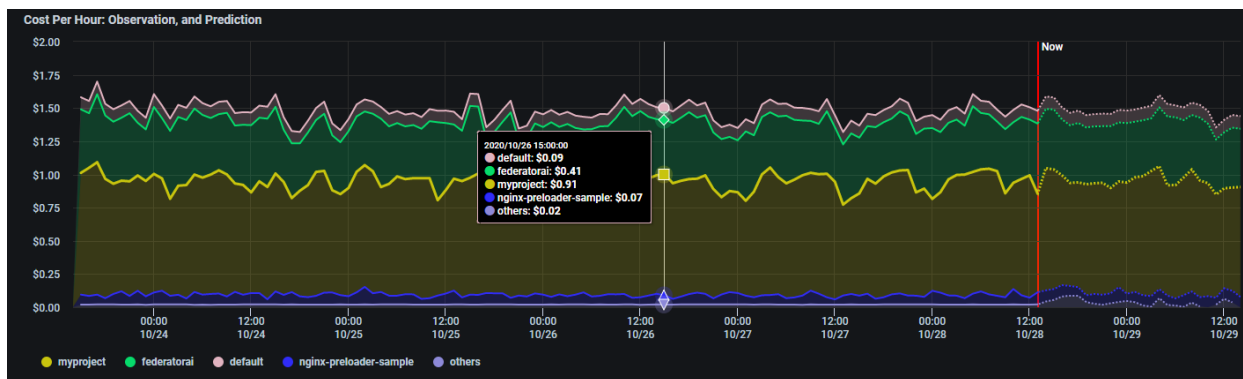
The *Namespace Cost Percentage (month)* chart displays this month's actual percentage of cost to-date for each namespace and when the system is idle. The *others* tag represents the unmonitored namespaces in a cluster. Move your cursor over each section to show the percentage of total cost for a namespace or when the system is idle.



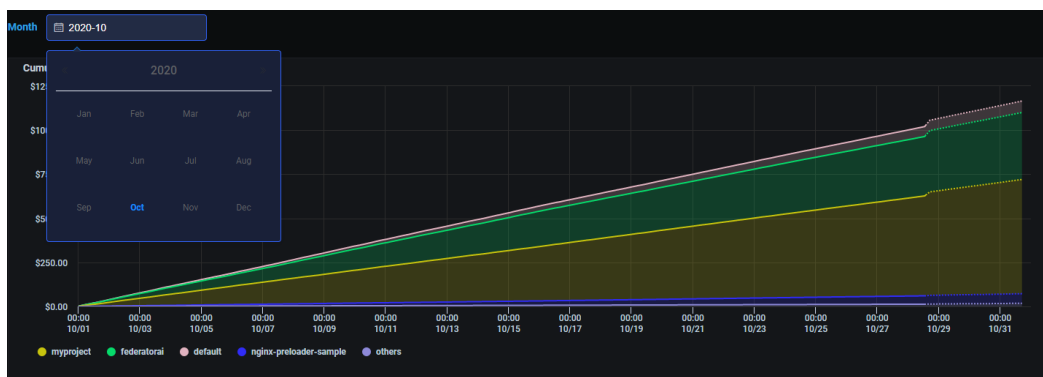
The *Namespace Cost Percentage (Last Hour)* chart displays the actual cost for each namespace and when system was idle during the past hour. The *others* tag represents the unmonitored namespaces in a cluster. Move your cursor over each bar to show the cost of a namespace or when system is idle.



The *Cost Per Hour* chart displays hourly costs for each namespace during the past five days and the predicted hourly cost for the next 24 hours. The solid lines represent the observed actual cost while the dotted line shows the future predicted cost based on the namespace resource usage prediction. The *others* tag represents the unmonitored namespaces in a cluster. Move your cursor to a specific date/hour to see the detailed cost of each namespace.



The *Cumulative Cost Projection* chart displays a cumulative cost from the beginning of the month and a cost projection for this month for each namespace. The solid lines represent the observed actual cost while the dotted lines show the future predicted cost based on the namespace resource usage prediction. The *others* tag represents the unmonitored namespaces in a cluster. Click on the calendar icon to choose the cost analysis for a specific month.



Related topics:

[Multicloud Cost Analysis](#)

[Application Cost Analysis](#)

[Common Administration Portal Functions](#)

[Terminology](#)

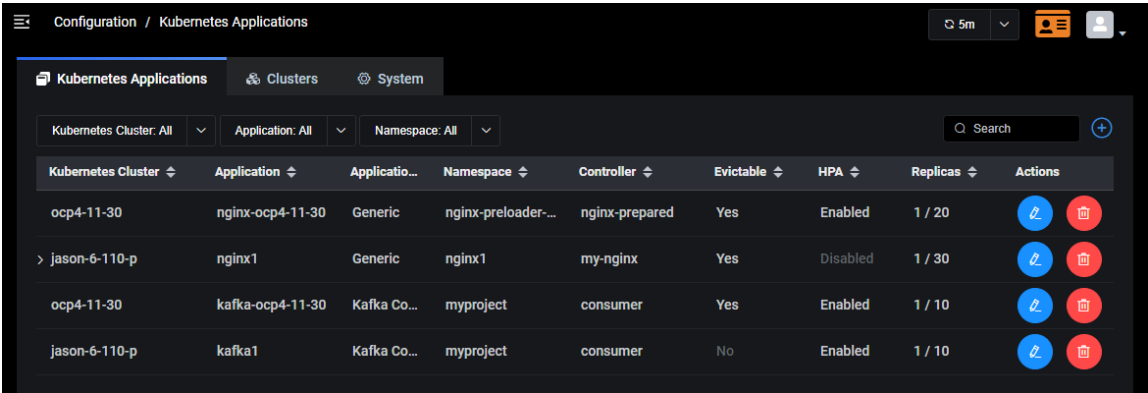
[Search/Sort Information in Tables](#)

[Show/Hide Information in Charts](#)




[Price Books](#)

Configuration – Kubernetes Applications

The *Kubernetes Applications* page displays all Kubernetes applications being monitored by Federator.ai and allows you to manage them.



You can perform the following functions:

Icon	Function
	Add an application.
	Edit an application.
	Remove an application.

Add an Application

1. On the *Configuration / Kubernetes Applications* page, click the + (Add Application) icon.

A screenshot of the 'Add Application' form. It has a title bar 'Add Application' with a close button. The form contains three fields: '* Application Name' (a text input field with a yellow warning icon), '* Cluster' (a dropdown menu with 'Select' and a downward arrow), and 'Instance Type' (a toggle switch with 'Generic' and 'Kafka' options). At the bottom right, there are 'Cancel' and 'Save' buttons.

Application Name - Name of the application you want to manage. The name must be a maximum of 253 characters, contain only lowercase alphanumeric characters, “-”, or “.”, and start and end with an alphanumeric character. Note, an application is not a native Kubernetes object. You will define the controllers that are part of your application later.

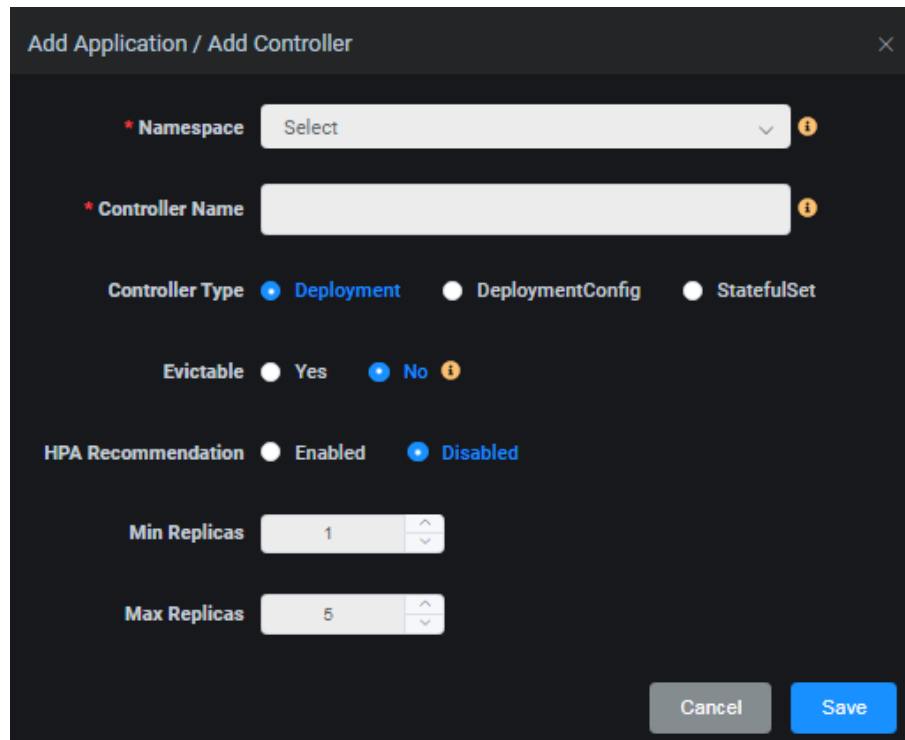
Cluster – Select an existing cluster where this application resides.

Instance Type - For a *Generic* application, CPU, memory, and usage are monitored. For scaling Kafka consumers, choose *Kafka Consumer* as the instance type

2. Click the + icon to add controllers (generic) or consumer groups (Kafka Consumer).

The options will vary, depending upon the selected *Instance Type*.

Generic:



The screenshot shows a dark-themed dialog box titled "Add Application / Add Controller" with a close button (X) in the top right corner. The form contains the following fields and options:

- Namespace:** A dropdown menu with "Select" as the placeholder text and an information icon (i) to the right.
- Controller Name:** A text input field with an information icon (i) to the right.
- Controller Type:** Three radio button options: "Deployment" (selected with a blue dot), "DeploymentConfig", and "StatefulSet".
- Evictable:** Two radio button options: "Yes" and "No" (selected with a blue dot). An information icon (i) is to the right.
- HPA Recommendation:** Two radio button options: "Enabled" and "Disabled" (selected with a blue dot).
- Min Replicas:** A numeric input field with a value of "1" and up/down arrow buttons.
- Max Replicas:** A numeric input field with a value of "5" and up/down arrow buttons.
- Buttons:** "Cancel" and "Save" buttons at the bottom right.

- *Namespace* – Select the Kubernetes namespace where the controller is deployed.
- *Controller Name* – Specify the name of controller to be monitored.
- *Controller Type* – Supported controller types are *Deployment*, *StatefulSet*, and *DeploymentConfig* (OpenShift only).
- *Evictable* – Indicate if the controller can be interrupted if the node is shut down. Evictable controllers are good candidates to be deployed in Spot instances.
- *HPA Recommendation* – Indicate if you want to enable Horizontal Pod Autoscaling (HPA). When enabled, CPU and memory usage is monitored, and the number of pods is automatically increased/decreased based on the CPU/memory usage workload.
- *Mix/Max Replicas* – Specify the minimum and maximum number of pods when HPA is enabled.

Kafka Consumer:

Add Application / Add Consumer Group

* Kafka Broker Namespace

* Consumer Group Namespace

* Consumer Group Name

* Kafka Consumer ID

Controller Type ☒ Deployment ☐ DeploymentConfig ☐ StatefulSet

Evictable ☐ Yes ☒ No

HPA Recommendation ☐ Enabled ☒ Disabled

* Topic Name

Min Replicas

Max Replicas

Cancel Save




- *Kafka Broker Namespace* – Select the namespace for the Kafka broker that receives and stores messages from producers and allows consumers to fetch the messages.
- *Consumer Group Namespace* – Select the namespace of the consumer group used for checking the offset on topics and processing messages.
- *Consumer Group Name* – Specify the name of the Kafka consumer group used for checking the offset on topics and processing messages.
- *Kafka Consumer ID* – Specify the unique ID of the Kafka consumer group for the set of consumers within the same consumer group.
- *Controller Type* – Specify the controller type of the consumer group. Supported controller types are *Deployment*, *StatefulSet*, and *DeploymentConfig* (OpenShift only).
- *Evictable* – Indicate if the controller can be interrupted if the node is shut down. Evictable controllers are good candidates to be deployed in Spot instances.
- *HPA Recommendation* – Indicate if you want to enable Horizontal Pod Autoscaling (HPA). When enabled, the Kafka message production rate is monitored, and the number of Kafka consumer pods is automatically increased/decreased based on the Kafka message production rate.

- *Topic Name* – Indicate the topic (where records are stored and published) that will be processed by consumers.
 - *Mix/Max Replicas* – Specify the minimum and maximum number of consumers to be auto-scaled.
3. Click *Save*.
 4. Continue adding controllers/consumer groups and click *Save* when done.

Manage Applications

You can do the following from the *Kubernetes Applications* page:

- Add, edit, or remove a controller/consumer group. To do this, click the *Edit Application* icon and click the appropriate icon:

Icon	Function
	Add a controller/consumer group.
	Edit a controller/consumer group.
	Remove a controller/consumer group.

- Remove an application. To do this, click the *Remove Application* icon and confirm the removal.

Related topics:

[Terminology](#)

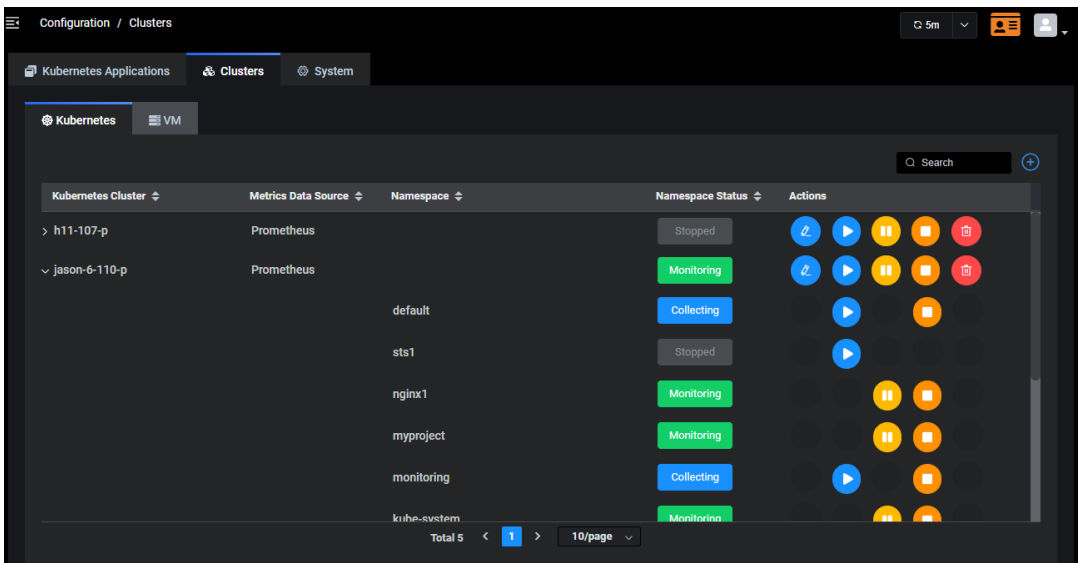
[Search/Sort Information in Tables](#)

Configuration - Kubernetes Clusters







The *Clusters* page displays all Kubernetes and VM clusters being monitored by Federator.ai. Select Kubernetes tab to see the list of monitored Kubernetes clusters. Expand a cluster to see the namespaces for a cluster.

You can also see namespace status for the cluster and individual namespaces. The status will be *Monitoring* when the system is collecting metrics and providing workload predictions. The status will be *Collecting* when all the namespaces in the cluster are in collecting state. When a namespace is in collecting state, the system will still collect metrics, but prediction tasks are paused. A namespace will also collect data if it was added into a cluster that does not automatically monitor for new namespaces.

A status of *Stopped* means there is no collecting of metrics and no predictions. A cluster is *Stopped* if all of its namespaces are stopped.

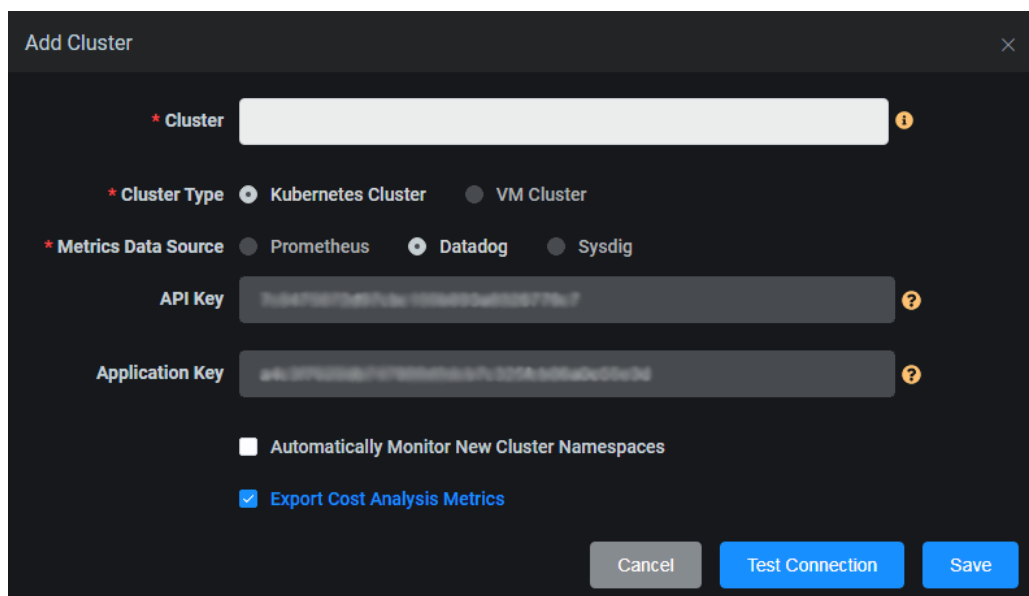


Functions are different for clusters and namespaces. You can perform the following functions:

Icon	Function
	Add a cluster.
	Edit settings for monitoring new namespaces and exporting cost analysis metrics.
	Start monitoring and prediction for all namespaces or a specific namespace.
	Pause monitoring and prediction for all namespaces or a specific namespace.
	Stop collecting metrics and making predictions for all namespaces or a specific namespace.
	Remove a cluster.

Add a Cluster

1. On the *Configuration / Clusters* page, click the + (Add Cluster) icon.



The 'Add Cluster' dialog box is shown with a dark background. It contains the following fields and options:

- * Cluster**: A text input field with a placeholder and an information icon.
- * Cluster Type**: Two radio buttons, 'Kubernetes Cluster' (selected) and 'VM Cluster'.
- * Metrics Data Source**: Three radio buttons, 'Prometheus', 'Datadog' (selected), and 'Sysdig'.
- API Key**: A text input field with a placeholder and a question mark icon.
- Application Key**: A text input field with a placeholder and a question mark icon.
- ☐ **Automatically Monitor New Cluster Namespaces**
- ☒ **Export Cost Analysis Metrics**
- Buttons at the bottom: 'Cancel', 'Test Connection', and 'Save'.

Cluster – Specify the name of a cluster to be managed. *Cluster Type* – By default, *Kubernetes Cluster* is selected for you. You can specify *VM Cluster* if you want to add a VM cluster.

Metrics Data Source – Specify the source of metrics for this cluster.

API Key/Application Key – For Datadog, the API key and application key are required for authentication. By default, they are set as Datadog API key and application key from the *Metrics Data Source* tab under *Configuration → System*. However, they can be specified with a different API key and an application key.

URL/Token – For Sysdig, a URL and token are required for authentication. For the Prometheus open-source monitoring system, the URL is required but the token is optional. By default, they are set as Sysdig URL/token or Prometheus URL from the *Metrics Data Source* tab under *Configuration → System*. However, they can be specified with different values.

Automatically Monitor New Cluster Namespaces – Specify if you want to automatically monitor new namespaces in this cluster. If unchecked (default), you must manually specify which namespaces to monitor.

Export Cost Analysis Metrics – Specify if you want cost analysis metrics to be automatically exported to Datadog, so that the information can appear in Datadog's user interface.

2. Click *Test Connection* to confirm that all information is correct.
3. Click *Save* when the system can connect to the cluster.

Manage Clusters

You can do the following from the *Configuration / Clusters* page:

- Edit the settings for monitoring new namespaces and exporting cost analysis metrics. To do this, click the *Edit Cluster* icon.
- Start monitoring and prediction for all namespaces in the cluster or a specific namespace. To do this, click the *Start Monitoring and Predictions* icon.
- Pause monitoring and prediction for all namespaces or a specific namespace. To do this, click the *Pause Monitoring and Predictions* icon for a cluster or for a namespace.
- Stop collecting metrics and making predictions for all namespaces or a specific namespace. To do this, click the *Stop Collecting Metrics and Predictions* icon for a cluster or for a namespace.
- Remove a cluster that does not have any applications configured. To do this, click the *Remove Cluster* icon.

Related topic:

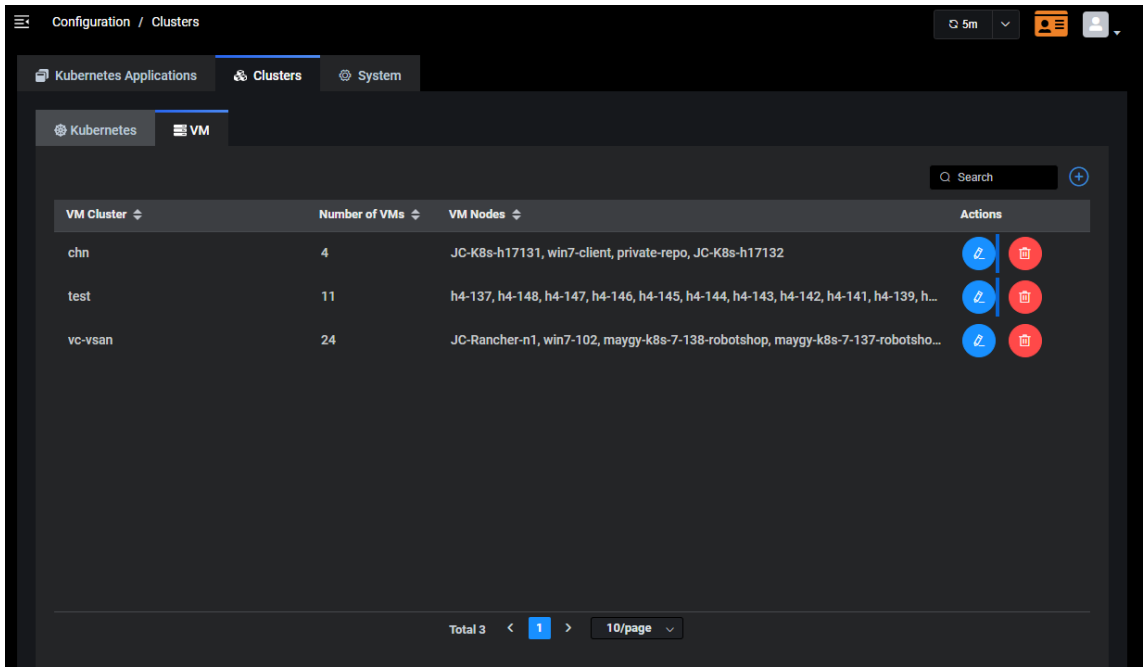
[Terminology](#)

[Search/Sort Information in Tables](#)

[Configure Applications](#)

Configuration - VM Clusters

The *Clusters* page displays Kubernetes clusters VMware clusters being monitored by Federator.ai. Select VM tab to see the list of VM clusters and the number and names of the virtual machines (VMs) in each cluster.

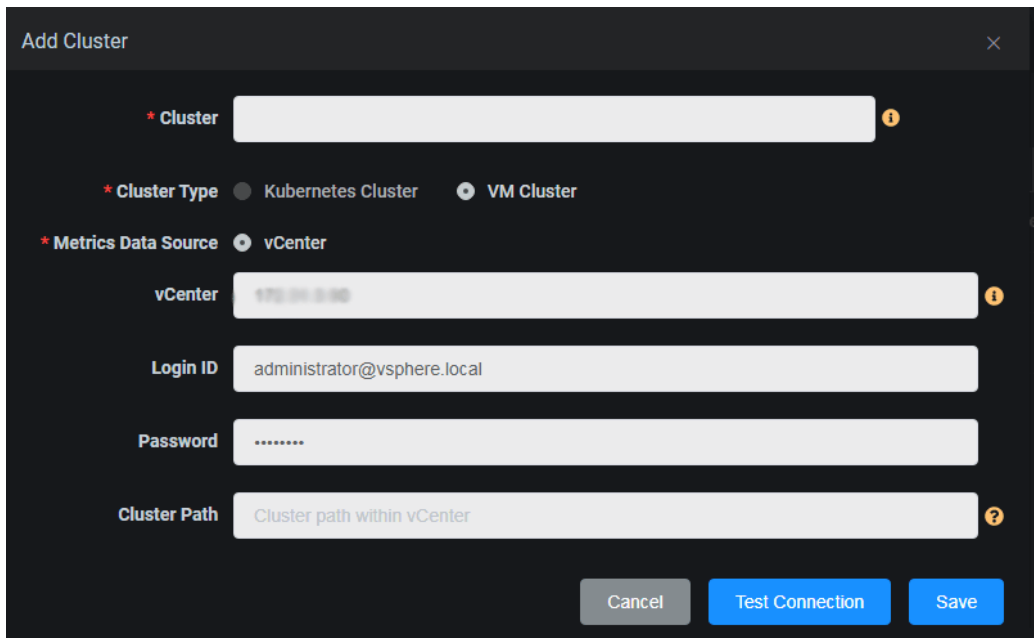


You can perform the following functions:

Icon	Function
	Add a cluster.
	Add/remove the cluster nodes to monitor.
	Remove a cluster.

Add a Cluster

1. On the *Configuration / Clusters* page, click the + (Add Cluster) icon.



Cluster – Specify name of the VM cluster to be managed. There is a maximum of 253 lowercase characters, "-", or "." allowed. The name must start and end with an alphanumeric character.

Cluster Type – By default, *VM Cluster* is selected for you. You can change to *Kubernetes* if you want to add a Kuberene cluster.

Metrics Data Source – Select the source of metric data source for this cluster. If this is a VM cluster, *VCenter* is selected for you.

vCenter – Specify the vCenter IP address. You can have multiple vCenters in your system.

Login ID and Password – Specify the login credentials.

Cluster Path – Specify the path to the cluster, within vCenter. If needed, you can click on the link to the vCenter website, which is included in the popup help text.

2. Click *Test Connection* to confirm that all information is correct.
3. Click *Save* when the system can connect to the cluster.

The system will discover the VMs in this cluster, which may take a few minutes to complete. Use the *Edit Cluster* function to select which nodes that you want to monitor.

Manage Clusters

You can do the following from the *Configuration / Clusters* page:

- Add/remove nodes to monitor. To do this, click the *Edit Cluster* icon. Select one, multiple, or all non-members to add or members to remove.

The screenshot shows the 'Edit Cluster' dialog box with the following fields and options:

- Cluster:** A text field containing 'vc55' with an information icon.
- Cluster Type:** Radio buttons for 'Kubernetes Cluster' and 'VM Cluster' (selected).
- Metrics Data Source:** Radio buttons for 'vCenter' (selected) and another option.
- vCenter:** A text field containing '192.168.1.100' with an information icon.
- Login ID:** A text field containing 'administrator@vsphere.local'.
- Password:** A password field with masked characters.
- Cluster Path:** A text field containing 'New Cluster/192.168.1.100' with a help icon.
- Non-Members:** A section with a '0/0' count and a 'No data' message.
- Members:** A section with a '0/4' count and a list of four nodes: 'Bella-StellarFlash-11-213', 'Bella-StellarFlash-11-214', 'K8S-12-41', and 'windows_2k8r2_11_218'. Each node has an unchecked checkbox.
- Buttons:** 'Add >', '< Remove', 'Cancel', 'Test Connection', and 'Save'.

- Remove a cluster. To do this, click the *Remove Cluster* icon.

Configuration - System

The *System* page has tabs that allow you to:

- Change the admin password.
- Update metrics data source information.
- Manage the system license.
- Set the policy/update price books.

Admin Password

To access this page, select *Configuration, System, Admin Password* tab. The *Admin Password* page allows you to change the admin password.

You must know the current password and *New Password* must match *Confirm Password*.

Metrics Data Source

The *Metrics Data Source* page allows you to set default values of authentication needed to access metrics from different data sources (e.g. Prometheus, Datadog, Sysdig, or VMware Vcenter). To access this page, select *Configuration, System, Metrics Data Source* tab. For Kubernetes, the available data sources are Prometheus, Datadog, and Sysdig For VM cluster, the available data source is VMware vCenter.





- For Datadog, the *API Key* and *Application Key* are required for authentication.
- For Sysdig, a *URL* and *Token* are required for authentication.
- For the Prometheus open-source monitoring system, the *URL* is required but the *Token* is optional.
- For VMware vCenter, a login ID and password are required for authentication. A cluster path is needed for identifying VM's to be managed.

For VMware, the *Metrics Data Source* page allows you to modify cluster information if there are no VMs being monitored for the cluster.

Licenses

To access this page, select *Configuration, System, Licenses* tab. The *Licenses* page displays your current system license, including license type, status, and expiration date. It also shows what is included in your license.

You can perform the following functions:

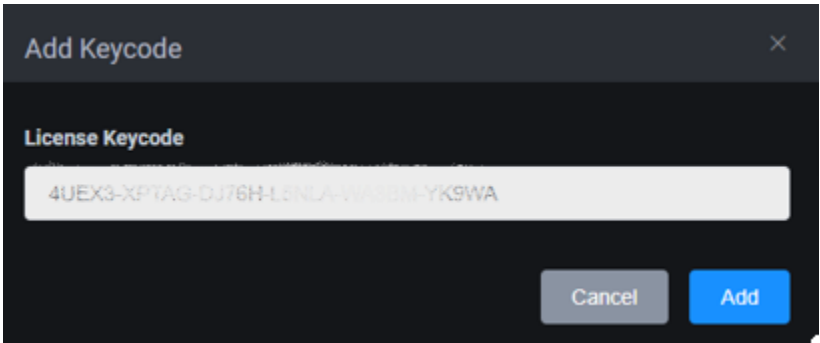
Icon	Function
	Add a license.
	Activate a license.
	Show license keycode.
	Remove a license.

Add a License

There is one license per system; you must delete the existing license before adding a new one.

Follow the steps below to add and register a license:

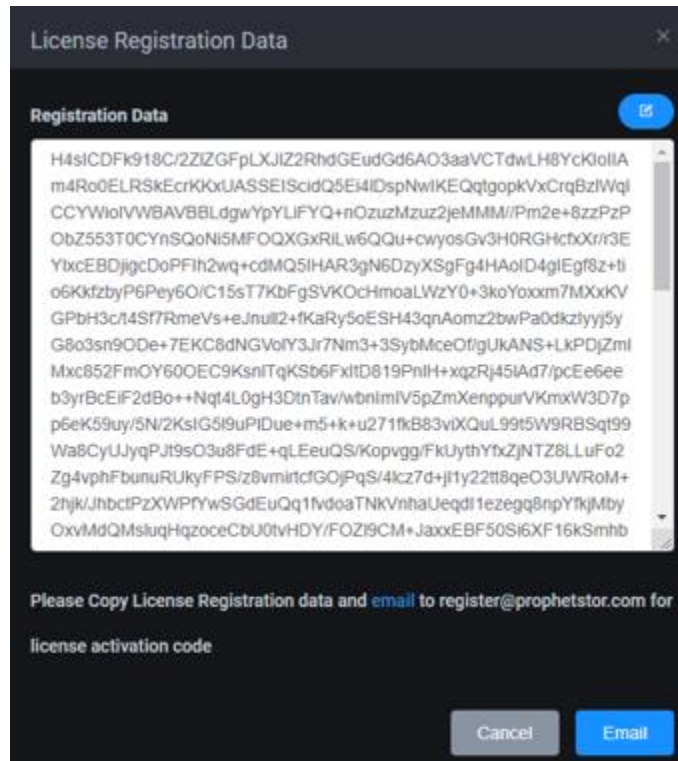
1. Click the *Add License* icon and enter the keycode.



2. Click *Add*.

A trial license is valid immediately after it is added. A regular license needs to be registered in order to be activated.

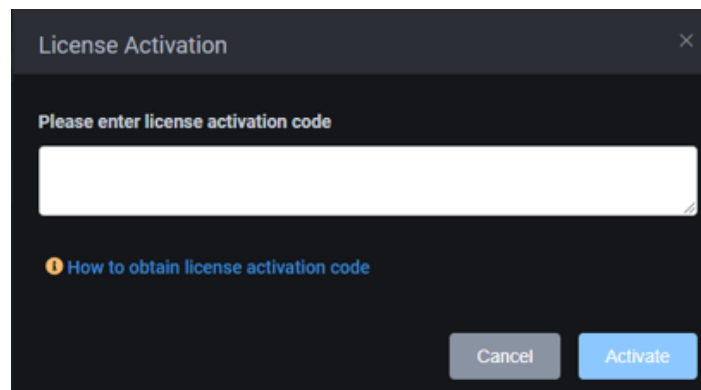
3. Email the registration data to register@prophetstor.com for a license activation code.



Click the *Email* button to launch your email program. Click the *Copy* icon to copy the registration text and paste it into your email before sending.

The license status will now say *Pending Activation*. It must be activated within 30 days.

4. Once you have received the license activation code, click the *Activate License* icon and paste the activation code.



5. Click *Activate*.

Manage Licenses

You can do the following from the *Licenses* page:

- View the keycode for your license. To do this, click the *Show License Keycode* icon.
- Remove the license. To do this, click the *Remove License* icon and confirm the removal.

Price Books

To access this page, select *Configuration, System, Price Books* tab.

Public Cloud Price Books

The *Public Cloud Price Books* page displays the versions of the on-demand and spot instance price books that have been collected from the following cloud providers: Amazon Web Services (AWS), Google Cloud, and Microsoft Azure.

The screenshot shows the 'Price Books' configuration page. On the left, there is a sidebar with 'Public Cloud Price Books' and 'Custom Price Book'. The main area is divided into two columns: 'On-Demand Instances' and 'Spot Instances'. Each column displays the current 'Version' and the 'Latest Version' of the price book. Below these columns, there is a 'Price Book Update Policy' section with radio buttons for 'Auto' (selected) and 'Manual'. At the bottom right, there are 'Cancel' and 'Save' buttons.

Category	Version	Latest Version
On-Demand Instances	v4.5.20210306T190021Z	v4.5.20210306T190021Z
Spot Instances	v4.5.20210309T164519Z	v4.5.20210309T164519Z

Price Book Update Policy: ☒ Auto ☐ Manual

The version numbers represent the ProphetStor version of the compiled price books.

The *Price Book Update Policy* allows you to specify how you want the price books updated. *Auto* regularly checks availability and automatically downloads the latest version, which requires Internet access.

If *Manual* is selected and the *Latest Version* is different from *Version*, Federator.ai is aware of a newer price book but it has not been applied. Click the *Update Price Books* button to update.

Note that the browser needs Internet access in order to display the *Latest Version*. If your computer does not have access, the data must be pushed to system.

Custom Price Book

The *Custom Price Book* page allows you to define your hourly operating costs for CPU, memory, and storage and use these numbers for calculating costs/savings instead of using cloud provider pricing.

The screenshot shows the 'Custom Price Book' configuration page. On the left, there is a sidebar with 'Public Cloud Price Books' and 'Custom Price Book'. The main area has three rows for defining costs: 'CPU', 'Memory', and 'Storage'. Each row has a text input field for the cost in dollars and a unit dropdown menu. At the bottom, there are 'Cancel' and 'Save' buttons.

Resource	Cost (\$)	Unit
CPU	0.063222	/ vCPU hour
Memory	0.008474	/ GB hour
Storage	0.00027396	/ GB hour

When determining your hourly costs, be sure to include electricity, cooling/heating, labor, hardware, etc.

Events

The *Events* page displays all system events that have occurred.

There are five levels of events:

- Fatal - Issues that may stop the system from operating properly.
- Error - Indicates that a failure has occurred.
- Warning - Indicates that something occurred that may require maintenance or corrective action; however, the system is still operational.
- Info - Day-to-day activities, which require no action.
- Debug - Detailed activities used for troubleshooting.

You can select the event level, which includes all levels above, and the time frame to display. For example, if you select the warning level, all warnings, errors, and fatal events will be displayed for the specified time period.

To specify a custom range, select *Custom* under *Time Range* and then specify a date range.

Time	Cluster	Level	Message	Resource Type	Source	Namespace	Event Ty...
2021-03-12 11:54:59		Info	Updated spot price book from v4.5.20210312T10...	Pod	Federator.ai	federatorai	
2021-03-12 05:57:26		Info	Updated spot price book from v4.5.20210312T04...	Pod	Federator.ai	federatorai	
2021-03-11 23:59:47		Info	Updated spot price book from v4.5.20210311T22...	Pod	Federator.ai	federatorai	
2021-03-11 17:56:24		Info	Updated spot price book from v4.5.20210311T16...	Pod	Federator.ai	federatorai	
2021-03-11 11:58:45		Info	Updated spot price book from v4.5.20210311T10...	Pod	Federator.ai	federatorai	
2021-03-11 05:56:25		Info	Updated spot price book from v4.5.20210311T04...	Pod	Federator.ai	federatorai	
2021-03-11 00:07:20		Info	Updated spot price book from v4.5.20210310T22...	Pod	Federator.ai	federatorai	
2021-03-10 18:00:09		Info	Updated spot price book from v4.5.20210310T16...	Pod	Federator.ai	federatorai	
2021-03-10 11:57:59		Info	Updated spot price book from v4.5.20210310T10...	Pod	Federator.ai	federatorai	
2021-03-10 10:23:00		Error	Cluster "jason-17-100", data source: "datadog", co...	Pod	Federator.ai	federatorai	
2021-03-10 06:00:58		Info	Updated spot price book from v4.5.20210310T04...	Pod	Federator.ai	federatorai	
2021-03-10 00:02:54		Info	Updated spot price book from v4.5.20210309T22...	Pod	Federator.ai	federatorai	
2021-03-09 17:55:42		Info	Updated spot price book from v4.5.20210309T16...	Pod	Federator.ai	federatorai	
2021-03-09 11:53:31		Info	Updated spot price book from v4.5.20210309T10...	Pod	Federator.ai	federatorai	
2021-03-09 05:56:24		Info	Updated spot price book from v4.5.20210309T04...	Pod	Federator.ai	federatorai	

At the bottom of the page, the total number messages being displayed is shown, along with navigation to other pages. You can also determine how many events to show per page.

Related topics:

[Terminology](#)

[Search/Sort Information in Tables](#)