

# Federator.ai Release v5.0 Installation Guide

# Contents

Overview	3
Federator.ai	3
Supported Metrics Data Sources	3
Prometheus	3
Datadog	4
Sysdig	5
VMWare vCenter	5
AWS CloudWatch	6
Requirements and Recommended Resource Configuration	7
Supported Platform	7
Data Source	7
Federator.ai Resource Requirements	7
Federator.ai Version	7
Datadog Agent Version(reference)	7
Prometheus Version(reference)	8
Sysdig Agent Version(reference)	8
Federator.ai Installation and Configuration	9
Summary of Installation Steps	9
Pre-installation Check List	9
Before You Start	13
Datadog	13
Sysdig	14
VMware vCenter	15
AWS CloudWatch	15
Brand New Installation	17
Federator.ai internal/external communication ports	19
Appendix	22
Datadog Dashboards Overview	22
ProphetStor Federator.ai Cluster Overview	22
ProphetStor Federator.ai Application Overview	22
ProphetStor Federator.ai Kafka Overview	23
ProphetStor Federator.ai Cost Analysis Overview	23
Sysdig Dashboard Overview	24
Federator.ai Cluster Overview	24
Federator.ai Application Overview	24

Federator.ai Application Overview	25
Federator.ai installation/uninstallation using Helm Chart	25
Federator.ai installation/uninstallation using Ansible	27
Prerequisite	27
Preparation (Ansible Control Node):	28
Installing Federator.ai	28
Configure nativeHPA using external metrics for Datadog	30
Deploy KEDA for Remote and Sysdig monitored Cluster Pod Autoscaling	31
Example: Configure Generic Application HPA using KEDA ScaledObject	31

# **Overview**

# Federator.ai

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 or AWS EC2 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 or AWS EC2 clusters
- Resource recommendations based on workload prediction, application, Kubernetes, and other related metrics
- Correlation and causality analysis of microservices/controllers of Kubernetes applications
- Automatic scaling of Kubernetes application containers, Kafka consumer groups, and Ingress upstream
- 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
- Correlation and causality analysis of microservices/controllers of Kubernetes applications
- Statistical analysis and predictions based on the correlation between resource usage and application workload

# **Supported Metrics Data Sources**

There are five different metrics data sources supported in released v5.0.0-ga: Prometheus, Datadog, Sysdig, VMWare vCenter, and AWS CloudWatch.

## Prometheus

Prometheus is a free and open-source event monitoring tool for containers or microservices. It uses scraping to collect numerical data based on time series. Metrics are collected in regular timestamps and stored locally. Federator.ai supports using Prometheus gathering Kubernetes cluster metrics and leveraging collected data for workload predictions, recommendations for resource planning, autoscaling containers/pods, and cost analysis for clusters deployed in a multicloud environment.

The following diagram shows how the metrics are collected from Prometheus by Federator.ai in a Kubernetes environment.



## Datadog

Fedederator.ai has integrated with Datadog and utilizes the metrics collected by Datadog Agent for workload predictions. The following diagram shows how application metrics are used by Federator.ai to predict workload and to scale applications for better performance automatically. Specifically,

- Datadog Agent sends cluster/applications metrics to Datadog Services
- Federator.ai's Data-adapter queries cluster/applications metrics from Datadog Services and forwards to Federator.ai AI engine
- Data-adapter posts the prediction/recommendation/plan created by Federator.ai to Datadog Services
- Datadog Cluster Agent gets prediction/recommendation/plan from Datadog Services
- Datadog Dashboard displays cluster/applications metrics and prediction/recommendation/plan by Federator.ai



## Sysdig

Fedederator.ai has integrated with Sysdig and utilizes the metrics collected by Sysdig Agent for workload predictions. The following diagram shows how application metrics are used by Federator.ai to predict workload and to autoscale applications for better performance and saving resources.



## VMWare vCenter

VMware vCenter Server provides integrated management of all hosts and virtual machines in the data center from a single console, allowing IT administrators to improve control, simplify daily work, and reduce the complexity and cost of managing the IT environment. Federator.ai data adapter connects to VMware vCenter servers via VMware SDK to retrieve VMs workload metrics data for predictions, recommendations, and cost analysis for VM clusters.



## **AWS CloudWatch**

AWS CloudWatch is a monitoring service for AWS cloud resources and the applications running on the AWS cloud. It provides visibility into resource utilization, operational performance, and overall demand patterns.

The metrics collected by CloudWatch by default do not include memory usage of EC2 instances. Therefore, CloudWatch agent is required for Federator.ai to collect memory usage metrics. Federator.ai supports two types of AWS VM clusters:

- Auto Scaling Group
- Individual VM



# **Requirements and Recommended Resource Configuration**

# **Supported Platform**

- OpenShift : 4.6~4.9
- Kubernetes : 1.11 ~ 1.22
- Rancher v2.4.8,v2.5.8,v2.5.9
- EKS/AKS/GKE

# Data Source

- Datadog
- Sysdig
- Prometheus
- VMWare vCenter 5.5/6.0/6.5/6.7/7.0
- AWS CloudWatch

# Federator.ai Resource Requirements

- Total Resource Requirments
  - Request: 5.1 CPU cores (Limit: 22 cores)
  - Request: 5.1 GB Memory (Limit: 42GB)
  - StorageClass: 176GB (require ReadWriteOnce access mode)
- Resource requirements for AI Engine
  - There must be at least one worker node with at least 2 CPU(Limit: 8 cores) cores and 1
     GB of memory available
  - The 2 CPU cores and 1 GB memory are included in the total 5.1 CPU cores and 5.1 GB memory requirements
- Persistent volumes
  - The StorageClass that provides the persistent volumes must support RWO (ReadWriteOnce) access mode.
  - It is recommended to use persistent volumes instead of ephemeral storage to store the data in the production environment. Any data on ephemeral storage will be lost after Federator.ai pods are restarted.

# Federator.ai Version

- Version: Release 5.0-ga
- Tag : v5.0.0-ga

# **Datadog Agent Version(reference)**

- Datadog Agent helm chart version: v2.4.24, v2.13.0
- Datadog Agent version: v7.21.1, v7.27.0
- Datadog Cluster Agent version: v1.8.0, v1.12.0

• Datadog Watermark Pod Autoscaler version: v0.1.0

# Prometheus Version(reference)

- OpenShift
  - Default installed Prometheus
- SUSE Rancher
  - Cattle-Prometheus
- Kubernetes
  - prometheus-operator-8.5.11
  - kube-prometheus-release-0.6
  - kube-prometheus-stack-12.5.0/15.4.6/17.0.3

# Sysdig Agent Version(reference)

• Sysdig agent: 11.2.0/11.3.0

# Federator.ai Installation and Configuration

# **Summary of Installation Steps**

- Step 0: Review pre-installation checklist items, make sure the environment and required information are ready.
- Step 1: Preparation
  - For Datadog, obtain API Key, Application Key of Datadog Cloud Service account. Instructions are provided below.
  - For Prometheus, obtain Prometheus service URL ( ex : http://<prometheus\_svc\_name>.<namesapce>:9090)
  - For Sysdig, obtain Sysdig API URL and Token.
  - For VMware vCenter, obtain adminstrator login credential and vCenter IP or FQDN.
  - For AWS CloudWatch, obtain Access Key ID and Secret Access Key of AWS account.

Step 2:

- For Datadog, install and configure Datadog Agent/Cluster Agent if they have not been installed. Please follow the Datadog documentation on how to install Datadog Agent and Cluster Agent.
- For Sysdig, install and configure Sysdig Agent. Please follow Sysdig documentation on how to install Sysdig Agent.
- Step 3: Install Federator.ai.
- Step 4: Configure Federator.ai Data Adapter for the external metrics data source via Federator.ai Initial Setup Wizard.
- Step 5: Check integrated Federator.ai dashboard on Datadog/Sysdig Cloud.

# **Pre-installation Check List**

## **Kubernetes Access Management Requirement**

ClusterRole/ClusterRoleBinding, Role/RoleBinding

Federator.ai Feature	Component	Scope	Operation
Installation, Lifecycle Management	federatorai-operator	Cluster	admin
Data Collection/Controller	alameda-datahub	Cluster	edit
	alameda-operator	Cluster	edit
	federatorai-data-adapter	Cluster	view
Dashboard	federatorai-dashboard- frontend	Local	view
	federatorai-dashboard- backend	Cluster	view

Prediction	alameda-ai	Local	view
	ai-dispatcher	Local	view
Recommendation	alameda-recommender	Local	view
Autoscaling Execution	alameda-executor	Cluster	edit
Workload Anomaly	alameda-analyzer	Cluster	view
Notification	alameda-rabbitmq	Local	view
	alameda-notifier	Cluster	edit
Cost Analysis	fedemeter-api	Local	view
	federatorai-agent	Local	view
Database	alameda-influxdb	Local	view
	federmeter-influxdb	Local	view
REST API	federatorai-rest	Cluster	edit

Installation Script Run Permission Requirement

- Installation scripts do NOT require root permissions to run in the host where "kubectl" command is set up
  - The default script download path is "/opt", if the non-root user does not have the permission to write "/opt", input a different download path during installation.

```
$ curl -s https://raw.githubusercontent.com/containers-
ai/prophetstor/master/deploy/federatorai-launcher.sh | bash
Please enter Federator.ai version tag [default: latest]:
Federator.ai version = v5.0.0-ga
Please enter the path of Federator.ai directory [default: /opt]:
```

Installataion scripts support Linux OS only
 macOS and Windows are not supported

Kubernetes:

#	Checklist Item	Requirement	Details
1	What is the Kubernetes version?	Server Version: 1.11~1.22	Use the command below to get the Kubernetes version: <b>\$ kubectl version</b>
			<pre> Server Version: version.Info{Major:"1", Minor:"17", GitVersion:"v1.17.2", GitCommit:"59603c6e503c87169aea6106f57b9f242f6 4df89", GitTreeState:"clean", BuildDate:"2020- 01-18T23:22:30Z", GoVersion:"go1.13.5", Compiler:"gc", Platform:"linux/amd64"}</pre>

2	Does installation on this Kubernetes cluster require a private image repository?	If a private image repository is required, the following information is needed during installation - Private image repository URL - Credential of the private image repository	Input the URL and credential when the Federator.ai installation script asks for the information.
3	StorageClass and Persistent Volumes requirement	StorageClass supports ReadWriteOnce access mode. Available storage size is larger than 176GB.	The minimum storage size for Federator.ai Release v5.0 is 176GB, including database, data, and logs.
4	Kubernetes cluster CPU/memory requirement	<ul> <li>Minimum CPU/mem/storage:</li> <li>CPU: 5,100 (mcores)</li> <li>Memory: 5.1 (GB)</li> <li>Storage Class Capacity: 168GB</li> <li>At least one worker node with</li> <li>CPU: 2 Cores</li> <li>Memory: 1GB</li> </ul>	To be able to run the AI Engine pod, there must be at least one worker node that has more than 2 CPU cores and 1 GB of memory available. 2 CPU Cores and 1GB for AI Engine are included in the total 5.1 CPU Cores and 5.1 GB memory requirements.
5	Is this Kubernetes cluster allowed for NodePort configuration?	Federator.ai creates two NodePorts for GUI and REST API by default - REST API - https:// <server>:31011 - GUI - https://<server>:31012</server></server>	If NodePort is not allowed, answer 'N' when the installation script prompts for creating NodePorts. Users need to expose Federator.ai GUI and REST API service manually.
6	Will there be a resource quota imposed for the namespace where Federator.ai is installed?	CPU/mem request quota should be more than the minimum resource requirement - CPU: 5.1 Cores - Memory: 5.1 GB	The CPU/memory required for Federator.ai depends on the number of clusters and applications being monitored/managed. Suggestion for initial namespace quota is - CPU 8 cores - Memory 12G The quota could be adjusted if the number of managed clusters/applications increases. Use the command to get namespace resource quota \$ kubectl get resourcequotaall-namespaces
7	Does this deployment require resource request/limit specified?	By default, Federator.ai deployments do not specify resource requests/limits. It can be done by setting up an environment variable before installation starts.	To turn on resource request/limit settings for all Federator.ai deployments, manually export environment variable before running 'federatorai- launcher.sh' \$ export ENABLE_RESOURCE_REQUIREMENT=y \$ ./federatorai-launcher.sh

## Prometheus:

#	Checklist Item	Requirement	Details
1	What is the Prometheus version? (for Kubernetes)	<ul> <li>Recommended version</li> <li>Prometheus operator helm chart version: 8.5.11</li> <li>Prometheus operator version: 0.34.0</li> <li>Prometheus server version: 2.13.1</li> </ul>	Use the command below to get Prometheus version: \$ helm Is -A  grep -i prometheus prometheus-adapter monitoring 1 2020-03-13 15:35:05.28963154 +0800 CST deployed prometheus-adapter-2.1.3 v0.6.0 prometheus-operator monitoring 1 2020-03-13 14:34:16.132479221 +0800 CST deployed prometheus-operator-8.12.1 0.37.0

	<pre>\$ kubectl get deployment -A -o custom- columns=IMAGE:.spec.template.spec.containers[0].im</pre>
	age  grep -i prometheus
	directxman12/k8s-prometheus-adapter-amd64: v0.6.0 quay.io/coreos/prometheus-operator: v0.37.0

# Datadog Agent:

#	Checklist Item	Requirement	Details
1	Is Datadog Agent installed?	Datadog Agent is mandatory	Kubernetes resources and workload metrics are collected by Datadog Agent.
2	Is Datadog Cluster Agent installed?	Cluster Agent is mandatory for the HPA autoscaling feature	Cluster Agent provides metrics to HPA Autoscaler for autoscaling.
3	Is Datadog WPA controller installed? (Option)	Datadog WPA is required if auto- scaling is done by WPA	Datadog WPA is the HPA Autoscaler developed by Datadog. Users can use Datadog WPA or Kubernetes native HPA to do autoscaling.
4	Datadog Kafka Consumer integration is enabled?(Option)	Datadog Kafka Consumer integration is mandatory if user wants to use Kafka consumer optimization feature	Use the command to confirm Kafka integration is enabled \$ <b>kubectl exec</b> < <i>datadog-agent-pod&gt;</i> -n < <i>datadog- agent-namespace&gt;</i> agent integration show datadog- kafka-consumer Refer to <u>https://www.datadoghq.com/blog/monitor- kafka-with-datadog/</u> for Kafka Consumer integration installation
5	Datadog account API key	An API key is mandatory for connecting Datadog Service	Follow the steps described in the "Before You Start" session to obtain the API key.
6	Datadog account Application key	An application key is mandatory for connecting Datadog Service	Follow the steps described in the "Before You Start" session to obtain the Application key.
7	Is one of cluster name is configurated for the Datadog agent/cluster agent? 1.>DD_TAGS with value ="kube_cluster: <cluster_name &gt;" in values.yaml or 2.&gt;"cluster_name" in values.yaml, or 3.&gt;"DD_CLUSTER_NAME" in Datadog cluster agent deployment</cluster_name 	"kube_cluster","cluster_name","kub e_cluster_name(DD_CLSUTER_NAM E)" one of them is required for Federator.ai to identify Kubernetes cluster.	Case 1.>New Datadog Agent installation: Install Datadog agent and cluster agent by "helm install -f values.yaml", in values.yaml.  clusterName: <cluster-name>  clusterAgent: enabled: false true Case 2.&gt; In Datadog Agent installed environment, with no Cluster Agent and no cluster_name setting Update Datadog Agent to enable Cluster agent by "helm upgrade -f values.yaml", in values.yaml - assign a cluster name  datadog: clusterName: <cluster-name> - enable cluster agent  clusterAgent: enabled: false true  - \$helm upgrade</cluster-name></cluster-name>

- Check "DD_Cluster_Name"
\$kubectl get daemonset
<datadog_agent_daemonset_name> -n</datadog_agent_daemonset_name>
<datadog_agnet_namespace> -o yaml</datadog_agnet_namespace>
- name: DD_CLUSTER_NAME
value: <cluster-name></cluster-name>
3.>In Datadog Agent and Cluster Agent installed
environment, with no cluster_name setting
Update Datadog Agent by "helm upgrade"
<ul> <li>assign a cluster name</li> </ul>
datadog:
clusterName: <cluster-name></cluster-name>
- \$helm upgrade
<ul> <li>Check "DD_Cluster_Name"</li> </ul>
\$kubectl get daemonset
<datadog_agent_daemonset_name> -n</datadog_agent_daemonset_name>
<datadog_agnet_namespace> -o yaml</datadog_agnet_namespace>
As:
- name: DD_CLUSTER_NAME
value: <cluster-name></cluster-name>
4.> In Datadog Agent and Cluster Agent installed
environment, with cluster_name setting
Use the command below to confirm DD_Cluster_Name
<ul> <li>\$kubectl get daemonset</li> </ul>
<datadog_agent_daemonset_name> -n</datadog_agent_daemonset_name>
<datadog_agnet_namespace> -o yaml</datadog_agnet_namespace>
As:
<ul> <li>name: DD_CLUSTER_NAME</li> </ul>
value: <cluster-name></cluster-name>

# **Before You Start**

Datadog

- The admin role for installing Fedeator.ai is "Cluster Admin."
- Datadog agent must be ready if Federator.ai runs in the same Kubernetes cluster monitored.
- Obtain Datadog account API Key, Application Key.
  - 1. A Datadog account is required for connecting and using Datadog Cloud Service. If you don't have an account, visit Datadog website and sign up for a free trial account. https://www.datadoghq.com/
  - 2. Log in Datadog Cloud Service with your account and get an API Key and Application Key for using Datadog API

# https://docs.datadoghq.com/account management/api-app-keys/



# Copy the API Key and Application Key for Federator.ai metrics data source configuration

✓ API Keys						
Your APT keys are unique to your organization. An APT key is required by the Detadeg Agent to submit metrics and events to Datadeg.						
Name Key	Created by	Created at (UTC)				
8a94db9a5er 11 89e8a6ặc	kyv745.chen@gmail.com	2020-05-21 08:31:24	Revoke			
New API key						
API key name Create API Ke	ry .					
Application Keys						
Application keys, in conjunction with your org's API key, give you full access to Datadog's progr	ammatic API. Application keys are associated with the user account the	hat created them and can be named. The applicatio	n key is used to log all requests made to the API.			
Name 2 Key Created by						
pod		kyv745.chen@gmail.com	Revoke			
New application key						

Sysdig

- Different Sysdig API URL is needed for different regions :
  - For US East, Sysdig API URL is <u>https://app.sysdigcloud.com</u>
  - For US West, Sysdig API URL is <u>https://us2.app.sysdig.com</u>
  - For European Union, Sysdig API URL is <u>https://eu1.app.sysdig.com</u>
- Copy Sysdig Monitor API Token for Federator.ai metrics data source configuration

$\leftarrow \   \rightarrow $	C	1 https://app.sysdig	gcloud.com/#/settings/user		口 Q ☆	7 6 6
<b>\$</b>	Set	tings				
MONITOR	٢	User Profile	User profile			
Overview	•	Users	Account Name	stella.chen@prophetstor.com		
<u>th</u>	*	Teams	Role	Admin		
Dashboards	۰	Notification Channels	Current Team	Monitor Operations		
Explore		AWS				
ل Alerts		Subscription	Admin Settings			
	Ţ	Agent Installation	Hide Agent Install	Enabling this feature hides the Access Key and Agent Installation page for non-admin users.		
Events	07	Authentication				
			Sysdig Monitor API			
رچی Get Started			Sysdig Monitor is based on a REST API that ci your API token for another team, switch to tha Click Reset Token to generate a new token if n NOTE: When reset, the previous token issued of See also Developer Documentation.	an be accessed from custom scripts or programs. Access requires the API security token below. Tol t team. exceed. will immediately become invalid and you will need to make appropriate changes to your programs o	kens are team-specifi r scripts.	ic; to retrieve
SC			Sysdig Monitor API Token			COPY
0					R	ESET TOKEN

### VMware vCenter

• You can define a VM cluster from any VMs under the same cluster path. See below for an example of cluster path on vCenter.



AWS CloudWatch

• Obtain CloudWatch Account Key ID and Secret Access Key.

- 1. Use your AWS account ID or account alias, your IAM username, and your password to sign in to the <u>IAM console</u>.
- Go to "Access management > Users > Security credentials" to get Access Key ID and Secret Access Key

aws Services V	Q Search for services, features, marketp	place products, and docs [Alt+S]	D & Clinnin Carolonyand) ▼ Global ▼ Support ▼							
Identity and Access Management (IAM)	Users > (Cherlin)									
Dashboard - Access management	User ARN am:aws:lam	user/diamite C	Delete user 🛛 😧							
User groups Users Roles Roles	Creation time 2018-01-09 10:35 UT Permissions Groups (2) Tags Securit	TC+0800 ty credentials Access Advisor								
Identity providers Account settings	Sign-in credentials	Console sign-in link: https://console.com/openationals.ama	zon.com/console (2)							
Access reports     Access analyzer     Archive rules	Console password Enabled (last signed in Today)   Manage Assigned MFA device Not assigned   Manage Signing certificates None 4									
Analyzers Settings Credential report	Access keys Use access keys to make programmatic calls to AWS fr inactive) at a time.	rom the AWS CLI, Tools for PowerShell, AWS SDKs, or dire	ct AWS API calls. You can have a maximum of two access keys (active or							
Organization activity Service control policies (SCPs)	For your protection, you should never share your secret If you lose or forget your secret key, you cannot r Create access key	ret keys with anyone. As a best practice, we recommend fre retrieve it. Instead, create a new access key and mak	quent key rotation. e the old key inactive. Learn more							
AWS account ID:	Access key ID Created AKIAICYTE RA 2018-02-13 11:37	Last used 7 UTC+0800 2021-05-18 22:12 UTC+0800 with storagega	Iteway in us-east-2 Active   Make inactive x							
	AKIAJYQ 16YFQ 2018-02-27 17:38	8 UTC+0800 2020-01-28 14:35 UTC+0800 with sts in us-e	ast-1 Active   Make inactive 🗙							

https://docs.aws.amazon.com/IAM/latest/UserGuide/id\_credentials\_access-keys.html

## **Brand New Installation**

### 1. Connect to Kubernetes cluster

### 2. Install the Federator.ai for Kubernetes by using the following command

```
$ curl https://raw.githubusercontent.com/containers-
ai/prophetstor/master/deploy/federatorai-launcher.sh | bash
```

```
~$ curl https://raw.githubusercontent.com/containers-
ai/prophetstor/master/deploy/federatorai-launcher.sh | bash
 % Total % Received % Xferd Average Speed Time
                                                               Time Current
                                                      Time
                               Dload Upload Total Spent
                                                               Left Speed
100 17101 100 17101 0 0 30118 0 --:--:- --:-- 30107
Please enter Federator.ai version tag [default: latest]:
Federator.ai version = v5.0.0-ga
Please enter the path of Federator.ai directory [default: /opt]:
Downloading v5.0.0-ga tgz file ...
Done
Do you want to use a private repository URL? [default: n]:
Do you want to launch the Federator.ai installation script? [default: y]:
Executing install.sh ...
Checking environment version...
...Passed
Enter the namespace you want to install Federator.ai [default: federatorai]:
tag_number = v5.0.0-ga
install_namespace = federatorai
Is the above information correct? [default: y]:
Downloading v5.0.0-ga tgz file ...
Done
Applying Federator.ai operator yaml files...
Applying 00-namespace.yaml...
namespace/federatorai created
Applying 01-serviceaccount.yaml...
serviceaccount/federatorai-operator created
Applying 02-alamedaservice.crd.yaml...
customresourcedefinition.apiextensions.k8s.io/alamedaservices.federatorai.containers.ai
created
Applying 03-federatorai-operator.deployment.yaml...
deployment.apps/federatorai-operator created
Applying 04-clusterrole.yaml...
clusterrole.rbac.authorization.k8s.io/federatorai-operator created
clusterrole.rbac.authorization.k8s.io/alameda-gc created
Applying 05-clusterrolebinding.yaml...
clusterrolebinding.rbac.authorization.k8s.io/federatorai-operator created
Applying 06-role.yaml...
role.rbac.authorization.k8s.io/federatorai-operator created
Applying 07-rolebinding.yaml...
rolebinding.rbac.authorization.k8s.io/federatorai-operator created
Applying 08-service.yaml...
```

service/federatorai-operator-service created Applying 09-secret.yaml... secret/federatorai-operator-service-cert created Applying 10-mutatingwebhook.yaml... mutatingwebhookconfiguration.admissionregistration.k8s.io/federatorai-operatorservicesmutation created Applying 11-validatingwebhook.yaml... validatingwebhookconfiguration.admissionregistration.k8s.io/federatorai-operatorservicesvalidation created Checking pods... Waiting for pod federatorai-operator-669566b7c-rmphp in namespace federatorai to be ready. phase: [Running] Waiting for pods in namespace federatorai to be ready... All pods under namespace(federatorai) are ready. Install Federator.ai operator v5.0.0-ga successfully Downloading Federator.ai alamedaservice sample file ... Done Downloading Federator.ai alamedascaler sample files ... Done \_\_\_\_\_ Which storage type you would like to use? ephemeral or persistent? [default: persistent]: Specify log storage size [e.g., 2 for 2GB, default: 2]: Specify AI engine storage size [e.g., 10 for 10GB, default: 10]: Specify InfluxDB storage size [e.g., 100 for 100GB, default: 100]: Specify storage class name: managed-nfs-storage Do you want to expose dashboard and REST API services for external access? [default: y]: \_ \_ \_ \_ \_ \_ \_ \_ \_ \_ \_ \_ install\_namespace = federatorai storage type = persistent log storage size = 2 GB AI engine storage size = 10 GB InfluxDB storage size = 100 GB storage class name = managed-nfs-storage expose service = y -----Is the above information correct [default: y]: Processing... Waiting for datahub(v5.0.0-ga) pod to appear ... datahub pod is present. Checking pods... Waiting for pod alameda-ai-66f5c7b6b4-rx87j in namespace federatorai to be ready. phase: [Pending] Waiting for pods in namespace federatorai to be ready... Waiting for pod alameda-ai-66f5c7b6b4-rx87j in namespace federatorai to be ready. phase: [Pending] Waiting for pods in namespace federatorai to be ready... Waiting for pod alameda-ai-66f5c7b6b4-rx87j in namespace federatorai to be ready. phase: [Pending] Waiting for pods in namespace federatorai to be ready...

Waiting for pod alameda-ai-66f5c7b6b4-rx87j in namespace federatorai to be ready. phase: [Running] Waiting for pods in namespace federatorai to be ready... Waiting for pod alameda-operator-7ff69f4bb5-v22ws in namespace federatorai to be ready. phase: [Running] Waiting for pods in namespace federatorai to be ready... All pods under namespace(federatorai) are ready. The default alamedaorganization under namespace federatorai is ready. \_\_\_\_\_ You can now access GUI through https://<YOUR IP>:31012 The default login credential is admin/admin Also, you can start to apply alamedascaler CR for the target you would like to monitor. Review the administration guide for further details. -----\_\_\_\_\_ You can now access Federatorai REST API through https://<YOUR IP>:31011 The default login credential is admin/admin The REST API online document can be found in https://<YOUR IP>:31011/apis/v1/swagger/index.html -----Install Federator.ai v5.0.0-ga successfully Downloaded YAML files are located under /opt/federatorai/installation Downloaded files are located under /opt/federatorai/repo/v5.0.0-ga

### 3. Verify Federator.ai pods are running properly

~# kubectl get pod -n federatorai				
NAME	READY	STATUS	RESTARTS	AGE
alameda-ai-66f784f79d-xphsb	1/1	Running	0	7m37s
alameda-ai-dispatcher-694474f89-njmjb	1/1	Running	0	7m37s
alameda-datahub-679875645c-cc6kt	1/1	Running	0	8m12s
alameda-executor-bf699b5fb-26r4n	1/1	Running	0	7m37s
alameda-influxdb-0	1/1	Running	0	8m12s
alameda-notifier-b87744cdc-wwrgt	1/1	Running	0	7m37s
alameda-rabbitmq-7f5869cdc8-5wt57	1/1	Running	0	8m12s
fedemeter-api-64665c6ff4-gbfsn	1/1	Running	0	8m12s
fedemeter-influxdb-0	1/1	Running	0	8m12s
federatorai-agent-774cdfb9dd-bblkh	1/1	Running	0	7m37s
federatorai-dashboard-backend-67b749597c-sqhlq	1/1	Running	0	7m37s
federatorai-dashboard-frontend-856458bf5c-pb9kb	1/1	Running	0	7m37s
federatorai-data-adapter-9f977cb5b-jbtqz	1/1	Running	0	7m37s
federatorai-operator-858cf75ffc-2tn6v	1/1	Running	0	11m
federatorai-postgresql-0	1/1	Running	0	8m12s
federatorai-recommender-dispatcher-5c4d6c965f-6mfwf	1/1	Running	0	7m37s
federatorai-recommender-worker-585ccc7f55-q7lf7	1/1	Running	0	7m25s
federatorai-rest-855bfb6956-Lcprj	1/1	Running	0	7m37s

# Federator.ai internal/external communication ports

If NodePort is not permitted due to security reason, disable NodePorts from Federator.ai
 Operator

edit AlamesaService CR and remove "serviceExposures session "

```
~# kubectl edit alamedaservice my-alamedaservice -n federatorai
. . .
serviceExposures:
  - name: federatorai-dashboard-frontend
   nodePort:
     ports:
      - nodePort: 31012
        port: 9001
    type: NodePort
  - name: federatorai-rest
   nodePort:
     ports:
      - nodePort: 31011
       port: 5056
   type: NodePort
  storages:
. . . .
```

Or during installing process: "Expose dashboard and REST API services: (default:y)" step to answer 'N' to disable NodePort

```
~# curl https://raw.githubusercontent.com/containers-
ai/prophetstor/master/deploy/federatorai-launcher.sh | bash
...
...
Which storage type you would like to use? ephemeral or persistent?
[default: persistent]:
Specify log storage size [e.g., 2 for 2GB, default: 2]:
Specify AI engine storage size [e.g., 10 for 10GB, default: 10]:
Specify InfluxDB storage size [e.g., 100 for 100GB, default: 100]:
Specify storage class name: managed-nfs-storage
Do you want to expose dashboard and REST API services for external access? [default: y]:
...
```

\*NGINX Ingress or LB tools may be used for external access to Federator.ai GUI if disabled NodePort.

Connecting to Federator.ai Web portal

```
Kubernetes/Rancher Cluster
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
```

<pre># kubect</pre>	# kubectl get nodes -o wide										
NAME	STATUS	ROLES	AGE	VERSION	INTERNAL-IP	EXTERNAL-IP	OS-IMAGE				
KERNEL-V	ERSION	CC	NTAINE	R-RUNTIME							
h7-130	Ready	master	35d	v1.18.5	172.31.7.130	<none></none>	CentOS Linux 7 (Core)				
3.10.0-9	57.el7.x8	6_64 do	cker:/	/19.3.13							
h7-131	Ready	<none></none>	35d	v1.18.5	172.31.7.131	<none></none>	CentOS Linux 7 (Core)				
3.10.0-9	57.el7.x8	6_64 do	cker:/	/19.3.13							
h7-132	Ready	<none></none>	35d	v1.18.5	172.31.7.132	<none></none>	CentOS Linux 7 (Core)				
3.10.0-9	57.el7.x8	6_64 do	cker:/	/19.3.13							
h7-133	Ready	<none></none>	35d	v1.18.5	172.31.7.133	<none></none>	CentOS Linux 7 (Core)				
3.10.0-957.el7.x86_64 docker://19.3.13											
			-								

(example) 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 federa	torai grep fede	eratorai-dashboard-frontend		
The output will le	ook somet	hing like this:			
NAME		HOST/PORT		PATH	SERVICES
	PORT	TERMINATION	WILDCARD		
federatorai-dashboar	d-frontend	federatorai-dashboo	ard-frontend-federatorai.apps.ocp4.172-31-2-86.nip.io		federatorai-
dashboard-frontend	frontend-htt	p edge/Redirect	None		
federatorai-rest		federatorai-rest-fe	ederatorai.apps.ocp4.172-31-2-86.nip.io		federatorai-
rest	restapi-http	edge/Redirect	None		

# (example) The URL will be https://federatorai-dashboard-frontend-federatorai.apps.ocp4.172-31-2-86.nip.io

## Federator.ai Web Portal





# Appendix

# **Datadog Dashboards Overview**

The following Custom Datadog Dashboards are available after Federator.ai is installed.

# ProphetStor Federator.ai Cluster Overview

978	* ProphetStor Federator.ai Cluster Overview 🗸 Edit Widgets +									1h Past 1 Hour	Q		
	Save or select views	• \$kube_cluste	er jean3-61 * \$	host * • \$prediction_win	dow 24_hours *							∞ <b>↓</b> ♦	
DATADOG													
New Stuffl		Cluster Resource Usage Predictions and Recommendations											
H Watchdog	- F		. 1	Cluster Resource Usa	esource Usage Predictions and Recommendations								
Events				KUBE_CLUSTER PREDICT	ION AVG CPU (MC	:0 MIN CPU (MC	MAX CPU (MC R	EC CPU (MCO A	VG MEM (BYT MIN MI	M (BY MAX MEM (BY	REC MEM (BYT		
🖮 Dashboards 🔸	Pro	phetS	tor	jean3-61 24_hour	6.	53K 4.89K	7.83K	8.12K	40.34G	14.54G 66.95	G 73.19G		
😵 Infrastructure 🖡													
Monitors													
(7) Metrics	Cluster Node I	Resource Usa	age Predictions	and Recommendatio	ns								
📥 Integrations 🔸	KUBE_CLUSTER	HOST	PREDICTION_WI	N AVG CPU (MCORE)	MIN CPU (MCORE)	MAX CPU (MCORE)	REC CPU (MCORE)	AVG MEM (BY	TE) MIN MEM (BYTE)	MAX MEM (BYTE)	REC MEM (BYTE)		
E APM	jean3-61	jek8s-361	24_hours	4.1K	3.47K	4.77K	4.9K	14.3	2G C	31.07G	33.57G		
<ul> <li>All and a set of a set of</li></ul>	jean3-61	jek8s-364	24_nours	1.09K 851.84	205.74	1.13K	984	7.	190 7.260 166 3.350	14.04G 8.34G	9.176		
INOCEDOOKS .	jean3-61	jek8s-362	24_hours	496.32	326.25	673.79	1.01K	8.4	15G 3.93G	13.49G	14.81G		
5D Logs													
C Security													
🥮 UX Monitoring 🕴													
					Nodes	Resource Utilizat	tion History						
	Node Current,	Predicted Cl	PU Usage (Daily	/) 2d	Node Current/Pr	edicted CPU Usag	e (Weekly)	Tw No	ode Current/Predict	ed CPU Usage (Mon	thly) Imo		
	8K			hum	1			1	kubernetes.cpu.usage.tota	Lmillicore{	1.1		
	4× WAM	Man Man	MAMARA	Mary man	10K	,		1	kube_cluster:jean3-61, hos	(jek8s-362) 1.3K	I.I. ALL		
O Hala	2K 40444	and more	Antermant	Company of the same	SK What Harry	santing (	mandminh	in s	ĸ	my harded	Monthat 1		
G Heip	0K We	d 19	12:00 Th	20 12:00	OK Sat 15	Mon 17	Wed 19	(	к	Aug 2 09/06 18:00 9	Aug 16		
It Team	Value Min	Avg Max	Metric	Tags J	Value Min A	wg Max Metric	Tags ↓		Value Min Avg	lax Metric	Tags 4		
ProphetStor Data	5.49K 1.22 3.72K 2.98	K 4.35K 7.49K K 4.21K 5.16K	federatorai.predicti.	host:jek8s-361,kube	2.71K 2.86K	4,48K 11.93K kubernet 3,94K 5.24K federator	es.cpu.usa host:jek8s-3 ai.predicti host:jek8s-3	161,kube	55.76 41.25 862.98	5.23K federatoral.predict	host-jek8s-361,kub		

# ProphetStor Federator.ai Application Overview

9.		* ProphetStor Federator.ai Application Overview 🗸 🛛 🖬 Past 1 Hour									<b>- - - - - - - - - -</b>				
		Save or select views	\$kube_cluster jean	3-61 • \$kube_names	pace * • \$kube_d	eployment * 👻 !	kube_stateful_set *	• /							
DATA	DOG														
						А	pplication Worl	doad Pr	ediction/Re	source Recom	mendation				
	Stuff!														
M Watch	hdog			Workloa	d Prediction for I	Next 24 Hours									
	ts			KUBE_NAM	ME KUBE_DEPLO	KUBE_STATEF A	VG CPU (MC MIN	CPU (M	MAX CPU (M	REC CPU (MC	AVG MEM (B	MIN MEM (B	MAX MEM (B	REC MEM (BY	
📥 Dashb	boards >			nginx-pre	oader nginx-prepared	N/A	335.86	240.93	437.68	394	3.21M	584.89K	6.18M	11.22M	
	structure >	Prop	hetStor	myproject	consumer1-top	i N/A	87.12	64.47	199.68	133	2.26G	2.12G	2.35G	2.38G	
	tors >			myproject	consumer2-top	i N/A	71.87	58.88	111.28	88	1.77G	1.17G	2.77G	2.98G	
(4) Metric	cs →														
📥 Integr	ratione >														
	, adding			_											
		Workload Predic	tion for Next 7 l	Days	AVG CRIL (MCORES)	MIN CRIL (MCORES			DII (MCORES)	AVG MEM (BYTES)	MIN MEM (DV	TES) MAY M	EM (BYTES)	DEC MEM (DVTES)	
🗐 Noteb	books <sup>&gt;</sup>	nginx-preloader-san	nginx-prepared	N/A	244.9	134.7	398	.5	688	6.42M	6.	42M	6.42M	5.87M	
	•	myproject	consumer2-topic00	0 N/A	68.94	57.3	8 87	.1	56	2.35G	2	01G	2.63G	1.82G	
	rity ≻	myproject	consumer1-topic00	0 N/A	59.57	59.5	7 59.5	7	51	1.7G	1	43G	1.97G	1.84G	
	onitoring →														
		Workload Predi	ction for Next 30	Days											
		KUBE_NAMESPACE	KUBE_DEPLOYMENT	KUBE_STATEFUL_S	AVG CPU (MCORES)	MIN CPU (MCORES	MAX CPU (MCORE	S) REC C	PU (MCORES)	AVG MEM (BYTES)	MIN MEM (BY	TES) MAX M	EM (BYTES)	REC MEM (BYTES)	
		nginx-preloader-san	nginx-prepared	N/A	429.82	391.5	471.5	9	451	7.96M	;	7.7M	8.26M	9.09M	
		myproject	consumer2-topic00	0 N/A	79.01	75.0	5 87.1	5	84	1.86G	1	57G	2.15G	2.36G	
<b>O</b> Help		myproject	consumer1-topic00	0 N/A	78.89	78.8	9 78.1	9	79	1.44G	858.	94M	2.02G	2.23G	
98 Team															
root@	prophet														
Prophe	etStor Data					Applicatio	n Resource Utili	zation H	listory						

# ProphetStor Federator.ai Kafka Overview

	0.18	★ ProphetStor Federator.ai Kafka Ove	Clone Dash	board	1h Past 1 Hour	Q
		\$kube_cluster jean3-61 v \$kube_namespace myproject v \$	sube_deployment consumer1-to	ppic0001-group-0001 v \$topic topic0001 v \$	\$consumer_group group0001 v	··· 🖓 🗘
D	ATADOG					
<b>1</b>	New Stuff!	E	With integration of Prophe	etStor Federator.ai, users can easily track the	Kafka message	
A	Watchdog		dashboard. Based on the j	prediction or message production rate, Feder	rator.ai automatically scales Kafka Federator ai dashboard where the	
	Events	ProphetStor	recommended consumer overall consumer lags as y	replicas and the current number of consumer vell as the average latency in the gueue befor	r replicas are shown. Additionally, re a message is received by a	
	Dashboards 🛛 🕨		consumer are also shown	on the dashboard for better performance mo	onitoring.	
٠	Infrastructure →	Recommended Replicas vs Current/Desired Replicas	5	Production vs Consumption vs Produ	iction Prediction	
0	Monitors +	8				
<i>(</i> 4)	Metrics >	6				
<b>ń</b> -	Integrations	2		100K		
-	APM >	0	15 14:30	0K 13:45 14:00	14:15 14:30	
	Notebooks 💦	Kafka Consumer Lag		Consumer Queue Latency (msec)		
Þ	Logs >			20K	,	
۲	Security >	4K		10K	16.01K	
÷	UX Monitoring $ ightarrow$	2K		consumer_group:group	p0001, kube_cluster:jean3-61, topic:topic0001	,
		0K	15 14:30	0K 13:45 14:00	14:15 14:33:20	
		Deployment Memory Usage		Deployment CPU Usage		
Ø	Help →	2		100		
21	Team			when when when when when when when when	my man war in	
<b>a</b>	root@prophet ProphetStor Dat,			, , , , , , , , , , , , , , , , , , ,	1 minute the design of the	

# ProphetStor Federator.ai Cost Analysis Overview

	ProphetSto	r Federator.ai C	ost Analysis	Overview ~	dit Widgets +					1h Past 1 Hour		
	ave or select views 💌	\$kube_cluster jean3-61	<ul> <li>\$pricing_option</li> </ul>	-demand-instance * Sco	untry usa 💌	1						
G								Multiclou	d Cost Analysis			
	_		Current Clust	er Cost 4h					Current Cluste	er Configuration		
					HOST	KUBE_N	ODE_ROLE	AVAILABILITY-ZONE	INSTANCE-TYPE	SIZE	CPU CAPACITY	MEMORY CAPACI
					jek8s-361	master		N/A	N/A	N/A	8 cpus	31.26 G
ne i k	ProphetStor		2070		jek8s-363	N/A		N/A	N/A	N/A	4 cpus	15.51 G
			2078	5.52 \$/mo	Jek85-362	N/A		N/A	N/A	N/A	4 cpus	15.51 G
					јеказ-304	N/A		N/A	N/A	N/A	4 cpus	15.51 G
	Recommended Cl	uster - AWS 1d	Recommende	ed Cluster Configura	ion - AWS							
	830	22	PROVIDER	DISPLAY_NAME	REGION		INSTANCE_TYPE	ONDEMAND_INSTAN	RESERVED_INSTANCE	ONDEMAND_INSTAN	RESERVED_INSTANCE	COUNTRY
	039	• <b>4 4</b> \$/mo	aws	m5.4xlarge_16.0_co	res_ us_west_	oregon	m5.4xlarge	1	0	0	0	usa
	aurs	Savings 10	aws	m5.xlarge_4.0_core	s_16 us_west_	oregon	m5.xlarge	0	0	2	0	usa
		1239.30	aws	c5.large_2.0_cores_	4.0_ç us_west_	oregon	c5.large	1	0	2	U	usa
ng i	Recommended Cl	luster - Azure 1d	Recommende	ed Cluster Configura	ion - Azure							
			PROVIDER	DISPLAY_NAME	REGION		INSTANCE_TYPE	ONDEMAND_INSTAN	RESERVED_INSTANCE	ONDEMAND_INSTAN	RESERVED_INSTANCE	COUNTRY
	/86	.82 s/mo	azure	standard-d16s-v3_	6.0_ east_us		standard-d16s-v3	1	0	0	0	usa
	1	Savings 1d	azure	standard-d4s-v3_4.	0_co_east_us		standard-d4s-v3	0	0	2	0	usa
	Azure	1291.70	azure	standard-f2s-v2_2.	_cor east_us		standard-f2s-v2	1	0	2	0	usa
	Recommended Cl	luster - GCP 1d	Recommende	ed Cluster Configura	ion - GCP							
	570	45	PROVIDER	DISPLAY_NAME	REGION		INSTANCE_TYPE	ONDEMAND_INSTAN	RESERVED_INSTANCE	ONDEMAND_INSTAN	RESERVED_INSTANCE	COUNTRY
	570	. 15 s/mo	gcp	e2-standard-16_16	0_co_us_west_	1c	e2-standard-16	1	0	0	0	usa
	~	Savings 1d	gcp	e2-standard-4_4.0_	core us_west_	.1c	e2-standard-4	0	0	2	0	usa
			aro	e2-standard-2 2.0	ore us west	1c	e2-standard-2	1	0	2	0	usa

# Sysdig Dashboard Overview

The following Custom Sysdig Dashboards are available after Federator.ai is installed.

### Federator.ai Cluster Overview



## Federator.ai Application Overview



### Federator.ai Application Overview



# Federator.ai installation/uninstallation using Helm Chart

### Prerequisites

- Kubernetes version 1.18 or later
- OpenShift version 4.x.x or later
- Helm version is 3.x.x or later

#### Add Helm chart repository

~# helm repo add prophetstor https://prophetstor-ai.github.io/federatoraioperator-helm/

#### Test the Helm chart repository

```
~# helm search repo federatorai
```

### Installing with the release name my-name:

```
~# helm install `my-name` prophetstor/federatorai --namespace=federatorai --
create-namespace
```

## To uninstall/delete the my-name deployment:

```
~# helm ls --all-namespaces
helm delete `my-name` --namespace=federatorai
```

## Configuration

The following table lists the configurable parameters of the chart and their default values are specified in values.yaml.

Parameter	Description
<pre>image.pullPolicy</pre>	Container pull policy
image.repository	Image for Federator.ai operator
image.tag	Image Tag for Federator.ai operator
federatorai.imageLocation	Image Location for services containers
federatorai.version	Image Tag for services containers
federatorai.persistence.enabled	Enable persistent volumes
federatorai.persistence.storageClass	Storage Class Name of persistent volumes
<pre>federatorai.persistence.storages.logStorage.size</pre>	Log volume size
federatorai.persistence.aiCore.dataStorage.size	AICore data volume size
<pre>federatorai.persistence.influxdb.dataStorage.size</pre>	Influxdb data volume size
federatorai.persistence.fedemeterInfluxdb.dataStorage.size	Fedemeter influxdb data volume size
services.dashboardFrontend.nodePort	Port of the Dashboard service

Specify each parameter using the --set key=value[,key=value] argument to helm install

Tip: You can use the default values.yaml

Sample :valume.yaml

```
## Default values for Federator.ai
## This is a YAML-formatted file.
## Declare variables to be passed into your templates.
##
image:
    pullPolicy: IfNotPresent
    repository: quay.io/prophetstor/federatorai-operator-ubi
    tag: v5.0.0-ga
## Set default values
##
federatorai:
    imageLocation: quay.io/prophetstor
    version: v5.0.0-ga
```

```
## If the persistence is enabled, a default StorageClass
 ## is needed in the k8s cluster to provision volumes.
 persistence:
    enabled: true
    storageClass: "standard"
    storages:
      logStorage:
        size: 2Gi
    aiCore:
      dataStorage:
        size: 10Gi
    influxdb:
      dataStorage:
        size: 100Gi
    fedemeterInfluxdb:
      dataStorage:
       size: 10Gi
services:
 dashboardFrontend:
    ## Specify the nodePort value for the dashboard frontend
   ## Comment out the following line to disable nodePort service
    nodePort: 31012
 rest:
   ## Specify the nodePort value for the REST service
   ## Comment out the following line to disable nodePort service
    nodePort: 31011
```

## Alternative installation with configuration file

A YAML file that specifies the values for the parameters can be provided while installing the chart. For example

```
~# helm install `my-name` prophetstor/federatorai -f values.yaml --
namespace=federatorai --create-namespace
```

# Federator.ai installation/uninstallation using Ansible

### Only support Federator.ai since v4.4.0 or later

### Prerequisite

## Ansible Control Node

Software:	Version:	Query Command:
Ansible	2.10.2 or later	ansibleversion
Ansible Collection - community.kubernetes	1.1.1 or later	ansible-galaxy collection list or ansible-galaxy collection install community.kubernetes -vvv
Python	3.7 or later	python3version

OpenShift python client (Required by community.kubernetes collection)	0.11.2 or later	pip3 list   grep openshift
kubeconfig file (Need to copy target cluster's kubeconfig file to the Ansible Control Node)		e.g. file is put on /root/.kube/config.135

# Preparation (Ansible Control Node):

1. Install Ansible

https://docs.ansible.com/ansible/latest/installation\_guide/intro\_installation.html

2. Install collection "community.kubernetes"

~# ansible-galaxy collection install community.kubernetes

- 3. Install python & pip https://www.python.org/downloads/
- 4. Install OpenShift python client if you are using OpenShift clusters

~# pip3 install openshift

- 5. Download Ansible playbook for Federator.ai
- 6. Modify user\_variable.yaml file for customizing needed info.

## Installing Federator.ai

## Variables for in user\_variable.yml

Group	Variable Name	Sample value	Description	Mandatory
<u>Federator.ai</u> env	federatorai_version	v5.0.0-ga	Federator.ai version tag	Υ
Storage for <u>Federator</u> .ai pods	storage_type	ephemeral or persistent	Using ephemeral persistent volume type	Y
Storage info (Only be used when storage_type is persistent)	log_storage_size	10	Log size reserved for every pod. 10 means 10GB	Ν
Private repo	enable_private_repo	У	Using private repo to pull the <u>Federator.ai</u> required docker images	N
Pod resource	enable_resource_requi rement	У	Add pod resource requirement (limits & requests) for every <u>Federator.ai</u> pod	N

Expose services (Only be used when openshift_env is "n ")	expose_dashboard_an d_rest_api_services	У	Expose the dashboard and API services in the Kubernetes cluster.	Y
Cluster type	openshift_env	n	Input "y "if installed cluster is OpenShift cluster	Υ
	installed_namespace	federatorai	namespace where <u>Federator.ai</u> will be installed	Ν
	image_url_prefix		Input the private repo URL	N
	ai_engine_size	10	Storage size reserved for Alameda Al engine.	N
	influxdb_storage_size	100	Data size reserved for InfluxDB pod.	Ν
	storage_class_name	scname	To specifying storage class name for provisioning persistent volumes	Y

## Steps:

1. Go to Ansible playbook folder

~# cd ansible\_for\_federatorai

- 2. Modify user\_variable.yaml (under uninstaller folder) file for customizing needed info.
- 3. Export K8S\_AUTH\_KUBECONFIG to specify kubeconfig file path for Ansible collection (community.kubernetes).

~# export K8S\_AUTH\_KUBECONFIG=/root/.kube/config.135

4. Run Ansible playbook

~# ansible-playbook federtorai\_installation.yaml

## Uninstalling Federator.ai

For Uninstallation, please use the file under ansible\_for\_federatorai/uninstaller directory.

Variables in user\_variable.yml.

Group	Variable Name	Sample value	Description	Mandatory
Storage for <u>Federator</u> .ai pods	storage_type	ephemeral or persistent	Specify current <u>Federator.ai</u> storage type (ephemeral or persistent)	Y
Preserve current persistent volume (Only be used when storage_type is persistent)	preserve_pv	Y	Specify whether to preserve <u>Federator.ai</u> PVs	

## Steps:

1. Get to Ansible playbook uninstallation folder

~# cd ansible\_for\_federatorai/uninstaller

- 2. Modify user\_variable.yaml (under uninstaller folder) file for customizing needed info.
- 3. Export **K8S\_AUTH\_KUBECONFIG** to specify kubeconfig file path for Ansible collection (community.kubernetes)

~# export K8S\_AUTH\_KUBECONFIG=/root/.kube/config.135

4. Run Ansible playbook

~# ansible-playbook federatorai\_uninstaller.yaml

# Configure nativeHPA using external metrics for Datadog

Sample YAML file

```
#!Sample YAML
apiVersion: autoscaling/v2beta1
kind: HorizontalPodAutoscaler
metadata:
  name: <app1-hpa>
  namespace: <namespace>
spec:
  minReplicas: <num>
  maxReplicas: <num>
  scaleTargetRef:
    apiVersion: apps/v1
    kind: Deployment
    name: <deployment_name>
  metrics:
  - type: External
    external:
      metricName: federatorai.recommendation
      metricSelector:
```

```
matchLabels:
    resource: replicas
    kube_cluster: <clusterName>
    kube_deployment: <deployment_name>
    kube_namespace: <namespace>
targetAverageValue: 1
```

# Deploy KEDA for Remote and Sysdig monitored Cluster Pod Autoscaling

Step 1: Deploy KEDA.

\$helm repo add kedacore https://kedacore.github.io/charts
\$helm repo update
\$kubectl create namespace keda
\$helm install keda kedacore/keda --namespace keda

Example: Configure Generic Application HPA using KEDA ScaledObject



## Environment

- Kind: deployment
- Deployment Name: nginx-a1
- Namespace: nginx-shopping
- Ingress for Federator.ai API URL: "http://172.31.79.151:31011"

## Apply updated YAML

This file can be divided into 3 parts.

1. Secrets: storage for access token

Federatorai REST provide "basic" authentication mode, so we need to create "Secret" with

encoded base64 string (base64), in the example above, "YWRtaW4=" and "cGFzc3dvcmQ=" are

"admin" and "password" respectively.

```
$vi secret-shopping.yaml
apiVersion: v1
kind: Secret
metadata:
   name: keda-metric-api-secret
   namespace: nginx-shopping
data:
   username: "YWRtaW4="
```

```
password: "cGFzc3dvcmQ="
$kubect1 apply -f secret-shopping.yaml -n nginx-shopping
```

Get Federator.ai API certificated and patch secret

```
$CACRT=$(echo | openssl s_client -showcerts -connect 172.31.79.151:31011
2>/dev/null | openssl x509 | sed -n -e '/BEGIN\ CERTIFICATE/,/END\ CERTIFICATE/
p' | base64 | tr -d '\n')
```

\*" 31011" is Federator.ai API port access from external of cluster

```
$ kubectl -n nginx-shopping patch secret keda-metric-api-secret -p
"{\"data\":{\"ca.crt\":\"${CACRT}\"}}"
```

- 2. TriggerAuthentication: target to trigger authentication
- 3. ScaledObject: target to scale and how to trigger it with provided authentication method
- 4. URL format in ScaledObject configuration YAML:

"https://<clusterIP>:<federator\_api

\_port/apis/v1/recommendations/clusters/<cluster\_name>/namespaces/<namespace\_name

>/deployments/<demployment\_name>/metrics/<metric\_name>?limit=1&order=desc"

metric\_name:

- "controller\_desired\_replicas"
- "kafka\_consmer\_group\_desired\_replicas"
- "ingress\_upstream\_desired\_replicas"

	pollingInterval	cooldownPeriod
generic application	90	180
kafka consumer group	150	300
ingress upstream	30	60

```
$ vi app-scaledobject.yaml
#! - sample YAML
apiVersion: keda.sh/v1alpha1
kind: TriggerAuthentication
metadata:
  name: keda-metric-api-creds
  namespace: nginx-shopping
spec:
  secretTargetRef:
    - parameter: username
      name: keda-metric-api-secret
      key: username
    - parameter: password
      name: keda-metric-api-secret
      key: password
    - key: ca.crt
      name: keda-metric-api-secret
```

```
parameter: ca
- - -
apiVersion: keda.sh/v1alpha1
kind: ScaledObject
metadata:
 name: http-scaledobject
  namespace: nginx-shopping
 labels:
    deploymentName: nginx-a1
spec:
 minReplicaCount: 1
 maxReplicaCount: 10
 pollingInterval: 150
  cooldownPeriod: 300
  scaleTargetRef:
   name: nginx-a1
                   ←deployment name
 triggers:
    - type: metrics-api
      metadata:
        targetValue: "1"
        url: "https://172.31.79.151:31011/apis/v1/recommendations/clusters/k8s-
79161/namespaces/nginx-shopping/deployments/nginx-a1/metrics/
controller_desired_replicas?limit=1&order=desc"
        valueLocation: 'data.0.value'
        authMode: "basic"
      authenticationRef:
        name: keda-metric-api-creds
```

Apply YAML file to create KEDA scaled object

\$kubect] \$kubect]	l apply -f app l get ScaledOb	o-scaledobje oject -n ngi	ct.y nx-s	aml hopp	ing					
NAME http-scaledobject	SCALETARGETKIND apps/v1.Deployment	SCALETARGETNAME nginx-al	MIN 1	MAX 10	TRIGGERS metrics-api	AUTHENTICATION keda-metric-api-creds	READY True	ACTIVE True 🕅	FALLBACK False	AGE 25h
Make sure Re message.	ady, Active sta	tus is "True".	lf st	atus	is false, d	escribe ScaledOb	ject to	check	error	