Reference & Support

Data Infrastructure Insights

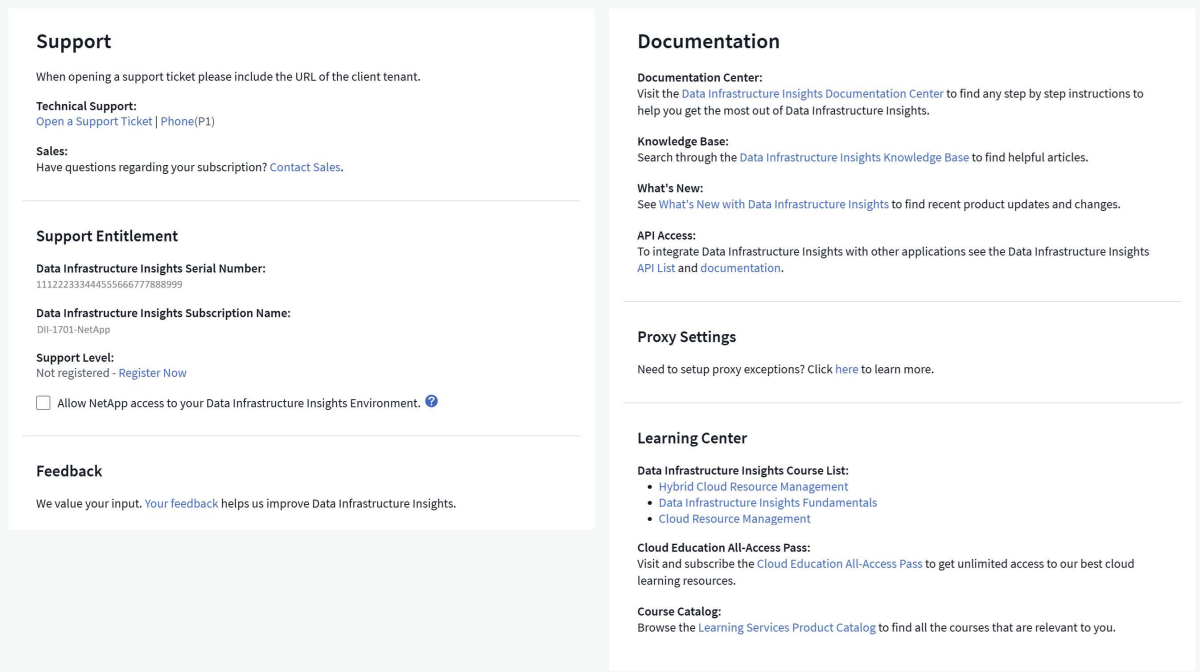
NetApp
February 03, 2026

Table of Contents

Reference & Support

Requesting Support

You can access available support options in Data Infrastructure Insights by selecting **Help > Support**.



Activating support entitlement

Data Infrastructure Insights offers self-service and email support when running in trial mode. Once you have subscribed to the service, it is strongly recommended that you activate support entitlement. Activating support entitlement enables you to access technical support via the web ticketing system and phone. The default support mode is self-service until registration is completed. See details below.

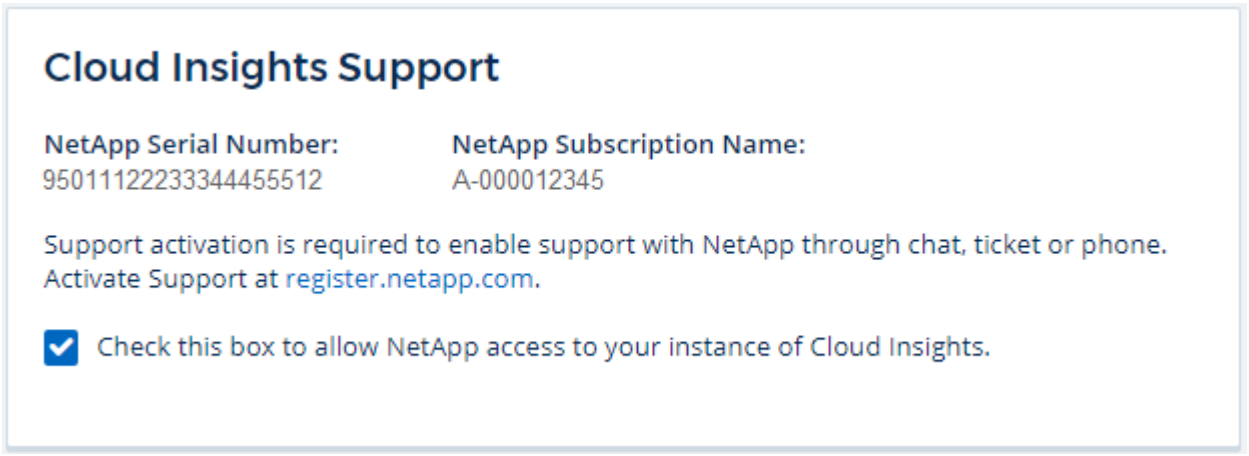
During the initial subscription process, your Data Infrastructure Insights instance will generate a 20-digit NetApp serial number starting with "950". This NetApp serial number represents the Data Infrastructure Insights subscription associated with your account. You must register the NetApp serial number to activate support entitlement. We offer two options for support registration:

1. User with pre-existing NetApp Support Site (NSS) SSO account (e.g. current NetApp customer)
2. New NetApp customer with no pre-existing NetApp Support Site (NSS) SSO account

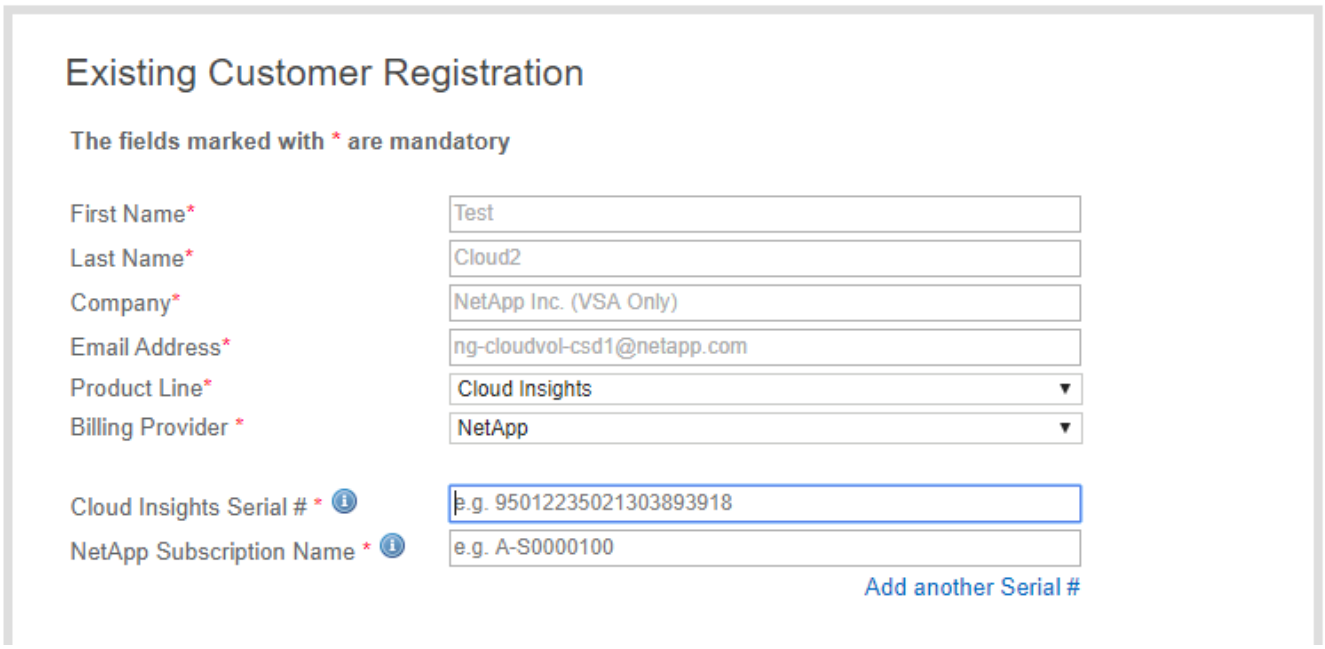
Option 1: Steps for a user with a pre-existing NetApp Support Site (NSS) SSO account

Steps

1. Navigate to the NetApp registration website https://register.netapp.com
2. Select “I am already registered as NetApp Customer” and choose *Data Infrastructure Insights* as the Product Line. Select your Billing Provider (NetApp or AWS) and provide your Serial Number and your NetApp Subscription Name or AWS Customer ID by referring to the “Help > Support” menu within the Data Infrastructure Insights user interface:



3. Complete the Existing Customer Registration form and click **Submit**.

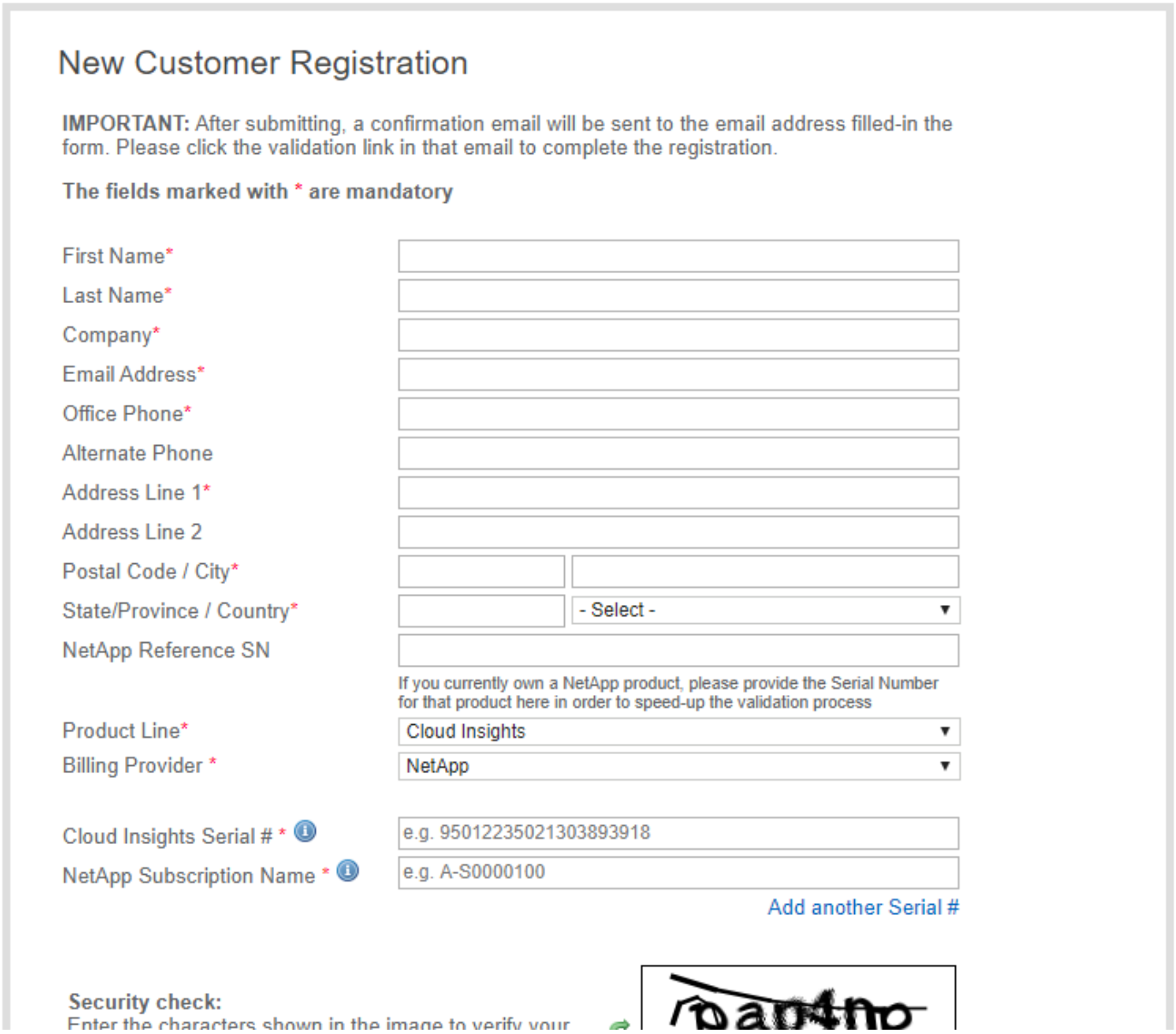


4. If no errors occur, user will be directed to a “Registration Submitted Successfully” page. The email address associated with the NSS SSO username used for registration will receive an email within a couple minutes stating “your product is now eligible for support”.
5. This is a onetime registration for the Data Infrastructure Insights NetApp serial number.

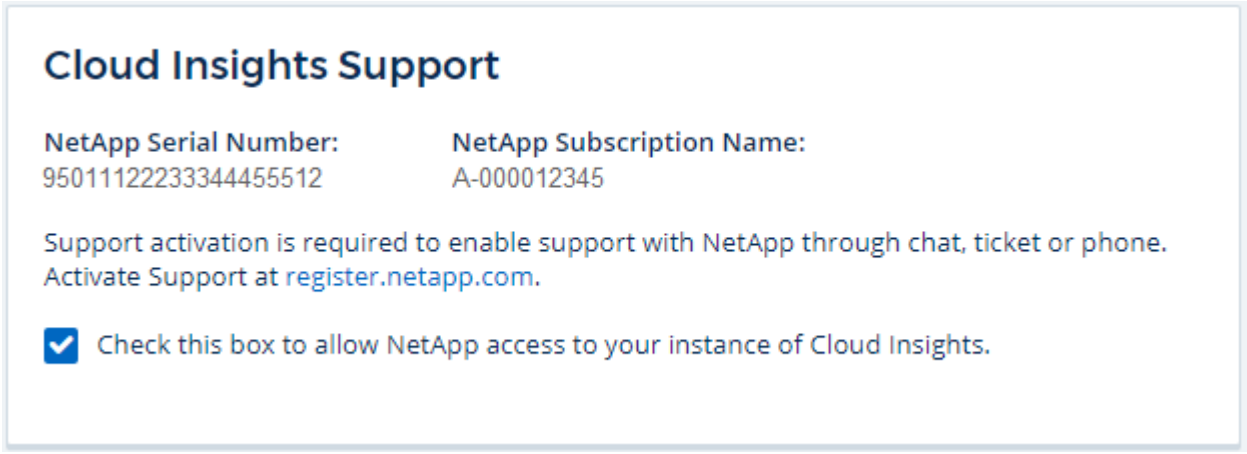
Option 2: Steps for a new NetApp customer with no pre-existing NetApp Support Site (NSS) SSO account

Steps

1. Navigate to the NetApp registration website https://register.netapp.com
2. Select “I am not a registered NetApp Customer” and complete the required information in example form below:



1. Select *Data Infrastructure Insights* as the Product Line. Select your Billing Provider (NetApp or AWS) and provide your Serial Number and your NetApp Subscription Name or AWS Customer ID by referring to the “Help > Support” menu within the Data Infrastructure Insights user interface:



2. If no errors occur, user will be directed to a “Registration Submitted Successfully” page. The email address associated with the NSS SSO username used for registration will receive an email within a few hours stating “your product is now eligible for support”.
3. As a new NetApp customer, you will also need to create a NetApp Support Site (NSS) user account for future registrations and access to support portal for technical support and web ticketing. This link is located at https://mysupport.netapp.com/eservice/public/now.do. You can provide your newly registered Data Infrastructure Insights serial number to expedite the process.
4. This is a one-time registration for the Data Infrastructure Insights NetApp serial number.

Obtaining Support Information

NetApp provides support for Data Infrastructure Insights in a variety of ways. Extensive free self-support options are available 24x7, such as knowledgebase (KB) articles or the NetApp community. For users who are subscribed to Data Infrastructure Insights, technical support is available via phone or web ticketing. A NetApp Support Site (NSS) SSO account is required for web ticket along with case management.

Self-Service Support:

These support options are available in Trial mode and are available for free 24x7:

- **KnowledgeBase**

  Clicking the links in this section takes you to the NetApp Knowledgebase, where you can search through relevant articles, how-to’s, and more.

- **Documentation**

  Clicking on the Documentation link takes you to this documentation center.

- **Community**

  Clicking on the community link takes you to the NetApp Data Infrastructure Insights community, where you can connect with peers and experts.

There is also a link to provide xref:./**Feedback** to help us improve Data Infrastructure Insights.

Subscription Support

In addition to the self-support options above, if you have a Data Infrastructure Insights subscription or paid support for monitored NetApp products or services, you can work with a NetApp Support Engineer to resolve your problem.

You must register in order to activate support for NetApp Cloud products. To register, go to NetApp’s Cloud Data Services Support Registration.

It is highly recommended that you check the box to allow a NetApp Support Engineer access to your Data Infrastructure Insights tenant during your support session. This will allow the engineer to troubleshoot the problem and help you resolve it quickly. When your issue is resolved or your support session has ended, you can un-check the box.

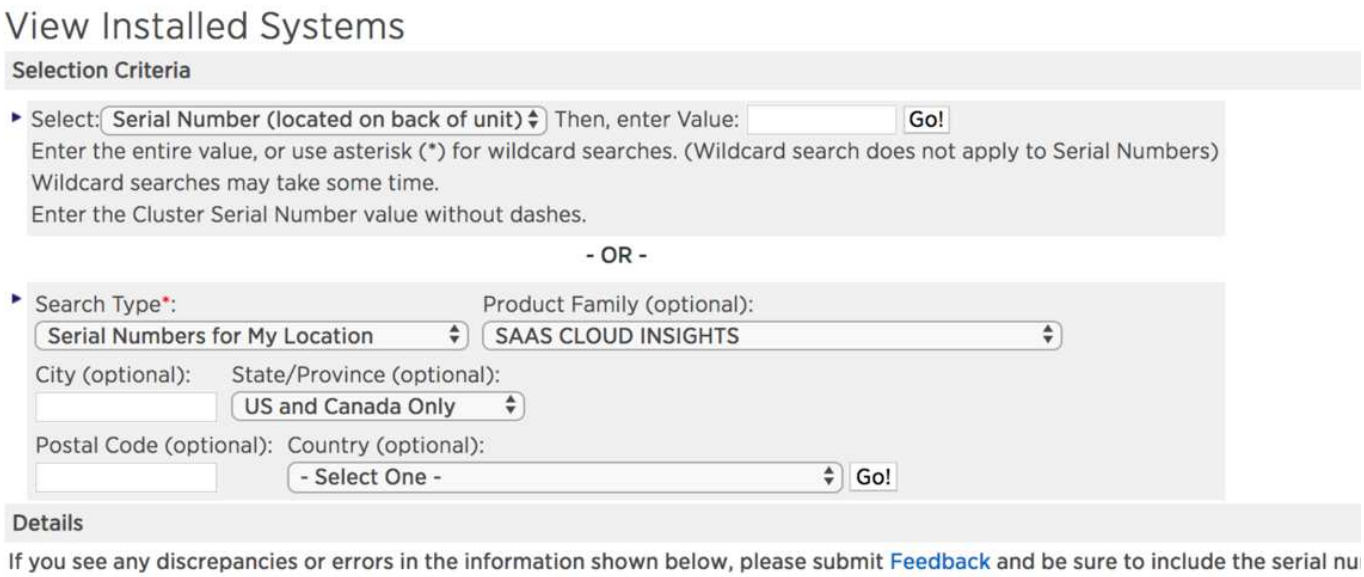
You can request support by any of the following methods. You must have an active Data Infrastructure Insights subscription to use these support options:

- Phone
- Support Ticket

You can also request sales support by clicking on the **Contact Sales** link.

Your Data Infrastructure Insights serial number is visible within the service from the **Help > Support** menu. If you are experiencing issues accessing the service and have registered a serial number with NetApp previously, you can also view your list of Data Infrastructure Insights serial numbers from the NetApp Support Site as follows:

- Login to mysupport.netapp.com
- From the Products > My Products menu tab, use Product Family “SaaS Data Infrastructure Insights” to locate all your registered serial numbers:



Data Infrastructure Insights Data Collector Support Matrix

You can view or download information and details about supported Data Collectors in the **Data Infrastructure Insights Data Collector Support Matrix**.

Learning Center

Regardless of your subscription, **Help > Support** links to several NetApp University course offerings to help you get the most out of Data Infrastructure Insights. Check them out!

Data Collector Reference - Infrastructure

Vendor-Specific Reference

The topics in this section provide vendor-specific reference information. In most cases, configuring a data collector is straightforward. In some cases, you may need additional information or commands to properly configure the data collector.

Click on a **vendor** in the menu to the left to see information for their data collectors.

Configuring the Amazon EC2 data collector

Data Infrastructure Insights uses the Amazon EC2 data collector to acquire inventory and performance data from EC2 instances.

Requirements

In order to collect data from Amazon EC2 devices, you must have the following information:

- You must have one of the following:
  - The **IAM Role** for your Amazon EC2 cloud account, if using IAM Role Authentication. IAM Role only applies if your acquisition unit is installed on an AWS instance.
  - The **IAM Access Key** ID and Secret Access Key for your Amazon EC2 cloud account, if using IAM Access Key authentication.
- You must have the "list organization" privilege
- Port 443 HTTPS
- EC2 Instances can be reported as a Virtual Machine, or (less naturally) a Host. EBS Volumes can be reported as both a VirtualDisk used by the VM, as well as a DataStore providing the Capacity for the VirtualDisk.

Access keys consist of an access key ID (for example, AKIAIOSFODNN7EXAMPLE) and a secret access key (for example, wJalrXUtnFEMI/K7MDENG/bPxRfiCYEXAMPLEKEY). You use access keys to sign programmatic requests that you make to EC2 if you use the Amazon EC2 SDKs, REST, or Query API operations. These keys are provided with your contract from Amazon.

Configuration

Enter data into the data collector fields according to the table below:

Field	Description
AWS Region	Choose AWS region
IAM Role	For use only when acquired on an AU in AWS. See below for more information on IAM Role.
AWS IAM Access Key ID	Enter AWS IAM Access Key ID. Required if you do not use IAM Role.
AWS IAM Secret Access Key	Enter AWS IAM Secret Access Key. Required if you do not use IAM Role.
I understand AWS bills me for API requests	Check this to verify your understanding that AWS bills you for API requests made by Data Infrastructure Insights polling.

Advanced Configuration

Field	Description
Include Extra Regions	Specify additional regions to include in polling.

Field	Description
Cross Account Role	Role for accessing resources in different AWS accounts.
Inventory Poll Interval (min)	The default is 60
Choose 'Exclude' or 'Include' to Apply to Filter VMs by Tags	Specify whether to include or exclude VM's by Tags when collecting data. If 'Include' is selected, the Tag Key field can not be empty.
Tag Keys and Values on which to Filter VMs	Click **+ Filter Tag** to choose which VMs (and associated disks) to include/exclude by filtering for keys and values that match keys and values of tags on the VM. Tag Key is required, Tag Value is optional. When Tag Value is empty, the VM is filtered as long as it matches the Tag Key.
Performance Poll Interval (sec)	The default is 1800
CloudWatch Agent Metrics Namespace	Namespace in EC2/EBS from which to collect data. Note that if the names of the default metrics in this namespace are changed, Data Infrastructure Insights may not be able to collect that renamed data. It is recommended to leave the default metric names.

IAM Access Key

Access keys are long-term credentials for an IAM user or the AWS account root user. Access keys are used to sign programmatic requests to the AWS CLI or AWS API (directly or using the AWS SDK).

Access keys consist of two parts: an access key ID and a secret access key. When you use *IAM Access Key* authentication (as opposed to *IAM Role* authentication), you must use both the access key ID and secret access key together for authentication of requests. For more information, see the Amazon documentation on Access Keys.

IAM Role

When using *IAM Role* authentication (as opposed to IAM Access Key authentication), you must ensure that the role you create or specify has the appropriate permissions needed to access your resources.

For example, if you create an IAM role named *InstanceEc2ReadOnly*, you must set up the policy to grant EC2 read-only list access permission to all EC2 resources for this IAM role. Additionally, you must grant STS (Security Token Service) access so that this role is allowed to assume roles cross accounts.

After you create an IAM role, you can attach it when you create a new EC2 instance or any existing EC2 instance.

After you attach the IAM role *InstanceEc2ReadOnly* to an EC2 instance, you will be able to retrieve the temporary credential through instance metadata by IAM role name and use it to access AWS resources by any application running on this EC2 instance.

For more information see the Amazon documentaiton on IAM Roles.

Note: IAM role can be used only when the Acquisition Unit is running in an AWS instance.

Mapping Amazon tags to Data Infrastructure Insights annotations

The Amazon EC2 data collector includes an option that allows you to populate Data Infrastructure Insights annotations with tags configured on EC2. The annotations must be named exactly as the EC2 tags. Data Infrastructure Insights will always populate same-named text-type annotations, and will make a "best attempt" to populate annotations of other types (number, boolean, etc). If your annotation is of a different type and the data collector fails to populate it, it may be necessary to remove the annotation and re-create it as a text type.

Note that AWS is case-sensitive, while Data Infrastructure Insights is case-insensitive. So if you create an annotation named "OWNER" in Data Infrastructure Insights, and tags named "OWNER", "Owner", and "owner" in EC2, all of the EC2 variations of "owner" will map to Cloud Insight's "OWNER" annotation.

Include Extra Regions

In the AWS Data Collector **Advanced Configuration** section, you can set the **Include extra regions** field to include additional regions, separated by comma or semi-colon. By default, this field is set to ***us-.**, which collects on all US AWS regions. To collect on *all* regions, set this field to ***.**.
If the **Include extra regions** field is empty, the data collector will collect on assets specified in the **AWS Region** field as specified in the **Configuration** section.

Collecting from AWS Child Accounts

Data Infrastructure Insights supports collection of child accounts for AWS within a single AWS data collector. Configuration for this collection is performed in the AWS environment:

- You must configure each child account to have an AWS Role that allows the main account ID to access EC2 details from the children account.
- Each child account must have the role name configured as the same string.
- Enter this role name string into the Data Infrastructure Insights AWS Data Collector **Advanced Configuration** section, in the **Cross account role** field.
- The account where the collector is installed needs to have *delegate access administrator* privileges. See the AWS Documentation for more information.

Best Practice: It is highly recommended to assign the AWS predefined *AmazonEC2ReadOnlyAccess* policy to the EC2 main account. Also, the user configured in the data source should have at least the predefined *AWSOrganizationsReadOnlyAccess* policy assigned, in order to query AWS.

Please see the following for information on configuring your environment to allow Data Infrastructure Insights to collect from AWS child accounts:

Tutorial: Delegate Access Across AWS Accounts Using IAM Roles

AWS Setup: Providing Access to an IAM User in Another AWS Account That You Own

Creating a Role to Delegate Permissions to an IAM User

Troubleshooting

Additional information on this Data Collector may be found from the Support page or in the Data Collector Support Matrix.

Amazon FSx for NetApp ONTAP data collector

This data collector acquires inventory and performance data from Amazon FSx for NetApp ONTAP. This data collector will be made available incrementally throughout the Data Infrastructure Insights service regions. Contact your sales person if you do not see the Icon for this collector in your Data Infrastructure Insights Environment.

This Data Infrastructure Insights collector requires an ONTAP user with a *Filesystem-Scoped* role. Please review the AWS Roles and Rules documentation for available options. At this time AWS supports only one kind of user role with Filesystem Scope, which is *fsxadmin*. This is the appropriate role to be used for the Data Infrastructure Insights collector. The user should also have all three of these applications assigned to it: http, ontapi, ssh.

Terminology

Data Infrastructure Insights acquires inventory and performance data from the FSx-NetApp data collector. For each asset type acquired, the most common terminology used for the asset is shown. When viewing or troubleshooting this data collector, keep the following terminology in mind:

Vendor/Model Term	Data Infrastructure Insights Term
Cluster	Storage
LUN	Volume
Volume	Internal Volume

FSx-NetApp Terminology

The following terms apply to objects or references that you might find on FSx-NetApp storage asset landing pages. Many of these terms apply to other data collectors as well.

Storage

- Model – A comma-delimited list of the unique, discrete model names within this cluster.
- Vendor – AWS
- Serial number – The array serial number.
- IP – generally will be the IP(s) or hostname(s) as configured in the data source.
- Raw Capacity – base 2 summation of all the SSD storage assigned to the FSx filesystem.
- Latency – a representation of what the host facing workloads are experiencing, across both reads and writes. Ideally, Data Infrastructure Insights is sourcing this value directly, but this is often not the case. In lieu of the array offering this up, Data Infrastructure Insights is generally performing an IOPs-weighted calculation derived from the individual internal volumes' statistics.
- Throughput – aggregated from internal volumes.
  Management – this may contain a hyperlink for the management interface of the device. Created programmatically by the Data Infrastructure Insights data source as part of inventory reporting.

Storage Pool

- Storage – what storage array this pool lives on. Mandatory.

- Type – a descriptive value from a list of an enumerated list of possibilities. Most commonly will be “Aggregate” or “RAID Group””.
- Capacity – the values here are the logical used, usable capacity and the logical total capacity, and the percentage used across these.
- IOPS – the sum IOPs of all the volumes allocated on this storage pool.
- Throughput – the sum throughput of all the volumes allocated on this storage pool.

Requirements

The following are requirements to configure and use this data collector:

- You must have access to an account with the "fsxadmin" role, with three applications assigned to it - ssh, ontapi, http
- Account details include username and password.
- Port requirements: 443

Configuration

Field	Description
NetApp Management IP	IP address or fully-qualified domain name of the NetApp cluster
User Name	User name for NetApp cluster
Password	Password for NetApp cluster

Advanced Metrics

This data collector collects the following advanced metrics from the FSx for NetApp ONTAP storage:

- fpolicy
- nfsv3
- nfsv3:node
- nfsv4
- nfsv4_1
- nfsv4_1:node
- nfsv4:node
- policy_group
- qtree
- volume
- workload_volume

Note that FSx CLI and API commands retrieve some capacity values that Data Infrastructure Insights ZAPI doesn’t collect, so certain capacity values (such as those for storage pools) may be different in Data Infrastructure Insights than they are on the FSx itself.

Troubleshooting

Some things to try if you encounter problems with this data collector:

Inventory

Problem:	Try this:
Receive 401 HTTP response or 13003 ZAPI error code and ZAPI returns “Insufficient privileges” or “not authorized for this command”	Check username and password, and user privileges/permissions.
ZAPI returns "cluster role is not cluster_mgmt LIF"	AU needs to talk to cluster management IP. Check the IP and change to a different IP if necessary
ZAPI command fails after retry	AU has communication problem with the cluster. Check network, port number, and IP address. User should also try to run a command from command line from the AU machine.
AU failed to connect to ZAPI via HTTP	Check whether ZAPI port accepts plaintext. If AU tries to send plaintext to an SSL socket, the communication fails.
Communication fails with SSLException	AU is attempting to send SSL to a plaintext port on a filer. Check whether the ZAPI port accepts SSL, or use a different port.
Additional Connection errors: ZAPI response has error code 13001, “database is not open” ZAPI error code is 60 and response contains “API did not finish on time” ZAPI response contains “initialize_session() returned NULL environment” ZAPI error code is 14007 and response contains “Node is not healthy”	Check network, port number, and IP address. User should also try to run a command from command line from the AU machine.

Additional information may be found from the Support page or in the Data Collector Support Matrix.

Configuring the Azure compute data collector

Data Infrastructure Insights uses the Azure compute data collector to acquire inventory and performance data from Azure compute instances.

Requirements

You need the following information to configure this data collector.

- Port requirement: 443 HTTPS
- Azure OAuth 2.0 Redirect URI (login.microsoftonline.com)

- Azure Management Rest IP (management.azure.com)
- Azure Resource Manager IP (management.core.windows.net)
- Azure Service Principal Application (Client) ID (Reader role required)
- Azure service principal authentication key (user password)
- You need to set up an Azure account for Data Infrastructure Insights discovery.

  Once the account is properly configured and you register the application in Azure, you will have the credentials required to discover the Azure instance with Data Infrastructure Insights. The following link describes how to set up the account for discovery.
  https://docs.microsoft.com/en-us/azure/active-directory/develop/howto-create-service-principal-portal

Configuration

Enter data into the data collector fields according to the table below:

Field	Description
Azure Service Principal Application (Client) ID (Reader role required)	Sign-in ID to Azure. Requires Reader Role access.
Azure tenant ID	Microsoft tenant ID
Azure Service Principal Authentication Key	Login authentication key
I understand Microsoft bills me for API requests	Check this to verify your understanding that Microsoft bills you for API requests made by Insight polling.

Advanced Configuration

Field	Description
Inventory Poll Interval (min)	The default is 60
Choose 'Exclude' or 'Include' to Apply to Filter VMs by Tags	Specify whether to include or exclude VM's by Tags when collecting data. If 'Include' is selected, the Tag Key field can not be empty.
Tag Keys and Values on which to Filter VMs	Click **+ Filter Tag** to choose which VMs (and associated disks) to include/exclude by filtering for keys and values that match keys and values of tags on the VM. Tag Key is required, Tag Value is optional. When Tag Value is empty, the VM is filtered as long as it matches the Tag Key.
Performance Poll Interval (sec)	The default is 300

Troubleshooting

Additional information on this Data Collector may be found from the Support page or in the Data Collector Support Matrix.

Broadcom

Brocade Network Advisor data collector

Data Infrastructure Insights uses the Brocade Network Advisor data collector to acquire inventory and performance data from Brocade switches.

Terminology

Data Infrastructure Insights acquires the following inventory information from the Brocade Network Advisor data collector. For each asset type acquired by Data Infrastructure Insights, the most common terminology used for this asset is shown. When viewing or troubleshooting this data collector, keep the following terminology in mind:

Vendor/Model Term	Data Infrastructure Insights Term
Switch	Switch
Port	Port
Virtual Fabric, Physical Fabric	Fabric
Logical Switch	Logical Switch

Note: These are common terminology mappings only and might not represent every case for this data collector.

Requirements

The following are required to configure this data collector:

- The Data Infrastructure Insights Acquisition Unit will initate connections to TCP port 443 on the BNA server. BNA server must be running version 14.2.1 or higher.
- Brocade Network Advisor Server IP address
- User name and password to an administrator account
- Port requirement: HTTP/HTTPS 443

Configuration

Field	Description
Brocade Network Advisor Server IP	IP address of the Network Advisor Server
User Name	User name for the switch
User Name	Administrator user name
Password	Administrator password

Advanced configuration

Field	Description
Connection Type	HTTPS (default port 443) or HTTP (default port 80)
Override Connection Port	If blank, use the default port in the Connection Type field, otherwise enter the connection port to use

Field	Description
Password	Password for the switch
Inventory poll interval (min)	The default is 40
Report Access Gateway	Check to include devices in Access Gateway mode
Performance Poll Interval (sec)	The default is 1800

Troubleshooting

Some things to try if you encounter problems with this data collector:

Inventory

Problem:	Try this:
Receive a message that more than 1 node is logged into the Access Gateway port, or data collector fails to discover Access Gateway device.	Check that the NPV device is operating correctly and that all connected WWNs are expected. Do not directly acquire the NPV device. Instead, acquisition of the core fabric switch will collect the NPV device data.

Additional information may be found from the Support page or in the Data Collector Support Matrix.

Brocade FC Switch data collector

Data Infrastructure Insights uses the Brocade FC Switch (SSH) data source to discover inventory for Brocade or rebranded switch devices running Factored Operating System (FOS) firmware 4.2 and later. Devices in both FC switch and Access Gateway modes are supported.

Terminology

Data Infrastructure Insights acquires the following inventory information from the Brocade FC Switch data collector. For each asset type acquired by Data Infrastructure Insights, the most common terminology used for this asset is shown. When viewing or troubleshooting this data collector, keep the following terminology in mind:

Vendor/Model Term	Data Infrastructure Insights Term
Switch	Switch
Port	Port
Virtual Fabric, Physical Fabric	Fabric
Zone	Zone
Logical Switch	Logical Switch
Virtual Volume	Volume
LSAN Zone	IVR Zone

Note: These are common terminology mappings only and might not represent every case for this data

collector.

Requirements

- The Data Infrastructure Insights Acquisition Unit (AU) will initiate connections to TCP Port 22 on Brocade switches to collect inventory data. The AU will also initiate connections to UDP port 161 for collection of performance data.
- There must be IP connectivity to all switches in the fabric. If you select the Discover all switches in the fabric check box, Data Infrastructure Insights identifies all the switches in the fabric; however, it needs IP connectivity to these additional switches to discover them.
- The same account is needed globally across all switches in the fabric. You can use PuTTY (open source terminal emulator) to confirm access.
- Ports 161 and 162 must be open to all switches in the fabric for SNMP performance polling.
- SNMP read-only Community String

Configuration

Field	Description
Switch IP	IP address or fully-qualified domain name of the EFC Server
User Name	User name for the switch
Password	Password for the switch
SNMP	SNMP version
SNMP Community String	SNMP read-only community string used to access the switch
SNMP User Name	SNMP user name
SNMP Password	SNMP password

Advanced configuration

Field	Description
Fabric name	Fabric name to be reported by the data collector. Leave blank to report the fabric name as WWN.
Inventory Poll Interval (min)	Interval between inventory polls. The default is 15.
Excluded Devices	Comma-separated list of device IDs to exclude from polling
Admin Domains Active	Select if using Admin Domains
Retrieve MPR Data	Select to acquire routing data from your multiprotocol router.
Enable Trapping	Select to enable acquisition upon receiving an SNMP trap from the device. If you select enable trapping, you must also activate SNMP.
Minimum Time Between Traps (sec)	Minimum time between acquisition attempts triggered by traps. The default is 10.

Field	Description
Discover all switches in the fabric	Select to discover all switches in the fabric
Choose Favoring HBA vs. Zone Aliases	Choose whether to favor HBA or zone aliases
Performance Poll Interval (sec)	Interval between performance polls. The default is 300.
SNMP Auth Protocol	SNMP authentication protocol (SNMP v3 only)
SNMP Privacy Password	SNMP privacy password (SNMP v3 only)
SNMP Retries	Number of SNMP retry attempts

Troubleshooting

Some things to try if you encounter problems with this data collector:

Inventory

Problem:	Try this:
The inventory acquisition of the Brocade datasource fails with the error: <date> <time> ERROR [com.onaro.sanscreen.acquisition.framework.datasource.BaseDataSource] Error 2 out of 2: <datasource name> [Internal error] - Unable to generate the model for device <IP>. Error detecting prompt ([Device name <name>]: Unable to generate the model for device <IP>. Error detecting prompt)	The issue may be caused when the Brocade switch takes too long to return with a prompt, exceeding the default timeout of 5 seconds. In the data collector's Advanced Configuration settings in Data Infrastructure Insights, try increasing the *SSH Banner Wait Timeout (sec)* to a higher value.
Error: "Data Infrastructure Insights received Invalid Chassis Role"	Check that the user configured in this data source has been granted the chassis role permission.
Error: "Mismatched Chassis IP Address"	DII does NOT support Network Address Translation or Port Address Translation between the Acquisition Unit and devices as a general rule. DII may be detecting that the hostname/IP address in the collector configuration does NOT match any of the addresses the device believes it has.
Receive a message that more than 1 node is logged into the Access Gateway port	Check that the NPV device is operating correctly and that all connected WWNs are expected. Do not directly acquire the NPV device. Instead, acquisition of the core fabric switch will collect the NPV device data.
Error: ….Max remote sessions for login…	FOS has different limits for the number of supported concurrent SSH sessions per user role. DII's SSH session to this device is being rejected at login for violating those limits. This can be a sign that you have duplicative collectors discovering the same asset, which should be avoided

Performance

Problem:	Try this:
Performance collection fails with "Timed out during sending SNMP request".	Depending on query variables and switch configuration, some queries may exceed the default timeout. Learn More.
Performance collection fails with …Row duplicates found in SNMP table…	DII has detected bad SNMP responses. You are likely running FOS 8.2.3e. Upgrade to 8.2.3e2 or higher.
Performance collections fails with …Unknown user name…	You have configured your DII collector with a "SNMP User Name" value which is not slotted into one of the SNMPv3 user slots. Simply creating a user on Brocade FOS does NOT necessarily enable it as a SNMPv3 user - you must place it in one of the v3 user slots.
Performance collections fails with …Unsupported Security Level…	You have configured your DII collector to use SNMPv3, but with encryption (aka privacy) and/or authorization settings which are not enabled on the device in question.
Performance collection fails with …Empty Privacy Password is only allowed for Privacy Protocol NONE	You have configured your DII collector to use SNMPv3, with an encryption aka privacy protocol (AES, etc), but you have an empty "SNMP Privacy Password" value, so DII cannot negotiate encrypted SNMPv3 data flows with this device
Performance collection fails with …..VF:nn, error: No access…	You have configured your DII collector to use SNMPv3 on a device with multiple Virtual Fabrics enabled, but the SNMPv3 user does NOT have rights for VF NN. DII does not support partial discovery of a physical asset - you should always grant DII access to all 128 possible VFs proactively as DII will always attempt to retrieve performance data for any extant VF on a given physical device

Additional information may be found from the Support page or in the Data Collector Support Matrix.

Brocade FOS REST Data Collector

Data Infrastructure Insights uses the Brocade FOS REST collector to discover inventory and performance for Brocade switch devices running FabricOS (FOS) firmware 8.2 and later. Please note that early 8.2 FOS releases may have REST API bugs; it is highly recommended to run the latest possible FOS release your platform supports.

Please note: FOS' default "user" level is insufficiently powerful for Data Infrastructure Insights to view all the logical aspects of a device - we need a user account with the "Chassis Role" enabled, as well as permissions on all the Virtual Fabrics configured on a switch.

Here is an example of how you can create a "least privilege" user account for Data Infrastructure Insights usage in a SSH session to a FOS device:

userConfig --add NetAppCIUser -r user -l 1-128 -c user -p Qwerty!

This will create a user “NetAppCIUser” with a password of “Qwerty!”. This user has the “user” role (-r) across all 128 possible virtual fabrics (-l). This user additionally has the required “Chassis” role (-c) with user level access assigned.

By default, this collector will attempt to discover all the FOS devices that are a part of all the fabrics the switch is part of.

Please note: FOS' default read-only user "user" does NOT have view permissions on all Virtual Fabrics, nor does it have "chassis role" permissions. This means you will have a low likelihood of success using "user" with Data Infrastructure Insights, which needs to understand both the physical and logical configuration of the FOS device.

Terminology

Data Infrastructure Insights acquires the following inventory information from the Brocade FOS REST data collector. For each asset type acquired by Data Infrastructure Insights, the most common terminology used for this asset is shown. When viewing or troubleshooting this data collector, keep the following terminology in mind:

Vendor/Model Term	Data Infrastructure Insights Term
Switch	Switch
Port	Port
Virtual Fabric, Physical Fabric	Fabric
Zone	Zone
Logical Switch	Logical Switch
LSAN Zone	IVR Zone

Note: These are common terminology mappings only and might not represent every case for this data collector.

Requirements

- There must be TCP connectivity to all switches in the fabric. This data collector type will seamlessly try both HTTP and HTTPS for each device in the fabric. If you select the *Discover all switches in the fabric* check box, Data Infrastructure Insights identifies all the switches in the fabric; however, it needs TCP connectivity to these additional switches to discover them.
- The same account is needed globally across all switches in the fabric. You can use the device’s Web interface to confirm access.

Configuration

Field	Description
Switch IP	IP address or fully-qualified domain name of the FOS switch
User Name	User name for the switch
Password	Password for the switch

Advanced configuration

Field	Description
Excluded Devices	Comma-separated list of device IPv4 addresses to exclude from polling.
Inventory Poll Interval (min)	Interval between inventory polls. The default is 60.
Discover all switches in the fabric	Select to discover all switches in the fabric.
Choose Favoring HBA vs. Zone Aliases	Choose whether to favor HBA or zone aliases.
Connection type	HTTP or HTTPS.
Note that this setting only changes which protocol CI attempts to use per device first - CI will attempt the opposite protocol automatically if the default fails	Override TCP Port
Specify a port if not using the default.	Performance Poll Interval (sec)

Troubleshooting

Some things to try if you encounter problems with this data collector:

Inventory

Problem:	Try this:
The Test feature warns me that a protocol is inaccessible	A given Brocade FOS 8.2+ device will only want to speak on HTTP or HTTPS - if a switch has a digital certificate installed, the switch will throw HTTP errors if one attempts to communicate to it with unencrypted HTTP versus HTTPS. The test feature attempts communication with both HTTP and HTTPS - if the Test tells you that one protocol passes, you can safely save the collector and not worry that the other protocol was unsuccessful - the collector will attempt both protocols during collection, and only fail if neither works.
Error: Inventory fails with 401 Unauthorized…Invalid Session Key…	This is distinct bug within some very early 8.2 FOS releases lie 8.2.1c that do NOT properly support HTTP basic authentication. Upgrade to a later 8.2 or 9.* release
Error: “Data Infrastructure Insights received Invalid Chassis Role”	Check that the user configured in this data source has been granted the chassis role permission.
Error: "Mismatched Chassis IP Address"	Change the data source configuration to use chassis IP address.
Inventory fails with a 403 Forbidden	This may simply be bad credentials, or it may be indicative that you are attempting to use an insufficiently powerful role - remember that "user" level users do NOT have the required "Chassis Role" right, or view access to non default Virtual Fabrics.

Additional information may be found from the Support page or in the Data Collector Support Matrix.

Cisco MDS Fabric Switches data collector

Data Infrastructure Insights uses the Cisco MDS Fabric Switches data collector to discover inventory for Cisco MDS Fabric Switches as well as a variety of Cisco Nexus FCoE switches on which the FC service is enabled.

Additionally, you can discover many models of Cisco devices running in NPV mode with this data collector.

Terminology

Data Infrastructure Insights acquires the following inventory information from the Cisco FC Switch data collector. For each asset type acquired by Data Infrastructure Insights, the most common terminology used for this asset is shown. When viewing or troubleshooting this data collector, keep the following terminology in mind:

Vendor/Model Term	Data Infrastructure Insights Term
Switch	Switch
Port	Port
VSAN	Fabric
Zone	Zone
Logical Switch	Logical Switch
Name Server Entry	Name Server Entry
Inter-VSAN Routing (IVR) Zone	IVR Zone

Note: These are common terminology mappings only and might not represent every case for this data collector.

Requirements

- An IP address of one switch in the fabric or individual switches
- Chassis discovery, to enable fabric discovery
- If using SNMP V2, read-only community string
- Port 161 is used to access the device

Configuration

Field	Description
Cisco Switch IP	IP address or fully-qualified domain name of the switch
SNMP Version	Select V1, V2, or V3. V2 or later is required for performance acquisition.
SNMP Community String	SNMP read-only community string used to access the switch (not applicable for SNMP v3)
User Name	User name for the switch (SNMP v3 only)

Field	Description
Password	Password used for the switch (SNMPv3 only)

Advanced configuration

Field	Description
Inventory Poll Interval (min)	Interval between inventory polls (default 40 minutes)
SNMP Auth Protocol	SNMP authentication protocol (SNMPv3 only)
SNMP Privacy Protocol	SNMP privacy protocol (SNMPv3 only)
SNMP Privacy Password	SNMP Privacy Password
SNMP Retries	Number of SNMP retry attempts
SNMP Timeout (ms)	SNMP timeout (default 5000 ms)
Enable Trapping	Select to enable trapping. If you enable trapping, you must also activate SNMP notifications.
Minimum Time Between Traps (sec)	Minimum time between acquisition attempts triggered by traps (default 10 seconds)
Discover All Fabric Switches	Select to discover all switches in the fabric
Excluded Devices	Comma-separated list of device IPs to exclude from polling
Included Devices	Comma-separated list of device IPs to include in polling
Check Device Type	Select to accept only those devices that explicitly advertise themselves as Cisco devices
First Alias Type	Provide a first preference for resolution of the alias. Choose from the following: **Device Alais** This is a user-friendly name for a port WWN (pWWN) that can be used in all configuration commands, as required. All switches in the Cisco MDS 9000 Family support Distributed Device Alias Services (device aliases). **None** Do not report any alias. **Port Description** A description to help identify the port in a list of ports. **Zone Alias (all)** A user-friendly name for a port that can be used only for the active configuration. This is the default.
Second Alias Type	Provide a second preference for resolution of the alias
Third Alias Type	Provide a third preference for resolution of the alias

Field	Description
Enable SANTap Proxy Mode Support	Select if your Cisco switch is using SANTap in proxy mode. If you are using EMC RecoverPoint, then you are probably using SANTap.
Performance Poll Interval (sec)	Interval between performance polls (default 300 seconds)

Troubleshooting

Some things to try if you encounter problems with this data collector:

Inventory

Problem:	Try this:
Error: Failed to discover chassis - no switches have been discovered	• Ping the device with the IP configured • Login to the device using Cisco Device Manager GUI • Login to the device using CLI • Try to run SNMP walk
Error: Device is not a Cisco MDS switch	• Make sure the data source IP configured for the device is correct • Login to the device using Cisco Device Manager GUI • Login to the device using CLI
Error: Data Infrastructure Insights is not able to obtain the switch's WWN.	This may not be a FC or FCoE switch, and as such may not be supported. Make sure the IP/FQDN configured in the datasource is truly a FC/FCoE switch.
Error: Found more than one nodes logged into NPV switch port	Disable direct acquisition of the NPV switch
Error: Could not connect to the switch	• Make sure the device is UP • Check the IP address and listening port • Ping the device • Login to the device using Cisco Device Manager GUI • Login to the device using CLI • Run SNMP walk

Performance

Problem:	Try this:
Error: Performance acquisition not supported by SNMP v1	• Edit Data Source and disable Switch Performance • Modify Data Source and switch configuration to use SNMP v2 or higher

Additional information may be found from the Support page or in the Data Collector Support Matrix.

Cohesity SmartFiles data collector

This REST API-based collector will acquire a Cohesity cluster, discovering the "Views"

(as Data Infrastructure Insights Internal Volumes), the various nodes, as well as collecting performance metrics.

Configuration

Field	Description
Cohesity Cluster IP	IP address of the Cohesity cluster
User Name	User name for the Cohesity cluster
Password	Password used for the Cohesity cluster

Advanced configuration

Field	Description
TCP Port	Port used for TCP communication with the Cohesity cluster
Inventory Poll Interval (min)	Interval between inventory polls. The default is 60 minutes.
Performance Poll Interval (min)	Interval between performance polls. The default is 900 seconds.

Troubleshooting

Additional information may be found from the Support page or in the Data Collector Support Matrix.

Dell

Dell EMC XC Series data collector

Data Infrastructure Insights uses this data collector to discover inventory and performance information for the Dell EMC XC Series storage arrays.

Configuration

Field	Description
Prism External IP Address	IP address of the XC server
User Name	User name for the XC server
Password	Password used for the XC server

Advanced configuration

Field	Description
TCP Port	Port used for TCP communication with the XC server
Inventory Poll Interval (min)	Interval between inventory polls. The default is 60 minutes.

Field	Description
Performance Poll Interval (min)	Interval between performance polls. The default is 300 seconds.

Troubleshooting

Additional information may be found from the Support page or in the Data Collector Support Matrix.

Dell EMC

DELL EMC Data Domain data collector

This data collector gathers inventory and performance information from DELL EMC Data Domain deduplication storage systems. To configure this data collector, there are specific configuration instructions and usage recommendations you must follow.

Terminology

Data Infrastructure Insights acquires the following inventory information from the Data Domain data collector. For each asset type acquired by Data Infrastructure Insights, the most common terminology used for this asset is shown. When viewing or troubleshooting this data collector, keep the following terminology in mind:

Vendor/Model Term	Data Infrastructure Insights Term
Disk	Disk
Array	Storage
FC Port	Port
File System	Internal Volume
Quota	Quota
NFS and CIFS share	FileShare

Note: These are common terminology mappings only and might not represent every case for this data colletor.

Requirements

You need the following information to configure this data collector:

- IP address of the Data Domain device
- Read-only user name and password to the Data Domain storage
- SSH port 22

Configuration

Field	Description
IP address	The IP address or fully-qualified domain name of the Data Domain storage array
User name	The user name for the Data Domain storage array

Field	Description
Password	The password for the Data Domain storage array

Advanced configuration

Field	Description
Inventory Poll Interval (min)	Interval between inventory polls. The default is 20.
SSH Port	SSH service port

Troubleshooting

Additional information on this Data Collector may be found from the Support page or in the Data Collector Support Matrix.

Configuring the EMC ECS data collector

This data collector acquires inventory and performance data from EMC ECS storage systems. For configuration, the data collector requires an IP address or hostname of the ECS cluster and a username and password.

Dell EMC ECS is metered at a different Raw TB to Managed Unit rate. Every 40 TB of unformatted ECS capacity is charged as 1 Managed Unit (MU).

Terminology

Data Infrastructure Insights acquires the following inventory information from the ECS data collector. For each asset type acquired, the most common terminology used for this asset is shown. When viewing or troubleshooting this data collector, keep the following terminology in mind:

Vendor/Model Term	Data Infrastructure Insights Term
Cluster	Storage
Tenant	Storage Pool
Bucket	Internal Volume
Disk	Disk

Note: These are common terminology mappings only and might not represent every case for this data collector.

Requirements

- An IP address or hostname of the ECS cluster
- A username and password for the ECS system
- Port 4443 (HTTPS). Requires outbound connectivity to TCP port 4443 on the ECS system.

Configuration

Field	Description
ECS Host	IP address or fully-qualified domain name of the ECS system
ECS Host Port	Port used for communication with ECS Host
ECS User ID	User ID for ECS
Password	Password used for ECS

Advanced configuration

Field	Description
Inventory Poll Interval (min)	The default is 360 minutes.

Troubleshooting

Some things to try if you encounter problems with this data collector:

Inventory

Problem:	Try this:
Error: User authentication failed.	Make sure your credentials for this device are correct.

Performance

Problem:	Try this:
Error: No enough data collected.	* Check collection timestamp in log file and modify polling interval accordingly * Wait for longer time
Error: Performance polling interval is too big.	Check collection timestamp in log file ${logfile} and modify polling interval accordingly

Additional information may be found from the Support page or in the Data Collector Support Matrix.

Dell EMC PowerScale data collector

Data Infrastructure Insights uses the Dell EMC PowerScale (previously Isilon) SSH data collector to acquire inventory and performance data from PowerScale scale-out NAS storage.

Terminology

Data Infrastructure Insights acquires the following inventory information from this data collector. For each asset type acquired by Data Infrastructure Insights, the most common terminology used for this asset is shown. When viewing or troubleshooting this data collector, keep the following terminology in mind:

Vendor/Model Term	Data Infrastructure Insights Term
Drive	Disk

Vendor/Model Term	Data Infrastructure Insights Term
Cluster	Storage
Node	Storage Node
File System	Internal Volume

Note: These are common terminology mappings only and might not represent every case for this data collector.

Requirements

You need the following information to configure this data collector:

- Administrator permissions to the PowerScale storage
- IP address of the PowerScale cluster
- SSH access to port 22

Configuration

Field	Description
IP address	The IP address or fully-qualified domain name of the PowerScale cluster
User Name	User name for the PowerScale cluster
Password	Password used for the PowerScale cluster

Advanced configuration

Field	Description
Inventory Poll Interval (min)	Interval between inventory polls. The default is 20.
Performance Poll Interval (sec)	Interval between performance polls. The default is 300.
SSH Port	SSH service port. The default is 22.

Troubleshooting

Some things to try if you encounter problems with this data collector:

Inventory

Problem:	Try this:
"Invalid login credentials" with error messages “Commands not enabled for role-based administration require root user access”	* Verify that the user has permissions to run the following commands on the device: > isi version osrelease > isi status -q > isi status -n > isi devices -d %s > isi license * Verify credentials used in the wizard are matching device credentials
"Internal Error" with error messages “Command <Your command> run failed with permission: <Your current permission>. Sudo command run permission issue”	Verify that the user has sudo permissions to run the following command on the device

Additional information may be found from the Support page or in the Data Collector Support Matrix.

Dell EMC Isilon / PowerScale REST data collector

Data Infrastructure Insights uses the Dell EMC Isilon / PowerScale REST data collector to acquire inventory and performance data from Dell EMC Isilon or PowerScale storage. This collector supports arrays running OneFS 8.0.0+.

Terminology

Data Infrastructure Insights acquires the following inventory information from this data collector. For each asset type acquired by Data Infrastructure Insights, the most common terminology used for this asset is shown. When viewing or troubleshooting this data collector, keep the following terminology in mind:

Vendor/Model Term	Data Infrastructure Insights Term
Drive	Disk
Cluster	Storage
Node	Storage Node
OneFS File System	Internal Volume
OneFS File System	Storage Pool
Qtree	Qtree

Note: These are common terminology mappings only and might not represent every case for this data collector.

Requirements

You need the following information to configure this data collector:

- A user account and password. This account does NOT need to be admin/root, but you MUST grant a substantial number of read only privileges to your service account - see table below
- IP address / Fully Qualified Domain Name of the Dell EMC Isilon / PowerScale cluster
- HTTPS access to port 8080

- Isilon / PowerScale cluster running OneFS 8.0.0 or higher

Privilege Name	Description	r(read) or rw (read+write)
ISI_PRIV_LOGIN_PAPI	Platform API	r
ISI_PRIV_SYS_TIME	Time	r
ISI_PRIV_AUTH	Auth	r
ISI_PRIV_ROLE	Privilege	r
ISI_PRIV_DEVICES	Devices	r
ISI_PRIV_EVENT	Event	r
ISI_PRIV_HDFS	HDFS	r
ISI_PRIV_NDMP	NDMP	r
ISI_PRIV_NETWORK	Network	r
ISI_PRIV_NFS	NFS	r
ISI_PRIV_PAPI_CONFIG	Configure Platform API	r
ISI_PRIV_QUOTA	Quota	r
ISI_PRIV_SMARTPOOLS	SmartPools	r
ISI_PRIV_SMB	SMB	r
ISI_PRIV_STATISTICS	Statistics	r
ISI_PRIV_SWIFT	Swift	r
ISI_PRIV_JOB_ENGINE	Job Engine	r

Configuration

Field	Description
Isilon IP address	The IP address or fully-qualified domain name of the Isilon storage
User Name	User name for the Isilon
Password	Password used for the Isilon

Advanced configuration

Field	Description
HTTPS Port	The default is 8080.
Inventory Poll Interval (min)	Interval between inventory polls. The default is 20.
Performance Poll Interval (sec)	Interval between performance polls. The default is 300.

Troubleshooting

Some things to try if you encounter problems with this data collector:

Inventory

Problem:	Try this:
"Invalid login credentials" with error messages “Commands not enabled for role-based administration require root user access”	* Verify that the user has permissions to run the following commands on the device: > isi version osrelease > isi status -q > isi status -n > isi devices -d %s > isi license * Verify credentials used in the wizard are matching device credentials
"Internal Error" with error messages “Command <Your command> run failed with permission: <Your current permission>. Sudo command run permission issue”	Verify that the user has sudo permissions to run the following command on the device

Additional information may be found from the Support page or in the Data Collector Support Matrix.

Dell EMC PowerStore data collector

The EMC PowerStore data collector gathers inventory information from EMC PowerStore storage. For configuration, the data collector requires the IP address of the storage processors and a read-only user name and password.

The EMC PowerStore data collector gathers the volume-to-volume replication relationships that PowerStore coordinates across other storage arrays. Data Infrastructure Insights shows a storage array for each PowerStore cluster, and collects inventory data for nodes and storage ports on that cluster. No storage pool or volume data is collected.

Terminology

Data Infrastructure Insights acquires the following inventory information from this data collector. For each asset type acquired by Data Infrastructure Insights, the most common terminology used for this asset is shown. When viewing or troubleshooting this data collector, keep the following terminology in mind:

Vendor/Model Term	Data Infrastructure Insights Term
host	host
host_volume_mapping	host_volume_mapping
hardware (it has Drives under "extra_details" object): Drives	Disk
Appliance	StoragePool
Cluster	Storage Array
Node	StorageNode
fc_port	Port

Vendor/Model Term	Data Infrastructure Insights Term
volume	Volume
InternalVolume	file_system

Note: These are common terminology mappings only and might not represent every case for this data collector.

Requirements

The following information is required to configure this data collector:

- IP address or fully-qualified domain name of storage processor
- Read-only user name and password

Configuration

Field	Description
PowerStore gateway(s)	IP addresses or fully-qualified domain names of PowerStore storage
User Name	User name for PowerStore
Password	Password used for PowerStore

Advanced configuration

Field	Description
HTTPS Port	Default is 443
Inventory Poll Interval (minutes)	Interval between inventory polls. The default is 60 minutes.

Cloud Insight's PowerStore performance collection makes use of PowerStore's 5-minute granularity source data. As such, Data Infrastructure Insights polls for that data every five minutes, and this is not configurable.

Troubleshooting

Additional information on this Data Collector may be found from the Support page or in the Data Collector Support Matrix.

Dell EMC RecoverPoint data collector

The EMC RecoverPoint data collector's primary use case is to discover volume-to-volume replication relationships that the RecoverPoint storage appliance facilitates. This collector will also discover the Recoverpoint appliance itself. Please note that Dell/EMC sells a VMware backup solution for VMs--"RecoverPoint for VMs"--which is not supported by this collector

For configuration, the data collector requires the IP address of the storage processors and a read-only user name and password.

The EMC RecoverPoint data collector gathers the volume-to-volume replication relationships that RecoverPoint coordinates across other storage arrays. Data Infrastructure Insights shows a storage array for each RecoverPoint cluster, and collects inventory data for nodes and storage ports on that cluster. No storage pool or volume data is collected.

Requirements

The following information is required to configure this data collector:

- IP address or fully-qualified domain name of storage processor
- Read-only user name and password
- REST API access via port 443

Configuration

Field	Description
Address of RecoverPoint	IP address or fully-qualified domain name of RecoverPoint cluster
User Name	User name for the RecoverPoint cluster
Password	Password used for the RecoverPoint cluster

Advanced configuration

Field	Description
TCP Port	TCP Port used to connect to Recoverpoint cluster
Inventory Poll Interval (minutes)	Interval between inventory polls. The default is 20 minutes.
Excluded Clusters	Comma-separated list of cluster IDs or names to exclude when polling.

Troubleshooting

Additional information on this Data Collector may be found from the Support page or in the Data Collector Support Matrix.

DELL EMC ScaleIO / PowerFlex data collector

The ScaleIO/PowerFlex data collector collects inventory information from ScaleIO & PowerFlex storage. For configuration, this data collector requires the ScaleIO/PowerFlex gateway address and an admin user name and password.

Terminology

Data Infrastructure Insights acquires the following inventory information from the ScaleIO/PowerFlex data collector. For each asset type acquired by Data Infrastructure Insights, the most common terminology used for this asset is shown. When viewing or troubleshooting this data collector, keep the following terminology in mind:

Vendor/Model Term	Data Infrastructure Insights Term
MDM (Meta Data Manager) Cluster	Storage
SDS (ScaleIO/PowerFlex Data Server)	Storage Node
Storage Pool	Storage Pool
Volume	Volume
Device	Disk

Note: These are common terminology mappings only and might not represent every case for this data collector.

Requirements

- Read-only access to the Admin user account
- Port requirement: HTTPS Port 443

Configuration

Field	Description
ScaleIO/PowerFlex Gateway(s)	IP addresses or FQDNs of ScaleIO/PowerFlex gateways, separated by comma (,) or semicolon (;)
User Name	Admin user name used to log in to the ScaleIO/PowerFlex device
Password	Password used to log in to the ScaleIO/PowerFlex device

Advanced configuration

Click the Inventory check box to enable inventory collection.

Field	Description
HTTPS port	443
Inventory poll interval (min)	The default is 60.
Connection Timeout (sec)	The default is 60.

Troubleshooting

Additional information may be found from the Support page or in the Data Collector Support Matrix.

Configuring the EMC Unity data collector

The DELL EMC Unity (formerly VNXe) data collector provides inventory support for VNXe unified storage arrays. Data Infrastructure Insights currently supports iSCSI and NAS protocols.

Requirements

- The Unity data collector is CLI based; you must install the Unisphere for Unity CLI, (uemcli.exe) onto the acquisition unit where your VNXe data collector resides.
- uemcli.exe uses HTTPS as the transport protocol, so the acquisition unit will need to be able to initiate HTTPS connections to the Unity.
- IP address or fully-qualified domain name of the Unity device
- You must have at least a read-only user for use by the data collector.
- HTTPS on Port 443 is required
- The EMC Unity data collector provides NAS and iSCSI support for inventory; fibre channel volumes will be discovered, but Data Infrastructure Insights does not report on FC mapping, masking, or storage ports.

Terminology

Data Infrastructure Insights acquires the following inventory information from the Unity data collector. For each asset type acquired by Data Infrastructure Insights, the most common terminology used for this asset is shown. When viewing or troubleshooting this data collector, keep the following terminology in mind:

Vendor/Model Term	Data Infrastructure Insights Term
Disk	Disk
Storage Array	Storage
Processor	Storage Node
Storage Pool	Storage Pool
General iSCSI Block info, VMware VMFS	Share
Replication Remote System	Synchronization
iSCSI Node	iSCSI Target Node
iSCSI Initiator	iSCSI Target Initiator

Note: These are common terminology mappings only and might not represent every case for this data source.

Configuration

Field	Description
Unity Storage	IP address or fully-qualified domain name of the Unity device
User Name	User name for the Unity device
Password	Password for the Unity device
Full Path to the Executable UEMCLI	Full path to the folder containing the *uemcli.exe* executable

Advanced configuration

Field	Description
Inventory Poll Interval (min)	Interval between inventory polls. The default is 40 minutes
Unity CLI Port	Port used for the Unity CLI
Performance poll interval (sec)	The default is 300.

Troubleshooting

Some things to try if you encounter problems with this data collector:

Inventory

Problem:	Try this:
"Failed to execute external utility" with error messages “Failed to find Unisphere executable uemcli”	* Verify correct IP address, username, and password * Confirm that Unisphere CLI is installed on the Data Infrastructure Insights Acquisition Unit * Confirm that Unisphere CLI installation directory is correct in the datasource configuration * Confirm that the IP of the VNXe is correct in the configuration of the datasource. From the Data Infrastructure Insights Acquisition Unit, open a CMD and change to to the configured installation directory: ${INSTALLDIR. Try to make a connection with the VNXe device by typing: uemcli -d <Your IP> -u <Your ID> /sys/general show

Additional information may be found from the Support page or in the Data Collector Support Matrix.

Dell EMC VMAX and PowerMax Family of Devices data collector

Data Infrastructure Insights discovers EMC VMAX and PowerMax storage arrays by using Solutions Enabler symcli commands in conjunction with an existing Solutions Enabler server in your environment. The existing Solutions Enabler server has connectivity to the VMAX/PowerMax storage array through access to gatekeeper volumes.

Requirements

Before configuring this data collector, you should ensure that Data Infrastructure Insights has TCP connectivity to port 2707 on the existing Solutions Enabler server. Data Infrastructure Insights discovers all the Symmetrix arrays that are "Local" to this server, as seen in "symcfg list" output from that server.

- The EMC Solutions Enabler (CLI) with SMI-S provider application must be installed on the Acquisition Unit server and the version must match or be earlier than the version running on the Solutions Enabler Server.
- A properly configured {installdir}\EMC\SYMAPI\config\netcnfg file is required. This file defines service names for Solutions Enabler servers, as well as the access method (SECURE / NOSECURE /ANY).
- If you require read/write latency at the storage node level, the SMI-S Provider must communicate with a running instance of the UNISPHERE for VMAX application.

- IP address of the managing Solutions Enabler server
- Administrator permissions on the Solutions Enabler (SE) Server
- Read-only user name and password to the SE software
- The UNISPHERE for VMAX application must be running and collecting statistics for the EMC VMAX and PowerMax sstorage arrays that are managed by the SMI-S Provider installation
- Access validation for performance: In a web browser on your Acquisition Unit, go to *https://<SMI-S Hostname or IP>:5989/ecomconfig* where "SMI-S Hostname or IP" is the IP address or hostname of your SMI-S server. This URL is for an administrative portal for the EMC SMI-S (aka "ECOM") service - you will receive a login popup.
- Permissions must be declared in the Solutions Enabler server's daemon configuration file, usually found here: */var/symapi/config/daemon_users*

  Here is an example file with the proper cisys permissions.

```
root@cernciaukc101:/root
14:11:25 # tail /var/symapi/config/daemon_users
###
###     Refer to the storrdfd(3) man page for additional details.
###
###     As noted above, only authorized users can perform stordaemon
control
###     operations (e.g., shutdown).
#############################################################################
########
# smith         storrdfd
cisys storapid <all>
```

Terminology

Data Infrastructure Insights acquires the following inventory information from the EMC VMAX/PowerMax data source. For each asset type acquired, the most common terminology used for this asset is shown. When viewing or troubleshooting this data collector, keep the following terminology in mind:

Vendor/Model Term	Data Infrastructure Insights Term
Disk	Disk
Disk Group	Disk Group
Storage	Array Storage
Director	Storage Node
Device Pool, Storage Resource Pool (SRP)	Storage Pool
Device TDev	Volume

Note: These are common terminology mappings only and might not represent every case for this data collector.

Configuration

Note: If SMI-S user authentication is not enabled, the default values in the Data Infrastructure Insights data collector are ignored.

Field	Description
Service Name	Service name as specified in *netcnfg* file
Full path to CLI	Full path to the folder containing the Symmetrix CLI
SMI-S Host IP Address	IP address of the SMI-S host

Advanced Configuration

Field	Description
Inventory Poll Interval (min)	Interval between inventory polls. The default is 40 minutes.
Choose 'Exclude' or 'Include' to specify a list	Specify whether to include or exclude the array list below when collecting data.
Inventory Filter Device List	Comma-separated list of device IDs to include or exclude
Connection Caching	Choose connection caching method: * LOCAL means that the Cloud Insights Acquisition service is running on the Solutions Enabler server, which has Fibre Channel connectivity to the Symmetrix arrays you seek to discover, and has access to gatekeeper volumes. This might be seen in some Remote Acquisition Unit (RAU) configurations. * REMOTE_CACHED is the default and should be used in most cases. This uses the NETCNFG file settings to connect using IP to the Solutions Enabler server, which must have Fibre Channel connectivity to the Symmetrix arrays you seek to discover, and has access to Gatekeeper volumes. * In the event that REMOTE_CACHED options make CLI commands fail, use the REMOTE option. Keep in mind that it will slow down the acquisition process (possibly to hours or even days in extreme cases). The NETCNFG file settings are still used for an IP connection to the Solutions Enabler server that has Fibre Channel connectivity to the Symmetrix arrays being discovered. **Note:** This setting does not change Data Infrastructure Insights behavior with respect to the arrays listed as REMOTE by the "symcfg list" output. Data Infrastructure Insights gathers data only on devices shown as LOCAL by this command.
SMI-S Protocol	Protocol used to connect to the SMI-S provider. Also displays the default port used.

Field	Description
Override SMIS-Port	If blank, use the default port in the Connection Type field, otherwise enter the connection port to use
SMI-S User Name	User name for the SMI-S Provider Host
SMI-S Password	User name for the SMI-S Provider Host
Performance Polling Interval (sec)	Interval between performance polls (default 1000 seconds)
hoose 'Exclude' or 'Include' to specify a list	Specify whether to include or exclude the array list below when collecting performance data
Performance Filter Device List	Comma-separated list of device IDs to include or exclude

Troubleshooting

Some things to try if you encounter problems with this data collector:

Problem:	Try this:
Error: The feature being requested is not currently licensed	Install the SYMAPI server license.
Error: No devices were found	Make sure Symmetrix devices are configured to be managed by the the Solutions Enabler server: - Run symcfg list -v to see the list of configured Symmetrix devices.
Error: A requested network service was not found in the service file	Make sure the Solutions Enabler Service Name is defined the netcnfg file for Solutions Enabler. This file is usually located under SYMAPI\config\ in the Solutions Enabler client installation.
Error: The remote client/server handshake failed	Check the most recent storsrvd.log* files on the Solutions Enabler host we are trying to discover.
Error: Common name in client certificate not valid	Edit the *hosts* file on the Solutions Enabler server so that the Acquisition Unit's hostname resolves to the IP address as reported in the storsrvd.log on the Solutions Enabler server.
Error: The function could not obtain memory	Make sure there is enough free memory available in the system to execute Solutions Enabler
Error: Solutions Enabler was unable to serve all data required.	Investigate the health status and load profile of Solutions Enabler
Error: • The "symcfg list -tdev" CLI command may return incorrect data when collected with Solutions Enabler 7.x from a Solutions Enabler server 8.x. • The "symcfg list -srp" CLI command may return incorrect data when collected with Solutions Enabler 8.1.0 or earlier from a Solutions Enabler server 8.3 or later.	Be sure you are using the same Solutions Enabler major release

Problem:	Try this:
I'm seeing data collection errors with the message: "unknown code"	You may see this message if permissions are not declared in the Solutions Enabler server's daemon configuration file (see the Requirements above.) This assumes your SE client version matches your SE server version. This error may also occur if the *cisys* user (which executes Solutions Enabler commands) has not been configured with the necessary daemon permissions in the /var/symapi/config/daemon_users configuration file. To fix this, edit the /var/symapi/config/daemon_users file and make sure the cisys user has <all> permission specified for the storapid daemon. Example: 14:11:25 # tail /var/symapi/config/daemon_users … cisys storapid <all>

Additional information may be found from the Support page or in the Data Collector Support Matrix.

Dell EMC VNX Block Storage (NaviCLI) data collector

Data Infrastructure Insights uses the Dell EMC VNX Block Storage (NaviSec) data collector (formerly CLARiiON) to acquire inventory and performance data.

Terminology

Data Infrastructure Insights acquires the following inventory information from the EMC VNX Block Storage data collector. For each asset type acquired by Data Infrastructure Insights, the most common terminology used for this asset is shown. When viewing or troubleshooting this data collector, keep the following terminology in mind:

Vendor/Model Term	Data Infrastructure Insights Term
Disk	Disk
Storage	Storage
Storage Processor	Storage Node
This Pool, RAID Group	Storage Pool
LUN	Volume

Note: These are common terminology mappings only and might not represent every case for this data source.

Requirements

The following requirements must be met in order to collect data:

- An IP address of each VNX block storage processor
- Read-only Navisphere username and password to the VNX block storage arrays
- NaviSecCli must be installed on the Data Infrastructure Insights AU
- Access validation: Run NaviSecCLI from the Data Infrastructure Insights AU to each array using the username and password.
- Port requirements: 80, 443
- NaviSecCLI version should correspond with the newest FLARE code on your array
- For performance, statistics logging must be enabled.

NaviSphere command line interface syntax

naviseccli.exe -h <IP address> -user <user> -password <password> -scope <scope,use 0 for global scope> -port <use 443 by default> command

Configuration

Field	Description
VNX Block Storage IP Address	IP address or fully-qualified domain name of the VNX block storage
User Name	Name used to log into the VNX block storage device.
Password	Password used to log into the VNX block storage device.
CLI Path to naviseccli.exe	Full path to the folder containing the *naviseccli.exe* executable

Advanced Configuration

Field	Description
Inventory Poll Interval (min)	Interval between inventory polls. Default is 40 minutes.
Scope	The secure client scope. The default is Global.
Performance Poll Interval (sec)	Interval between performance polls. The default is 300 seconds.

Troubleshooting

Some things to try if you encounter problems with this data collector:

Inventory

Problem:	Try this:
Error: • Agent Not Running • Failed to find naviseccli • Failed to execute any command	• Confirm that NaviSphere CLI is installed on the Cloud Insight Acquisition Unit • You have not selected the "Use secure client" option in the data collector configuration wizard and do not have a non-secure version of Naviphere CLI installed. • Confirm that NaviSphere CLI installation directory is correct in the data collector configuration • Confirm that the IP of the VNX block storage is correct in the data collector configuration: • From the Data Infrastructure Insights Acquisition Unit: ◦ Open a CMD. ◦ Change the directory to the configured installation directory ◦ Try to make a connection with the VNX block storage device by typing “navicli -h {ip} getagent” (replace the {ip} with the actual IP)
Error: 4.29 emc235848 emc241018 getall Failed to parse host alias info	This is likely caused by a FLARE 29 corruption issue of the host initiator database on the array itself. See EMC knowledge base articles: emc235848, emc241018. You can also check https://now.netapp.com/Knowledgebase/solutionarea.asp?id=kb58128
Error: Unable to retrieve Meta LUNs. Error Executing java -jar navicli.jar	• Modify the data collector configuration to use the secure client (recommended) • Install navicli.jar in the CLI path to navicli.exe OR naviseccli.exe • Note: navicli.jar is deprecated as of EMC Navisphere version 6.26 • The navicli.jar may be available on http://powerlink.emc.com
Error: Storage Pools not reporting disks on Service Processor at configured IP address	Configure the data collector with both Service Processor IPs, separated by a comma

Problem:	Try this:
Error: Revision mismatch error	• This is usually caused by updating the firmware on the VNX block storage device, but not updating the installation of NaviCLI.exe. This also might be caused by having different devices with different firmwares, but only one CLI installed (with a different firmware version). • Verify that the device and the host are both running identical versions of the software: ◦ From the Data Infrastructure Insights Acquisition Unit, open a command line window ◦ Change the directory to the configured installation directory ◦ Make a connection with the CLARiiON device by typing “navicli -h <ip> getagent” ◦ Look for the version number on the first couple of lines. Example: “Agent Rev: 6.16.2 (0.1)” ◦ Look for and compare the version on the first line. Example: “Navisphere CLI Revision 6.07.00.04.07”
Error: Unsupported Configuration - No Fibre Channel Ports	The device is not configured with any Fibre Channel ports. Currently, only FC configurations are supported. Verify this version/firmware is supported.

Additional information may be found from the Support page or in the Data Collector Support Matrix.

DELL EMC VNX File (formerly Celerra Unified Storage System) data collector

This data collector acquires inventory information from the VNX File Storage System. For configuration, this data collector requires the IP address of the storage processors and a read-only user name and password.

Terminology

Data Infrastructure Insights acquires the following inventory information from the VNX File data collector. For each asset type acquired by Data Infrastructure Insights, the most common terminology used for this asset is shown. When viewing or troubleshooting this data collector, keep the following terminology in mind:

Vendor/Model Term	Data Infrastructure Insights Term
Celerra Network Server/Celerra Storage Pool	Storage Pool
File System	Internal Volume
Data Mover	Controller
File System mounted on a data mover	File Share

Vendor/Model Term	Data Infrastructure Insights Term
CIFS and NFS Exports	Share
Disk Volume	Backend LUN

Note: These are common terminology mappings only and might not represent every case for this data collector.

Requirements

You need the following to configure this data collector:

- The IP address of the storage processor
- Read-only user name and password
- SSH port 22

Configuration

Field	Description
VNX File IP Address	IP address or fully-qualified domain name of the VNX File device
User Name	Name used to log in to the VNX File device
Password	Password used to log in to the VNX File device

Advanced configuration

Field	Description
Inventory Poll Interval (minutes)	Interval between inventory polls. The default is 20 minutes.

Troubleshooting

Some things to try if you encounter problems with this data collector:

Inventory

Problem:	Try this:
Error: Unable to proceed while DART update in progress	Possible solution: Pause the data collector and wait for the DART upgrade to complete before attempting another acquisition request.

Additional information may be found from the Support page or in the Data Collector Support Matrix.

Configuring the Dell EMC VNX Unified data collector

For configuration, the Dell EMC VNX Unified (SSH) data collector requires the IP address of the Control Station and a read-only username and password.

Terminology

Data Infrastructure Insights acquires the following inventory information from this data collector. For each asset type acquired by Data Infrastructure Insights, the most common terminology used for this asset is shown. When viewing or troubleshooting this data collector, keep the following terminology in mind:

Vendor/Model Term	Data Infrastructure Insights Term
Disk	Disk
Disk Folder	Disk Group
File system	Internal Volume
Storage	Storage
Storage Processor	Storage Node
Storage Pool, RAID Group	Storage Pool
LUN	Volume
Data Mover	Controller
File System mounted on a data mover	File Share
CIFS and NFS Exports	Share
Disk Volume	Backend LUN

Requirements

You need the following to configure the VNX (SSH) data collector:

- VNX IP address & Credentials to the Celerra Control Station.
- Read-only username and password.
- The data collector is able to run NaviCLI/NaviSecCLI commands against the backend array utilizing the DART OS NAS heads

Configuration

Field	Description
VNX IP Address	IP address or fully-qualified domain name of the VNX Control Station
User Name	User name for the VNX Control Station
Password	Password for the VNX Control Station

Advanced configiration

Field	Description
Inventory Poll Interval (min)	Interval between inventory polls. The default is 40 minutes.
Performance Poll Interval (sec).	Interval between performance polls. The default is 300 seconds.

Troubleshooting

Additional information on this Data Collector may be found from the Support page or in the Data Collector Support Matrix.

Configuring the EMC VPLEX data collector

This data collector acquires inventory and performance data from EMC VPLEX storage systems. For configuration, the data collector requires an IP address of the VPLEX server and an administrative level domain account.

> Data Infrastructure Insights' performance collection from Vplex clusters requires that the performance archive service be operational, in order to populate the .CSV files and logs that Data Infrastructure Insights retrieves via SCP-based file copies. NetApp has observed that many Vplex firmware upgrade/management station updates will leave this functionality non-operational. Customers planning such upgrades may want to proactively ask Dell/EMC if their planned upgrade will leave this functionality inoperable, and if so, how can they re-enable it to minimize gaps in performance visibility? Cloud Insight's Vplex performance code will assess on each poll whether all the expected files exist, and if they are being properly updated; if they are missing or stale, Data Infrastructure Insights will log performance collection failures.

Terminology

Data Infrastructure Insightst acquires the following inventory information from the VPLEX data collector. For each asset type acquired, the most common terminology used for this asset is shown. When viewing or troubleshooting this data collector, keep the following terminology in mind:

Vendor/Model Term	Data Infrastructure Insights Term
Cluster	Storage
Engine	Storage Node
Device, System Extent	Backend Storage Pool
Virtual Volume	Volume
Front-End Port, Back-End Port	Port
Distributed Device	Storage Synchronization
Storage View	Volume Map, Volume Mask
Storage Volume	Backend LUN
ITLs	Backend Path

Note: These are common terminology mappings only and might not represent every case for this data collector.

Requirements

- An IP address of the VPLEX Management Console
- Administrative level domain account for the VPLEX server
- Port 443 (HTTPS). Requires outbound connectivity to TCP port 443 on the VPLEX management station.

- For performance, read-only username and password for ssh/scp access.
- For performance, port 22 is required.

Configuration

Field	Description
IP address of VPLEX Management Console	IP address or fully-qualified domain name of the VPLEX Management Console
User Name	User name for VPLEX CLI
Password	Password used for VPLEX CLI
Performance Remote IP Address	Performance Remote IP address of the VPLEX Management Console
Performance Remote User Name	Performance Remote user name of VPLEX Management Console
Performance Remote Password	Performance Remote Password of VPLEX Management Console

Advanced configuration

Field	Description
Communication Port	Port used for VPLEX CLI. The default is 443.
Inventory Poll Interval (min)	The default is 20 minutes.
Number of connection retries	The default is 3.
Performance Poll Interval (sec)	Interval between performance polls. The default is 600 seconds.
Number of Retries	The default is 2.

Troubleshooting

Some things to try if you encounter problems with this data collector:

Inventory

Problem:	Try this:
Error: User authentication failed.	Make sure your credentials for this device are correct.

Performance

Problem:	Try this:
Error: VPLEX performance for version below 5.3 is not supported.	Upgrade VPLEX to 5.3 or above
Error: No enough data collected.	• Check collection timestamp in log file and modify polling interval accordingly • Wait for longer time

Problem:	Try this:
Error: Perpetual Log files not being updated.	Please contact EMC support to enable updating the perpetual log files
Error: Performance polling interval is too big.	Check collection timestamp in log file ${logfile} and modify polling interval accordingly
Error: Performance Remote IP address of VPLEX Management Console is not configured.	Edit the data source to set Performance Remote IP address of VPLEX Management Console.
Error: No performance data reported from director	• Check that the system performance monitors are running correctly • Please contact EMC support to enable updating the system performance monitor log files

Additional information may be found from the Support page or in the Data Collector Support Matrix.

Dell EMC XtremeIO data collector

The EMC XtremIO data collector acquires inventory and performance data from the EMC XtremIO storage system.

Requirements

To configure the EMC XtremIO (HTTP) data collector, you must have:

- The XtremIO Management Server (XMS) Host address
- An account with administrator privileges
- Access to port 443 (HTTPS)

Terminology

Data Infrastructure Insights acquires the following inventory information from the EMC XtremIO data collector. For each asset type acquired by Data Infrastructure Insights, the most common terminology used for this asset is shown. When viewing or troubleshooting this data source, keep the following terminology in mind:

Vendor/Model Term	Data Infrastructure Insights Term
Disk (SSD)	Disk
Cluster	Storage
Controller	Storage Node
Volume	Volume
LUN Map	Volume Map
Target FC Initiator	Volume Mask

Note: These are common terminology mappings only and might not represent every case for this data source.

Requirements

- The XtremIO Management Server (XMS) Host IP address

- Administrator user name and password for the XtremIO

Configuration

Field	Description
XMS Host	IP address or fully-qualified domain name of the XtremIO Management Server
User name	User name for the XtremIO Management Server
Password	Password for the XtremIO Management Server

Advanced configuration

Field	Description
TCP port	TCP Port used to connect to XTremIO Management Server. The default is 443.
Inventory poll interval (min)	Interval between inventory polls. The default is 60 minutes.
Performance poll interval (sec)	Interval between performance polls. The default is 300 seconds.

Troubleshooting

Additional information on this Data Collector may be found from the Support page or in the Data Collector Support Matrix.

Fujitsu Eternus data collector

The Fujitsu Eternus data collector acquires inventory data using administration-level access to the storage system.

Terminology

Data Infrastructure Insights acquires the following inventory information from the Fujitsu Eternus storage. For each asset type acquired by Data Infrastructure Insights, the most common terminology used for this asset is shown. When viewing or troubleshooting this data collector, keep the following terminology in mind:

Vendor/Model Term	Data Infrastructure Insights Term
Disk	Disk
Storage	Storage
Thin Pool, Flexible Tier Pool, Raid Group	Storage Pool
Standard Volume, Snap Data Volume (SDV), Snap Data Pool Volume (SDPV), Thin Provisioning Volume (TPV), Flexible Tier Volume (FTV), Wide Striping Volume (WSV)	Volume
Channel adapter	Controller

Note: These are common terminology mappings only and might not represent every case for this data collectior.

Requirements

The following are required to configure this data collector:

- An IP address of the Eternus storage, which cannot be comma delimited
- SSH Administration-level user name and password
- Port 22
- Ensure that the page scroll is disabled (clienv-show-more-scroll disable)

Configuration

Field	Description
IP Address of Eternus Storage	IP address of the Eternus storage
User Name	User name for Eternus storage
Password	Password for the Eternus storage

Advanced configuration

Field	Description
Inventory Poll Interval (min)	The default is 20 minutes.

Troubleshooting

Some things to try if you encounter problems with this data collector:

Inventory

Problem:	Try this:
"Error retrieving data" with error messages “Error Finding Prompt CLI” or “Error finding prompt at the end of shell results”	Likely caused by: Storage system has page scrolling enabled. Possible solution: * Try to disable page scrolling by running the following command: set clienv-show-more -scroll disable
"Connecting error" with error messages “Failed to instantiate an SSH connection to storage” or “Failed to instantiate a connection to VirtualCenter”	Likely causes: * Incorrect credentials. * Incorrect IP address. * Network problem. * Storage may be down or unresponsive. Possible solutions: * Verify credentials and IP address entered. * Try to communicate with storage using SSH Client.

Additional information may be found from the Support page or in the Data Collector Support Matrix.

NetApp Google Compute data collector

This data collector supports inventory and performance collection from Google Compute cloud platform configurations. This collector will seek to discover all the Compute resources within all the Projects within one Google organization. If you have multiple Google organizations you want to discover with Data Infrastructure Insights, you will want to deploy one Data Infrastructure Insights collector per organization.

Service Account Requirements

- You must create a service account by following the instructions in Creating/Managing Service Accounts. Such service account is identified by a unique ID, known as its *clientId*, which will be used as a username.
- Additionally, create a service account key by following the instructions in Creating/Managing Service Account Keys. This key can be downloaded as a json file, whose content will be used as a password.
- The service account must be scoped for *compute.readonly*, *monitoring.read*, and *cloud-platform*.

Configuration

Field	Description
Organization ID	The organization ID you want to discover with this collector. This field is required if your service account is able to see more than one organization
Choose 'Exclude' or 'Include' to filter GCP Projects by IDs	If you want to limit what projects' resources are brought into Data Infrastructure Insights.
Project IDs	The list of Project IDs that you want to filter in, or out from discovery, depending on the value of the "Choose 'Exclude'…." value. The default list is empty
Client ID	Client ID for the Google Cloud Platform configuration
Copy and paste the contents of your Google Credential File here	Copy your Google credentials for the Cloud Platform account to this field

Advanced configuration

Field	Description
Inventory Poll Interval (min)	Default is 60 minutes
Choose 'Exclude' or 'Include' to Apply to Filter VMs by Labels	Specify whether to include or exclude VM's by Labels when collecting data. If 'Include' is selected, the Label Key field can not be empty.
Label Keys and Values on which to Filter VMs	Click **+ Filter Label** to choose which VMs (and associated disks) to include/exclude by filtering for keys and values that match keys and values of labels on the VM. Label Key is required, Label Value is optional. When Label Value is empty, the VM is filtered as long as it matches the Label Key.
Performance Poll Interval (sec)	Default is 1800 seconds

Troubleshooting

Additional information on this Data Collector may be found from the Support page or in the Data Collector Support Matrix.

Google Cloud NetApp Volumes data collector

This data collector supports inventory and performance collection from Google Cloud NetApp Volumes configurations. The collector discovers NetApp volumes and storage resources across all projects within a Google organization. If you have multiple Google organizations that you want to monitor with Data Infrastructure Insights, deploy one collector per organization.

Service Account Requirements

- You must create a service account by following the instructions in Creating/Managing Service Accounts. This service account is identified by a unique ID, known as its *clientId*, which will be used as a username.
- Additionally, create a service account key by following the instructions in Creating/Managing Service Account Keys. This key can be downloaded as a json file, whose content will be used as a password.
- The service account must be scoped for *compute.readonly*, *monitoring.read*, and *cloud-platform*.

Configuration

Field	Description
Organization ID	The organization ID you want to discover with this collector. This field is required if your service account is able to see more than one organization
Choose 'Exclude' or 'Include' to filter GCNV assets by Location	This defaults to Exclude, as this collector by default intends to discover all GCNV volumes worldwide within your organization.
GCNV Exclude/Include Locations	This defaults to being empty, and is used in conjunction with the "Choose 'Exclude' or 'Include' option. If you want to discover assets only within certain regions, use these two options to limit the scope of this collector.
Project IDs	The list of Project IDs that you want to filter in, or out from discovery, depending on the value of the "Choose 'Exclude'…." value. The default list is empty
Client ID	Client ID for the Google Cloud Platform configuration
Copy and paste the contents of your Google Credential File here	Copy your Google credentials for the Cloud Platform account to this field

Advanced configuration

Field	Description
Inventory Poll Interval (min)	Default is 60 minutes

Field	Description
Use AU Proxy for REST API calls	Select this option to have the collector use the same proxy as the Acquisition Unit on which the collector resides. By default, this is disabled, which means the collector attempts to send HTTPS API calls directly to Google.
Label Keys and Values on which to Filter VMs	Click **+ Filter Label** to choose which VMs (and associated disks) to include/exclude by filtering for keys and values that match keys and values of labels on the VM. Label Key is required, Label Value is optional. When Label Value is empty, the VM is filtered as long as it matches the Label Key.
Performance Poll Interval (sec)	Default is 300 seconds

Troubleshooting

Additional information on this Data Collector may be found from the Support page or in the Data Collector Support Matrix.

HP Enterprise

HP Enterprise Alletra 9000 / Primera Storage data collector

Data Infrastructure Insights uses the HP Enterprise Alletra 9000 / HP Enterprise Primera (previously 3PAR) data collector to discover inventory and performance.

Terminology

Data Infrastructure Insights acquires the following inventory information from this data collector. For each asset type acquired by Data Infrastructure Insights, the most common terminology used for this asset is shown. When viewing or troubleshooting this data collector, keep the following terminology in mind:

Field	Description
Physical Disk	Disk
Storage System	Storage
Controller Node	Storage Node
Common Provisioning Group	Storage Pool
Virtual Volume	Volume

Note: These are common terminology mappings only and might not represent every case for this data collector.

Requirements

The following are required to configure this data collector:

- IP address or FQDN of the InServ cluster

- For inventory, read-only user name and password to the StoreServ Server
- For performance, read-write user name and password to the StoreServ Server
- Port requirements: 22 (inventory collection), 5988 or 5989 (performance collection) [Note: Performance is supported for StoreServ OS 3.x+]
- For performance collection confirm that SMI-S is enabled by logging into the array via SSH.

Configuration

Field	Description
Storage IP address	Storage IP address or fully-qualified domain name of the StoreServ cluster
User Name	User name for the StoreServ Server
Password	Password used for the StoreServ Server
SMI-S User Name	User name for the SMI-S Provider Host
SMI-S Password	Password used for the SMI-S Provider Host

Advanced configuration

Field	Description
Inventory Poll Interval (min)	Interval between inventory polls. The default is 40 minutes.
SMI-S Connectivity	Protocol used to connect to the SMI-S provider
Override SMI-S Default Port	If blank, use the default port from SMI-S Connectivity, otherwise enter the connection port to use
Performance Poll Interval (sec)	Interval between performance polls. The default is 300 seconds.

Troubleshooting

Some things to try if you encounter problems with this data collector:

Inventory

Problem:	Try this:
"showsys" command doesn’t return any result.	Run "showsys" and "showversion -a" from the command line and check if the version is supported by the array.

Performance

Problem:	Try this:
Failed to connect or login. Provider initialization failed.	An all-numeric array name can cause problems with SMI-S server. Try changing the array name.

Problem:	Try this:
SMI-S user configured does not have any domain	Grant appropriate domain privileges to the configured SMI-S user
Data Infrastructure Insights states that it cannot connect/login to SMI-S service.	Confirm there is no firewall between the CI AU and the array that would block the CI AU from making TCP connections to 5988 or 5989. Once that is done, and if you have confirmed there is no firewall, you should SSH to the array, and use the “showcim” command to confirm. Verify that: * Service is enabled * HTTPS is enabled * HTTPS port should be 5989 If those all are so, you can try to “stopcim” and then a “startcim” to restart the CIM (i.e. SMI-S service).

Additional information may be found from the Support page or in the Data Collector Support Matrix.

HP Enterprise Command View data collector

The HP Enterprise Command View Advanced Edition data collector supports discovering XP and P9500 arrays via Command View Advanced Edition (CVAE) server. Data Infrastructure Insights communicates with CVAE using the standard Command View API to collect inventory and performance data.

Terminology

Data Infrastructure Insights acquires the following inventory information from the HP Enterprise Command View data collector. For each asset type acquired by Data Infrastructure Insights, the most common terminology used for this asset is shown. When viewing or troubleshooting this data collector, keep the following terminology in mind:

Vendor/Model Term	Data Infrastructure Insights Term
PDEV	Disk
Journal Pool	Disk Group
Storage Array	Storage
Port Controller	Storage Node
Array Group, DP Pool	Storage Pool
Logical Unit, LDEV	Volume

Note: These are common terminology mappings only and might not represent every case for this data collector.

Inventory requirements

You must have the following in order to collect inventory data:

- IP address of the CVAE server
- Read-only user name and password for the CVAE software and peer privileges
- Port requirement: 2001

Performance requirements

The following requirements must be met in order to collect performance data:

- HDS USP, USP V, and VSP performance
  - Performance Monitor must be licensed.
  - Monitoring switch must be enabled.
  - The Export Tool (Export.exe) must be copied to the Data Infrastructure Insights AU and extracted to a location. On CI Linux AUs, ensure "cisys" has read and execute permissions.
  - The Export Tool version must match the microcode version of the target array.
- AMS performance:
  - Performance Monitor must be licensed.
  - The Storage Navigator Modular 2 (SNM2) CLI utility be installed on the Data Infrastructure Insights AU.
- Network requirements
  - The Export Tools are Java based, and use RMI to speak to the array. These tools may not be firewall-friendly as they may dynamically negotiate source and destination TCP ports on each invocation. Also, different model array's Export Tools may behave differently across the network - consult HPE for your model's requirements

Configuration

Field	Description
Command View Server	IP address or fully-qualified domain name of the Command View server
User Name	User name for the Command View server.
Password	Password used for the Command View server.
Devices - VSP G1000 (R800), VSP (R700), HUS VM (HM700) and USP storages	Device list for VSP G1000 (R800), VSP (R700), HUS VM (HM700) and USP storages. Each storage requires: * Array's IP: IP address of the storage * User Name: User name for the storage * Password: Password for the storage * Folder Containing Export Utility JAR Files

Field	Description
SNM2Devices - WMS/SMS/AMS Storages	Device list for WMS/SMS/AMS storages. Each storage requires: * Array's IP: IP address of the storage * Storage Navigator CLI Path: SNM2 CLI path * Account Authentication Valid: Select to choose valid account authentication * User Name: User name for the storage * Password: Password for the storage
Choose Tuning Manager for Performance	Override other performance options
Tuning Manager Host	IP address or fully-qualified domain name of tuning manager
Tuning Manager Port	Port used for Tuning Manager
Tuning Manager Username	User name for Tuning Manager
Tuning Manager Password	Password for Tuning Manager

Note: In HDS USP, USP V, and VSP, any disk can belong to more than one array group.

Advanced configuration

Field	Description
Command View Server Port	Port used for the Command View Server
HTTPs Enabled	Select to enable HTTPs
Inventory Poll Interval (min)	Interval between inventory polls. The default is 40.
Choose 'Exclude' or 'Include' to specify a list	Specify whether to include or exclude the array list below when collecting data.
Exclude or Include Devices	Comma-separated list of device ID's or array names to include or exclude
Query Host Manager	Select to query host manager
Performance Polling Interval (sec)	Interval between performance polls. The default is 300.

Troubleshooting

Some things to try if you encounter problems with this data collector:

Inventory

Problem:	Try this:
Error: User does not have enough permission	Use a different user account that has more privilege or increase the privilege of user account configured in the data collector

Problem:	Try this:
Error: Storages list is empty. Either devices are not configured or the user does not have enough permission	* Use DeviceManager to check if the devices are configured. * Use a different user account that has more privilege, or increase the privilege of the user account
Error: HDS storage array was not refreshed for some days	Investigate why this array is not being refreshed in HP CommandView AE.

Performance

Problem:	Try this:
Error: * Error executing export utility * Error executing external command	* Confirm that Export Utility is installed on the Data Infrastructure Insights Acquisition Unit * Confirm that Export Utility location is correct in the data collector configuration * Confirm that the IP of the USP/R600 array is correct in the configuration of the data collector * Confirm that the User name and password are correct in the configuration of the data collector * Confirm that Export Utility version is compatible with storage array micro code version * From the Data Infrastructure Insights Acquisition Unit, open a CMD prompt and do the following: - Change the directory to the configured installation directory - Try to make a connection with the configured storage array by executing batch file runWin.bat
Error: Export tool login failed for target IP	* Confirm that username/password is correct * Create a user ID mainly for this HDS data collector * Confirm that no other data collectors are configured to acquire this array
Error: Export tools logged "Unable to get time range for monitoring".	* Confirm performance monitoring is enabled on the array. * Try invoking the export tools outside of Data Infrastructure Insights to confirm the problem lies outside of Data Infrastructure Insights.
Error: * Configuration error: Storage Array not supported by Export Utility * Configuration error: Storage Array not supported by Storage Navigator Modular CLI	* Configure only supported storage arrays. * Use “Filter Device List” to exclude unsupported storage arrays.
Error: * Error executing external command * Configuration error: Storage Array not reported by Inventory * Configuration error:export folder does not contains jar files	* Check Export utility location. * Check if Storage Array in question is configured in Command View server * Set Performance poll interval as multiple of 60 seconds.

Problem:	Try this:
Error: * Error Storage navigator CLI * Error executing auperform command * Error executing external command	* Confirm that Storage Navigator Modular CLI is installed on the Data Infrastructure Insights Acquisition Unit * Confirm that Storage Navigator Modular CLI location is correct in the data collector configuration * Confirm that the IP of the WMS/SMS/SMS array is correct in the configuration of the data collector * Confirm that Storage Navigator Modular CLI version is compatible with micro code version of storage array configured in the data collector * From the Data Infrastructure Insights Acquisition Unit, open a CMD prompt and do the following: - Change the directory to the configured installation directory - Try to make a connection with the configured storage array by executing following command “auunitref.exe”
Error: Configuration error: Storage Array not reported by Inventory	Check if Storage Array in question is configured in Command View server
Error: * No Array is registered with the Storage Navigator Modular 2 CLI * Array is not registered with the Storage Navigator Modular 2 CLI * Configuration error: Storage Array not registered with StorageNavigator Modular CLI	* Open Command prompt and change directory to the configured path * Run the command “set=STONAVM_HOME=.” * Run the command “auunitref” * Confirm that the command output contains details of the array with IP * If the output does not contain the array details then register the array with Storage Navigator CLI: - Open Command prompt and change directory to the configured path - Run the command “set=STONAVM_HOME=.” - Run command “auunitaddauto -ip ${ip}”. Replace ${ip} with real IP

Additional information may be found from the Support page or in the Data Collector Support Matrix.

HPE Alletra 6000 data collector

The HP Enterprise Alletra 6000 (previously Nimble) data collector supports inventory and performance data for Alletra 6000 storage arrays.

Terminology

Data Infrastructure Insights acquires the following inventory information from this collector. For each asset type acquired by Data Infrastructure Insights, the most common terminology used for this asset is shown. When viewing or troubleshooting this data collector, keep the following terminology in mind:

Vendor/Model Term	Data Infrastructure Insights Term
Array	Storage
Disk	Disk

Vendor/Model Term	Data Infrastructure Insights Term
Volume	Volume
Pool	Storage Pool
Initiator	Storage Host Alias
Controller	Storage Node
Fibre Channel Interface	Controller

Note: These are common terminology mappings only and might not represent every case for this data collector.

Requirements

You must have the following in order to collect inventory and configuration data from the storage array:

- The array must be installed and configured, and reachable from the client through its fully qualified domain name (FQDN) or array management IP address.
- The array must be running NimbleOS 2.3.x or later.
- You must have a valid user name and password to the array with at least "Operator" level role. The "Guest" role does not have sufficient access to understand initiator configurations.
- Port 5392 must be open on the array.

You must have the following in order to collect performance data from the storage array:

- The array must be running NimbleOS 4.0.0 or later
- The array must have volumes configured. The only performance API NimbleOS has is for volumes, and any statistics Data Infrastructure Insights reports are derived from the statistics on volumes

Configuration

Field	Description
Array Management IP Address	Fully qualified domain name (FQDN) or array management IP address.
User Name	User name for the array
Password	Password for the array

Advanced configuration

Field	Description
Port	Port used by Nimble REST API. The default is 5392.
Inventory Poll Interval (min)	Interval between inventory polls. The default is 60 minutes.

Note: The default performance poll interval is 300 seconds and can not be changed. This is the only interval supported by HPE Alletra 6000.

Hitachi Data Systems

Hitachi Vantara Command Suite data collector

The Hitachi Vantara Command Suite data collector supports the HiCommand Device Manager server. Data Infrastructure Insights communicates with the HiCommand Device Manager server using the standard HiCommand API.

Terminology

Data Infrastructure Insights acquires the following inventory information from the Hitachi Vantara Command Suite data collector. For each asset type acquired by Data Infrastructure Insights, the most common terminology used for this asset is shown. When viewing or troubleshooting this data collector, keep the following terminology in mind:

Vendor/Model Term	Data Infrastructure Insights Term
PDEV	Disk
Journal Pool	Disk Group
Storage Array	Storage
Port Controller	Storage Node
Array Group, HDS Pool	Storage Pool
Logical Unit, LDEV	Volume

Note: These are common terminology mappings only and might not represent every case for this data collector.

Storage

The following terms apply to objects or references that you might find on HDS storage asset landing pages. Many of these terms apply to other data collectors as well.

- Name – comes directly from HDS HiCommand Device Manager's "name" attribute via the GetStorageArray XML API call
- Model - comes directly from HDS HiCommand Device Manager's "arrayType" attribute via the GetStorageArray XML API call
- Vendor – HDS
- Family - comes directly from HDS HiCommand Device Manager's "arrayFamily" attribute via the GetStorageArray XML API call
- IP – this is the management IP address of the array, not an exhaustive list of all IP addresses on the array
- Raw Capacity – a base2 value representing the sum of the total capacity of all disks in this system, regardless of disk role.

Storage Pool

The following terms apply to objects or references that you might find on HDS storage pool asset landing pages. Many of these terms apply to other data collectors as well.

- Type: The value here will be one of:
  - RESERVED – if this pool is dedicated for purposes other than data volumes, i.e, journaling, snapshots
  - Thin Provisioning – if this is a HDP pool
  - Raid Group – you will not likely see these for a few reasons:

    Data Infrastructure Insights takes a strong stance to avoid double counting capacity at all costs. On HDS, one typically needs to build Raid Groups from disks, create pool volumes on those Raid Groups, and construct pools (often HDP, but could be special purpose) from those pool volumes. If Data Infrastructure Insights reported both the underlying Raid Groups as is, as well as the Pools, the sum of their raw capacity would vastly exceed the sum of the disks.

    Instead, Data Infrastructure Insights' HDS Command Suite data collector arbitrarily shrinks the size of Raid Groups by the capacity of pool volumes. This may result in Data Infrastructure Insights not reporting the Raid Group at all. Additionally, any resulting Raid Groups are flagged in a way such that they are not visible in the Data Infrastructure Insights WebUI, but they do flow into the Data Infrastructure Insights Data Warehouse (DWH). The purpose of these decisions is to avoid UI clutter for things that most users do not care about – if your HDS array has Raid Groups with 50MB free, you probably cannot use that free space for any meaningful outcome.
- Node - N/A, as HDS pools are not tied to any one specific node
- Redundancy - the RAID level of the pool. Possibly multiple values for a HDP pool comprised of multiple RAID types
- Capacity % - the percent used of the pool for data usage, with the used GB and total logical GB size of the pool
- Over-committed Capacity - a derived value, stating "the logical capacity of this pool is oversubscribed by this percentage by virtue of the sum of the logical volumes exceeding the logical capacity of the pool by this percentage"
- Snapshot - shows the capacity reserved for snapshot usage on this pool

Storage Node

The following terms apply to objects or references that you might find on HDS storage node asset landing pages. Many of these terms apply to other data collectors as well.

- Name – The name of the Front-end director (FED) or Channel Adapter on monolithic arrays, or the name of the controller on a modular array. A given HDS array will have 2 or more Storage Nodes
- Volumes – The Volume table will show any volume mapped to any port owned by this storage node

Inventory Requirements

You must have the following in order to collect inventory data:

- IP address of the HiCommand Device Manager server
- Read-only user name and password for the HiCommand Device Manager software and peer privileges
- Port requirements: 2001 (http) or 2443 (https)
- Log into HiCommand Device Manager software using username and password
- Verify access to HiCommand Device Manager http://<HiCommand_Device_Manager_IP>:2001/service/StorageManager

Performance requirements

The following requirements must be met in order to collect performance data:

- HDS USP, USP V, and VSP performance
  - Performance Monitor must be licensed.
  - Monitoring switch must be enabled.
  - The Export Tool (Export.exe) must be copied to the Data Infrastructure Insights AU.
  - The Export Tool version must match the microcode version of the target array.
- AMS performance:
  - NetApp strongly recommends creating a dedicated service account on AMS arrays for Data Infrastructure Insights to use to retrieve performance data. Storage Navigator only allows a user account one concurrent login to the array. Having Data Infrastructure Insights use the same user account as management scripts or HiCommand may result in Data Infrastructure Insights, management scripts, or HiCommand being unable to communicate to the array due to the one concurrent user account login limit
  - Performance Monitor must be licensed.
  - The Storage Navigator Modular 2 (SNM2) CLI utility needs to be installed on the Data Infrastructure Insights AU.

Configuration

Field	Description
HiCommand Server	IP address or fully-qualified domain name of the HiCommand Device Manager server
User Name	User name for the HiCommand Device Manager server.
Password	Password used for the HiCommand Device Manager server.
Devices - VSP G1000 (R800), VSP (R700), HUS VM (HM700) and USP storages	Device list for VSP G1000 (R800), VSP (R700), HUS VM (HM700) and USP storages. Each storage requires: * Array's IP: IP address of the storage * User Name: User name for the storage * Password: Password for the storage * Folder Containing Export Utility JAR Files
SNM2Devices - WMS/SMS/AMS Storages	Device list for WMS/SMS/AMS storages. Each storage requires: * Array's IP: IP address of the storage * Storage Navigator CLI Path: SNM2 CLI path * Account Authentication Valid: Select to choose valid account authentication * User Name: User name for the storage * Password: Password for the storage
Choose Tuning Manager for Performance	Override other performance options

Field	Description
Tuning Manager Host	IP address or fully-qualified domain name of tuning manager
Override Tuning Manager Port	If blank, use the default port in the Choose Tuning Manager for Performance field, otherwise enter the port to use
Tuning Manager Username	User name for Tuning Manager
Tuning Manager Password	Password for Tuning Manager

Note: In HDS USP, USP V, and VSP, any disk can belong to more than one array group.

Advanced configuration

Field	Description
Connection Type	HTTPS or HTTP, also displays the default port
HiCommand Server Port	Port used for the HiCommand Device Manager
Inventory Poll Interval (min)	Interval between inventory polls. The default is 40.
Choose 'Exclude' or 'Include' to specify a list	Specify whether to include or exclude the array list below when collecting data.
Filter device List	Comma-separated list of device serial numbers to include or exclude
Performance Poll Interval (sec)	Interval between performance polls. The default is 300.
Export timeout in seconds	Export utility timeout. The default is 300.

Troubleshooting

Some things to try if you encounter problems with this data collector:

Inventory

Problem:	Try this:
Error: User does not have enough permission	Use a different user account that has more privilege or increase the privilege of user account configured in the data collector
Error: Storages list is empty. Either devices are not configured or the user does not have enough permission	* Use DeviceManager to check if the devices are configured. * Use a different user account that has more privilege, or increase the privilege of the user account
Error: HDS storage array was not refreshed for some days	Investigate why this array is not being refreshed in HDS HiCommand.

Performance

Problem:	Try this:
Error: * Error executing export utility * Error executing external command	* Confirm that Export Utility is installed on the Data Infrastructure Insights Acquisition Unit * Confirm that Export Utility location is correct in the data collector configuration * Confirm that the IP of the USP/R600 array is correct in the configuration of the data collector * Confirm that the User name and password are correct in the configuration of the data collector * Confirm that Export Utility version is compatible with storage array micro code version * From the Data Infrastructure Insights Acquisition Unit, open a CMD prompt and do the following: - Change the directory to the configured installation directory - Try to make a connection with the configured storage array by executing batch file runWin.bat
Error: Export tool login failed for target IP	* Confirm that username/password is correct * Create a user ID mainly for this HDS data collector * Confirm that no other data collectors are configured to acquire this array
Error: Export tools logged "Unable to get time range for monitoring".	* Confirm performance monitoring is enabled on the array. * Try invoking the export tools outside of Data Infrastructure Insights to confirm the problem lies outside of Data Infrastructure Insights.
Error: * Configuration error: Storage Array not supported by Export Utility * Configuration error: Storage Array not supported by Storage Navigator Modular CLI	* Configure only supported storage arrays. * Use “Filter Device List” to exclude unsupported storage arrays.
Error: * Error executing external command * Configuration error: Storage Array not reported by Inventory * Configuration error:export folder does not contains jar files	* Check Export utility location. * Check if Storage Array in question is configured in HiCommand server * Set Performance poll interval as multiple of 60 seconds.

Problem:	Try this:
Error: * Error Storage navigator CLI * Error executing auperform command * Error executing external command	* Confirm that Storage Navigator Modular CLI is installed on the Data Infrastructure Insights Acquisition Unit * Confirm that Storage Navigator Modular CLI location is correct in the data collector configuration * Confirm that the IP of the WMS/SMS/SMS array is correct in the configuration of the data collector * Confirm that Storage Navigator Modular CLI version is compatible with micro code version of storage array configured in the data collector * From the Data Infrastructure Insights Acquisition Unit, open a CMD prompt and do the following: - Change the directory to the configured installation directory - Try to make a connection with the configured storage array by executing following command “auunitref.exe”
Error: Configuration error: Storage Array not reported by Inventory	Check if Storage Array in question is configured in HiCommand server
Error: * No Array is registered with the Storage Navigator Modular 2 CLI * Array is not registered with the Storage Navigator Modular 2 CLI * Configuration error: Storage Array not registered with StorageNavigator Modular CLI	* Open Command prompt and change directory to the configured path * Run the command “set=STONAVM_HOME=.” * Run the command “auunitref” * Confirm that the command output contains details of the array with IP * If the output does not contain the array details then register the array with Storage Navigator CLI: - Open Command prompt and change directory to the configured path - Run the command “set=STONAVM_HOME=.” - Run command “auunitaddauto -ip <ip>”. Replace <ip> with the correct IP.

Additional information may be found from the Support page or in the Data Collector Support Matrix.

Configuring the Hitachi Vantara NAS data collector

The Hitachi Vantara NAS data collector is an inventory and configuration data collector that supports discovery of HDS NAS clusters. Data Infrastructure Insights supports discovering NFS and CIFS shares, file systems (Internal Volumes), and spans (Storage Pools).

Terminology

Data Infrastructure Insights acquires the following inventory information from the HNAS data collector. For each asset type acquired by Data Infrastructure Insights, the most common terminology used for this asset is shown. When viewing or troubleshooting this data collector, keep the following terminology in mind:

Vendor/Model Term	Data Infrastructure Insights Term
Tier	Disk Group

Vendor/Model Term	Data Infrastructure Insights Term
Cluster	Storage
Node	Storage Node
Span	Storage Pool
System Drive	Backend Lun
Files System	Internal Volume

Note: These are common terminology mappings only and might not represent every case for this data collector.

Requirements

- Device IP address
- Port 22, SSH protocol
- Username and password - privilege level: Supervisor
- Note: This data collector is SSH based, so the AU that hosts it must be able to initiate SSH sessions to TCP 22 on the HNAS itself, or the Systems Management Unit (SMU) that the cluster is connected to.

Configuration

Field	Description
HNAS Host	IP address or fully-qualified domain name of HNAS Management Host
User Name	User name for HNAS CLI
Password	Password used for HNAS CLI

Advanced configuration

Field	Description
Inventory Poll Interval (min)	Interval between inventory polls. The default is 30 minutes.

Troubleshooting

Some things to try if you encounter problems with this data collector:

Inventory

Problem:	Try this:
"Error connecting" with error messages “Error setting up shell channel:”or “Error opening shell channel”	Likely caused by network connectivity issues or SSH is misconfigured. Confirm connection with alternate SSH client
"Timeout" or “Error retrieving data" with error messages “Command: XXX has timed out."	* Try the command with alternate SSH client * Increase timeout

Problem:	Try this:
"Error connecting " or “Invalid login credentials" with error messages “Could not communicate with the device:"	* Check IP address * Check user name and password * Confirm connection with alternate SSH client

Additional information may be found from the Support page or in the Data Collector Support Matrix.

Hitachi Ops Center data collector

This data collector uses Hitachi Ops Center’s integrated suite of applications to access inventory and performance data of multiple storage devices. For inventory and capacity discovery, your Ops Center installation must include both the "Common Services" and "Administrator" components. For performance collection, you must additionally have "Analyzer" deployed.

Terminology

Data Infrastructure Insights acquires the following inventory information from this data collector. For each asset type acquired by Data Infrastructure Insights, the most common terminology used for this asset is shown. When viewing or troubleshooting this data collector, keep the following terminology in mind:

Vendor/Model Term	Data Infrastructure Insights Term
Storage Systems	Storage
Volume	Volume
Parity Groups	Storage Pool(RAID), Disk Groups
Disk	Disk
Storage Pool	Storage Pool(Thin, SNAP)
External Parity Groups	Storage Pool(Backend), Disk Groups
Port	Storage Node → Controller Node → Port
Host Groups	Volume Mapping and Masking
Volume Pairs	Storage Synchronization

Note: These are common terminology mappings only and might not represent every case for this data collector.

Inventory Requirements

You must have the following in order to collect inventory data:

- IP address or hostname of the Ops Center server hosting the "Common Services" component
- Root/sysadmin user account and password that exist on all servers hosting Ops Center components. HDS did not implement REST API support for usage by LDAP/SSO users until Ops Center 10.8+

Performance requirements

The following requirements must be met in order to collect performance data:

The HDS Ops Center "Analyzer" module must be installed
Storage arrays must be feeding the Ops Center "Analyzer" module

Configuration

Field	Description
Hitachi Ops Center IP Address	IP address or fully-qualified domain name of the Ops Center server hosting the "Common Services" component
User Name	User name for the Ops Center server.
Password	Password used for the Ops Center server.

Advanced configuration

Field	Description
Connection Type	HTTPS (port 443) is the default
Override TCP Port	Specify the port to use if not the default
Inventory Poll Interval (min)	Interval between inventory polls. The default is 40.
Choose 'Exclude' or 'Include' to specify a list	Specify whether to include or exclude the array list below when collecting data.
Filter device List	Comma-separated list of device serial numbers to include or exclude
Performance Poll Interval (sec)	Interval between performance polls. The default is 300.

Additional information may be found from the Support page or in the Data Collector Support Matrix.

Infinidat InfiniBox data collector

The Infinidat InfiniBox (HTTP) data collector is used to collect inventory information from the Infinidat InfiniBox storage system.

Terminology

Data Infrastructure Insights acquires the following inventory information from the Infinidat InfiniBox data collector. For each asset type acquired by Data Infrastructure Insights, the most common terminology used for this asset is shown. When viewing or troubleshooting this data collector, keep the following terminology in mind:

Vendor/Model Term	Data Infrastructure Insights Term
Storage Pool	Storage Pool
Node	Controller
Filesystem	Internal Volume
Filesystem	File Share

Vendor/Model Term	Data Infrastructure Insights Term
Filesystem Exports	Share

Requirements

The following are requirements when configuring this data collector.

- IP address or FQDN of InfiniBox management Node
- Admin userid and password
- Port 443 via REST API

Configuration

Field	Description
InfiniBox Host	IP address or fully-qualified domain name of the InfiniBox Management Node
User Name	User name for InfiniBox Management Node
Password	Password for the InfiniBox Management Node

Advanced configuration

Field	Description
TCP Port	TCP Port used to connect to InfiniBox Server. The default is 443.
Inventory Poll Interval	Interval between inventory polls. The default is 60 minutes.

Troubleshooting

Additional information on this Data Collector may be found from the Support page or in the Data Collector Support Matrix.

Huawei OceanStor data collector

Data Infrastructure Insights uses the Huawei OceanStor (REST/HTTPS) data collector to discover inventory and performance for Huawei OceanStor and OceanStor Dorado storage.

Terminology

Data Infrastructure Insights acquires the following inventory and performance information from the Huawei OceanStor. For each asset type acquired by Data Infrastructure Insights, the most common terminology used for this asset is shown. When viewing or troubleshooting this data collector, keep the following terminology in mind:

Vendor/Model Term	Data Infrastructure Insights Term
Storage Pool	Storage Pool
File System	Internal Volume
Controller	Storage Node
FC Port (Mapped)	Volume Map
Host FC Initiator (Mapped)	Volume Mask
NFS/CIFS Share	Share
iSCSI Link Target	iSCSI Target Node
iSCSI Link Initiator	iSCSI Initiator Node
Disk	Disk
LUN	Volume

Requirements

The following requirements are required to configure this data collector:

- Device IP address
- Credentials to access OceanStor device manager
- Port 8088 must be available

Configuration

Field	Description
OceanStor Host IP Address	IP address or fully-qualified domain name of the OceanStor Device Manager
User Name	Name used to log into the OceanStor Device Manager
Password	Password used to log into the OceanStor Device Manager

Advanced Configuration

Field	Description
TCP Port	TCP Port used to connect to OceanStor Device Manager. The default is 8088.
Inventory Poll Interval (min)	Interval between inventory polls. The default is 60 minutes.
Performance poll interval (sec).	The default is 300 seconds.

Troubleshooting

Additional information on this Data Collector may be found from the Support page or in the Data Collector Support Matrix.

IBM

IBM Cleversafe data collector

Data Infrastructure Insights uses this data collector to discover inventory and performance data for IBM Cleversafe storage systems.

IBM Cleversafe is metered at a different Raw TB to Managed Unit rate. Every 40 TB of unformatted IBM Cleversafe capacity is charged as 1 Managed Unit (MU).

Terminology

Data Infrastructure Insights acquires the following inventory information from the IBM Cleversafe data collector. For each asset type acquired by Data Infrastructure Insights, the most common terminology used for this asset is shown. When viewing or troubleshooting this data collector, keep the following terminology in mind:

Vendor/Model Term	Data Infrastructure Insights Term
Storage Pool	Storage Pool
Container	Internal Volume
Container	File Share
NFS Share	Share

Note: These are common terminology mappings only and might not represent every case for this data collector.

Requirements

- The external data services IP address for the cluster
- Administrator user name and password
- Port 9440

Configuration

Field	Description
Manager IP or host name	IP address or hostname of management node
User name	Username for the user account with super user or system administrator role
Password	Password for the user account with super user or system administrator role

Advanced configuration

Field	Description
Inventory poll interval (min)	Interval between inventory polls.
HTTP Connection Timeout (sec)	HTTP timeout in seconds.

Troubleshooting

Additional information on this Data Collector may be found from the Support page or in the Data Collector Support Matrix.

IBM CS data collector

Data Infrastructure Insights uses this data collector to discover inventory and performance data for IBM CS storage systems.

Terminology

Data Infrastructure Insights acquires the following inventory information from the IBM CS data collector. For each asset type acquired by Data Infrastructure Insights, the most common terminology used for this asset is shown. When viewing or troubleshooting this data collector, keep the following terminology in mind:

Vendor/Model Term	Data Infrastructure Insights Term
Storage Pool	Storage Pool
Container	Internal Volume
Container	File Share
NFS Share	Share

Note: These are common terminology mappings only and might not represent every case for this data collector.

Requirements

- The external data services IP address for the cluster
- Administrator user name and password
- Port 9440

Configuration

Field	Description
Prism External IP Address	The external data services IP address for the cluster
User name	User name for the Admin account
Password	Password for the Admin account

Advanced configuration

Field	Description
TCP port	TCP Port used to connect to the IBM CS array. The default is 9440.
Inventory poll interval (min)	Interval between inventory polls. The default is 60 minutes.

Field	Description
Performance poll interval(sec)	Interval between performance polls. The default is 300 seconds.

Troubleshooting

Additional information on this Data Collector may be found from the Support page or in the Data Collector Support Matrix.

IBM System Storage DS8000 Series data collector

The IBM DS (CLI) data collector supports inventory and performance data acquisition for DS6xxx and DS8xxx devices.

DS3xxx, DS4xxx, and DS5xxx devices are supported by the NetApp E-Series data collector. You should refer to the Data Infrastructure Insights support matrix for supported models and firmware versions.

Terminology

Data Infrastructure Insights acquires the following inventory information from the IBM DS data collector. For each asset type acquired by Data Infrastructure Insights, the most common terminology used for this asset is shown. When viewing or troubleshooting this data collector, keep the following terminology in mind:

Vendor/Model Term	Data Infrastructure Insights Term
Disk Drive Module	Disk
Storage Image	Storage
Extent Pool	Storage Node
Fixed Block Volume	Volume
Host FC Initiator (Mapped)	Volume Mask

Note: These are common terminology mappings only and might not represent every case for this data collecor.

Requirements

You need the following to configure this data collector:

- IP address of each DS array
- Read-only username and password on each DS array
- Third-party software installed on the Data Infrastructure Insights AU: IBM *dscli*
- Access validation: Run *dscli* commands using the username and password
- Port requirements: 80, 443, & 1750

Configuration

Field	Description
DS Storage	IP address or fully-qualified domain name of the DS device

Field	Description
User Name	User name for the DS CLI
Password	Password for the DS CLI
dscli executable path	Full path to the *dscli* executable

Advanced configuration

Field	Description
Inventory Poll Interval (min)	Interval between inventory polls (min). The default is 40.
Storage Display Name	Name of the IBM DS storage array
Inventory Exclude Devices	Comma-separated list of device serial numbers to exclude from inventory collection
Performance Poll Interval (sec)	The default is 300.
Performance Filter Type	Include: Data collected only from devices on list. Exclude: No data from these devices is collected
Performance Filter Device List	Comma-separated list of device IDs to include or exclude from performance collection

Troubleshooting

Some things to try if you encounter problems with this data collector:

Inventory

Problem:	Try this:
Error containing: CMUC00192E, CMUC00191E or CMUC00190E	* Verify credentials and IP address entered. * Try to communicate with the array through web management console https://<ip>:8452/DS8000/Console. Replace <ip> with data collector configured IP.
Error: * Cannot run program * Error executing command	* From Data Infrastructure Insights Acquisition Unit Open a CMD * Open CLI.CFG file in CLI's home dir/lib and check property JAVA_INSTALL, edit the value to match your environment * Display Java version installed on this machine, typing: "java -version" * Ping the IP address of the IBM Storage device specified in CLI command issued. * If all the above worked fine then manually run a CLI command

Additional information may be found from the Support page or in the Data Collector Support Matrix.

Configuring the IBM SAN Volume Controller data collector

The IBM SAN Volume Controller (SVC) data collector collects inventory and performance data using SSH, supporting a variety of devices that run the SVC operating system.

The list of supported devices includes models such as the SVC, the v7000, the v5000, and the v3700. Refer to the Data Infrastructure Insights support matrix for supported models and firmware versions.

Terminology

Data Infrastructure Insights acquires the following inventory information from the IBM SVC data collector. For each asset type acquired by Data Infrastructure Insights, the most common terminology used for this asset is shown. When viewing or troubleshooting this data collector, keep the following terminology in mind:

Vendor/Model Term	Data Infrastructure Insights Term
Drive	Disk
Cluster	Storage
Node	Storage Node
Mdisk Group	Storage Pool
Vdisk	Volume
Mdisk	Backend LUNs and paths

Note: These are common terminology mappings only and might not represent every case for this data collector.

Inventory Requirements

- IP address of each SVC cluster
- Port 22 available
- Read-only user name and password

Performance Requirements

- SVC Console, which is mandatory for any SVC cluster and required for the SVC discovery foundation package.
- Credentials will require administrative access level only for copying performance files from cluster nodes to the config node.
- Enable data collection by connecting to the SVC cluster by SSH and running: *svctask startstats -interval 1*

  Note: Alternatively, enable data collection using the SVC management user interface.

Configuration

Field	Description
Cluster IP Addresses	IP addresses or fully-qualified domain names of the SVC storage

Field	Description
Inventory User Name	User name for the SVC CLI
Inventory Password	Password for the SVC CLI

Advanced configuration

Field	Description
Inventory Poll Interval (min)	Interval between inventory polls. The default is 40 minutes.
Performance Poll Interval (sec)	Interval between performance polls. The default is 300 seconds.
To clean up dumped stats files	Select this checkbox to clean up dumped stats files

Troubleshooting

Some things to try if you encounter problems with this data collector:

Problem:	Try this:
Error: "The command cannot be initiated because it was not run on the configuration node."	The command must be executed on the configuration node.

Some things to try if you encounter problems with this data collector:

Problem:	Try this:
Error: "The command cannot be initiated because it was not run on the configuration node."	The command must be executed on the configuration node.

Additional information on this Data Collector may be found from the Support page or in the Data Collector Support Matrix.

Configuring the IBM XIV/A9000 data collector

IBM XIV and A9000 (CLI) data collector uses the XIV command-line interface to collect inventory data while performance collection is accomplished by making SMI-S calls to the XIV/A9000 array, which runs a SMI-S provider on port 7778.

Terminology

Vendor/Model Term	Data Infrastructure Insights Term
Disk	Disk
Storage System	Storage
Storage Pool	Storage Pool
Volume	Volume

Requirements

The following requirements must be met to configure and use this data collector:

- Port requirement: TCP port 7778
- Read-only user name and password
- The XIV CLI must be installed on the AU

Performance requirements

The following are requirements for performance collection:

- SMI-S Agent 1.4 or higher
- SMI-S compatible CIMService running on array. Most XIV arrays have a CIMServer installed by default.
- User login must be provided for the CIMServer. The login must have full read access to the array configuration and properties.
- SMI-S namespace. Default is root/ibm. This is configurable in the CIMServer.
- Port Requirements: 5988 for HTTP, 5989 for HTTPS.
- Refer to the following link on how to create an account for SMI-S performance collection: https://www.ibm.com/docs/en/products?topic=/com.ibm.tpc_V41.doc/fqz0_t_adding_cim_agent.html

Configuration

Field	Description
XIV IP address	IP address or fully-qualified domain name of the XIV storage
User Name	User name for the XIV storage
Password	Password for the XIV storage
Full Path to XIV CLI Directory	Full path to the folder containing the XIV CLI
SMI-S Host IP Address	IP address of the SMI-S host

Advanced configuration

Field	Description
Inventory Poll Interval (min)	Interval between inventory polls. The default is 40 minutes.
SMI-S Protocol	Protocol used to connect to the SMI-S provider. Also displays the default port.
Override SMI-S Port	If blank, use the default port in the Connection Type field, otherwise enter the connection port to use
Username	User name for the SMI-S Provider Host
Password	Password for the SMI-S Provider Host
Performance Poll Interval (sec)	Interval between performance polls. The default is 300 seconds.

Troubleshooting

Additional information on this Data Collector may be found from the Support page or in the Data Collector Support Matrix.

Lenovo data collector

Data Infrastructure Insights uses the Lenovo data collector to discover inventory and performance data for Lenovo HX storage systems.

Note: These are common terminology mappings only and might not represent every case for this data collector.

Requirements

- Prism External IP Address
- Administrator user name and password
- TCP Port requirement: 9440

Configuration

Field	Description
Prism External IP Address	The external data services IP address for the cluster
User name	User name for the Admin account
Password	Password for the Admin account

Advanced configuration

Field	Description
TCP port	TCP Port used to connect to array. The default is 9440.
Inventory poll interval (min)	Interval between inventory polls. The default is 60 minutes.
Performance poll interval (sec)	Interval between performance polls. The default is 300 seconds.

Troubleshooting

Additional information on this Data Collector may be found from the Support page or in the Data Collector Support Matrix.

Microsoft

Configuring the Azure NetApp Files data collector

Data Infrastructure Insights uses the Azure NetApp Files data collector to acquire inventory and performance data.

Requirements

You need the following information to configure this data collector.

- Port requirement: 443 HTTPS
- Azure Management Rest IP (management.azure.com)
- Azure service principal client ID (user account)
- Azure service principal authentication key (user password)
- You need to set up an Azure account for Data Infrastructure Insights discovery.

  Once the account is properly configured and you register the application in Azure, you will have the credentials required to discover the Azure instance with Data Infrastructure Insights. The following link describes how to set up the account for discovery:

https://docs.microsoft.com/en-us/azure/active-directory/develop/howto-create-service-principal-portal

Configuration

Enter data into the data collector fields according to the table below:

Field	Description
Azure Service Principal Client ID	Sign-in ID to Azure
Azure Tenant ID	Azure Tenant ID
Azure Service Principal Authentication Key	Login authentication key
I understand Microsoft bills me for API requests	Check this to verify your understanding that Microsoft bills you for API requests made by Insight polling.

Advanced Configuration

Field	Description
Inventory Poll Interval (min)	The default is 60

Troubleshooting

- The credentials used by your ANF data collector must have access to any Azure subscriptions that contain ANF volumes.
- If Reader access causes performance collection to fail, try granting contributor access on a resource group level.

Additional information on this Data Collector may be found from the Support page or in the Data Collector Support Matrix.

Microsoft Hyper-V data collector

The Microsoft Hyper-V data collector acquires inventory and performance data from the virtualized server computing environment. This data collector can discover a standalone Hyper-V host, or an entire cluster - create one collector per standalone host or cluster.

Terminology

Data Infrastructure Insights acquires the following inventory information from the Microsoft Hyper-V (WMI). For each asset type acquired by Data Infrastructure Insights, the most common terminology used for this asset is shown. When viewing or troubleshooting this data collector, keep the following terminology in mind:

Vendor/Model Term	Data Infrastructure Insights Term
Virtual Hard Disk	Virtual Disk
Host	Host
Virtual Machine	Virtual Machine
Cluster Shared Volumes (CSV), Partition Volume	Data Store
Internet SCSI Device, Multi Path SCSI LUN	LUN
Fiber Channel Port	Port

Note: These are common terminology mappings only and might not represent every case for this data collector.

Requirements

The following are required to configure this data collector:

- The Hyper-V requires port 5985 opened for data collection and remote access/management.
- IP address or FQDN of cluster or standalone hypervisor. Using the floating cluster hostname or IP is likely the most reliable approach versus pointing the collector at just one specific node in a cluster.
- Administrative-level user account that works on all the hypervisors in the cluster.
- WinRM needs to be enabled and listening on all hypervisors
- Port requirements: Port 135 via WMI & Dynamic TCP ports assigned 1024-65535 for Windows 2003 and older and 49152-65535 for Windows 2008.
- DNS resolution must succeed, even if the data collector is pointed at only an IP address
- Each Hyper-V hypervisor must have “Resource Metering” turned on for every VM, on every host. This allows each hypervisor to have more data available for Data Infrastructure Insights on each guest. If this is not set, fewer performance metrics are acquired for each guest. More information on Resource metering can be found in the Microsoft documentation:

  Hyper-V Resource Metering Overview

  Enable-VMResourceMetering

The Hyper-V data collector requires a Windows Acquisition Unit.

Configuration

Field	Description
Cluster IP address or floating cluster FQDN	The IP address or fully-qualified domain name for the the cluster, or a standalone, non-clustered hypervisor
User Name	Administrator user name for the hypervisor

Field	Description
Password	Password for the hypervisor
DNS domain suffix	The hostname suffix that combines with the simple hostname to render the FQDN of a hypervisor

Advanced configuration

Field	Description
Inventory Poll Interval (min)	The default is 20 minutes.

Troubleshooting

Additional information on this Data Collector may be found from the Support page or in the Data Collector Support Matrix.

NetApp

NetApp Cloud Volumes ONTAP data collector

This data collector supports inventory collection from Cloud Volumes ONTAP configurations.

Configuration

Field	Description
NetApp Management IP Address	IP address for Cloud Volumens ONTAP
User Name	User name for Cloud Volumes ONTAP
Password	Password for the above user

Advanced configuration

Field	Description
Connection Type	HTTPS recommended. Also shows default port.
Override Communication Port	Port to use if not default.
Inventory Poll Interval (min)	Default is 60 minutes.
Inventory Concurrent Thread Count	Number of concurrent threads.
Force TLS for HTTPS	Force TLS over HTTPS
Automatically Lookup Netgroups	Automatically Lookup Netgroups
Netgroup Expansion	Select Shell or File
HTTP read timeout seconds	Default is 30 seconds
Force responses as UTF-8	Force responses as UTF-8
Performance Poll Interval (min)	Default is 900 seconds.

Field	Description
Performance Concurrent Thread Count	Number of concurrent threads.
Advanced Counter Data Collection	Check this to have Data Infrastructure Insights collect the advanced metrics from the list below.

Troubleshooting

Additional information on this Data Collector may be found from the Support page or in the Data Collector Support Matrix.

NetApp ONTAP AFX data collector

This data collector acquires inventory, EMS logs, and performance data from storage systems running ONTAP 9.16.0 and higher using REST API calls.

Requirements

The following are requirements to configure and use this data collector:

- You must have access to a user account with the required level of access. Note that Admin permissions are required if creating a new REST user/role.
  - Functionally, Data Infrastructure Insights primarily makes read requests, but some write permissions are required for Data Infrastructure Insights to register with the ONTAP array. See the *Note About Permissions* immediately below.
- ONTAP version 9.16.0 or higher.
- Port requirements: 443

A Note About Permissions

Since a number of Data Infrastructure Insights' ONTAP dashboards rely on advanced ONTAP counters, you should keep **Enable Advanced Counter Data Collection** enabled in the data collector Advanced Configuration section.

To create a local account for Data Infrastructure Insights at the cluster level, log in to ONTAP with the Cluster management Administrator username/password, and execute the following commands on the ONTAP server:

1. Before you begin, you must be signed in to ONTAP with an *Administrator* account, and *diagnostic-level commands* must be enabled.
2. Retrieve the name of the vserver that is of type *admin*. You will using this name in subsequent commands.

```
vserver show -type admin
```

3. Create a role using the following commands:

```
security login rest-role create -role {role name} -api /api -access
readonly
security login rest-role create -role {role name} -api
/api/cluster/agents -access all
vserver services web access create -name spi -role {role name} -vserver
{vserver name as retrieved above}
security login create -user-or-group-name {username} -application http
-authentication-method password -role {role name}
```

4. Create the read-only user using the following command. Once you have executed the create command, you will be prompted to enter a password for this user.

```
security login create -username ci_user -application http
-authentication-method password -role ci_readonly
```

If AD/LDAP account is used, the command should be

```
security login create -user-or-group-name DOMAIN\aduser/adgroup
-application http -authentication-method domain -role ci_readonly
```

The resulting role and user login will look something like the following. Your actual output may vary:

```
security login rest-role show -vserver <vserver name> -role restRole

               Role                                    Access
Vserver        Name            API                     Level
----------     -------------   --------------------    ------
<vserver name> restRole        /api                    readonly
                               /api/cluster/agents     all
2 entries were displayed.

security login show -vserver <vserver name> -user-or-group-name restUser

Vserver: <vserver name>
                                                                  Second
User/Group                 Authentication                 Acct
Authentication
Name           Application Method        Role Name        Locked Method
-------------- ----------- ------------- ---------------- ------
--------------
restUser       http        password      restRole         no     none
```

Migration

To migrate from a previous ONTAP (ontapi) data collector to the newer ONTAP REST collector, do the following:

1. Add the REST Collector. It is recommended to enter information for a different user than the one configured for the previous collector. For example, use the user noted in the Permissions section above.
2. Pause the previous collector, so it doesn't continue to collect data.
3. Let the new REST collector acquire data for at least 30 minutes. Ignore any data during this time that does not appear "normal".
4. After the rest period, you should see your data stabilize as the REST collector continues to to acquire.

You can use this same process to return to the previous collector, should you wish.

Configuration

Field	Description
ONTAP management IP Address	IP address or fully-qualified domain name of the NetApp cluster. Must be Cluster Management IP/FQDN.
ONTAP REST User Name	User name for NetApp cluster
ONTAP REST Password	Password for NetApp cluster

Advanced configuration

Field	Description
Inventory Poll Interval (min)	Default is 60 minutes.
Performance Poll Interval (sec)	Default is 60 seconds.
Advanced Counter Data Collection	Select this to include ONTAP Advanced Counter data in polls. Enabled by default.
Enable EMS Event Collection	Select this to include ONTAP EMS log event data. Enabled by default.
EMS Poll Interval (sec)	Default is 60 seconds.

Terminology

Data Infrastructure Insights acquires inventory, logs and performance data from the ONTAP data collector. For each asset type acquired, the most common terminology used for the asset is shown. When viewing or troubleshooting this data collector, keep the following terminology in mind:

Vendor/Model Term	Data Infrastructure Insights Term
Cluster	Storage
Node	StorageNode
Volume/FlexVol/FlexGroup	InternalVolume
Qtree/FlexTree	Qtree (type : Explixit/Default)

Vendor/Model Term	Data Infrastructure Insights Term
S3 Bucket	Qtree (type: Bucket)
StoragePod/StorageAvailabilityZone	StoragePool
Share	Share
Quota	Quota
SVM (Storage Virtual Machine)	Storage Virtual Machine

ONTAP Data Management Terminology

The following terms apply to objects or references that you might find on ONTAP Data Management storage asset landing pages. Many of these terms apply to other data collectors as well.

Storage

- Model – A comma-delimited list of the unique, discrete node model names within this cluster. If all the nodes in the clusters are the same model type, just one model name will appear.
- Vendor – same Vendor name you would see if you were configuring a new data source.
- Serial number – The array UUID
- IP – generally will be the IP(s) or hostname(s) as configured in the data source.
- Microcode version – firmware.
- Raw Capacity – base 2 summation of all the physical disks in the system, regardless of their role.
- Latency – a representation of what the host facing workloads are experiencing, across both reads and writes. Ideally, Data Infrastructure Insights is sourcing this value directly, but this is often not the case. In lieu of the array offering this up, Data Infrastructure Insights is generally performing an IOPs-weighted calculation derived from the individual internal volumes' statistics.
- Throughput – aggregated from internal volumes.
  Management – this may contain a hyperlink for the management interface of the device. Created programmatically by the Data Infrastructure Insights data source as part of inventory reporting.

Storage Pool

- Storage – what storage array this pool lives on. Mandatory.
- Type – a descriptive value from a list of an enumerated list of possibilities. Most commonly will be "Aggregate" or "RAID Group"".
- Node – if this storage array's architecture is such that pools belong to a specific storage node, its name will be seen here as a hyperlink to its own landing page.
- Uses Flash Pool – Yes/No value – does this SATA/SAS based pool have SSDs used for caching acceleration?
- Redundancy – RAID level or protection scheme. RAID_DP is dual parity, RAID_TP is triple parity.
- Capacity – the values here are the logical used, usable capacity and the logical total capacity, and the percentage used across these.
- Over-committed capacity – If by using efficiency technologies you have allocated a sum total of volume or internal volume capacities larger than the logical capacity of the storage pool, the percentage value here will be greater than 0%.

- Snapshot – snapshot capacities used and total, if your storage pool architecture dedicates part of its capacity to segments areas exclusively for snapshots. ONTAP in MetroCluster configurations are likely to exhibit this, while other ONTAP configurations are less so.
- Utilization – a percentage value showing the highest disk busy percentage of any disk contributing capacity to this storage pool. Disk utilization does not necessarily have a strong correlation with array performance – utilization may be high due to disk rebuilds, deduplication activities, etc in the absence of host driven workloads. Also, many arrays' replication implementations may drive disk utilization while not showing as internal volume or volume workload.
- IOPS – the sum IOPs of all the disks contributing capacity to this storage pool.
  Throughput – the sum throughput of all the disks contributing capacity to this storage pool.

Storage Node

- Storage – what storage array this node is part of. Mandatory.
- HA Partner – on platforms where a node will fail over to one and only one other node, it will generally be seen here.
- State – health of the node. Only available when the array is healthy enough to be inventoried by a data source.
- Model – model name of the node.
- Version – version name of the device.
- Serial number – The node serial number.
- Memory – base 2 memory if available.
- Utilization – On ONTAP, this is a controller stress index from a proprietary algorithm. With every performance poll, a number between 0 and 100% will be reported that is the higher of either WAFL disk contention, or average CPU utilization. If you observe sustained values > 50%, that is indicative of undersizing – potentially a controller/node not large enough or not enough spinning disks to absorb the write workload.
- IOPS – Derived directly from ONTAP REST calls on the node object.
- Latency – Derived directly from ONTAP REST calls on the node object.
- Throughput – Derived directly from ONTAP REST calls on the node object.
- Processors – CPU count.

ONTAP Power Metrics

Several ONTAP models provide power metrics for Data Infrastructure Insights that can be used for monitoring or alerting. The lists of supported and unsupported models below are not comprehensive but should provide some guidance; in general, if a model is in the same family as one on the list, the support should be the same.

Supported Models:

A200
A220
A250
A300
A320
A400
A700
A700s

A800
A900
C190
FAS2240-4
FAS2552
FAS2650
FAS2720
FAS2750
FAS8200
FAS8300
FAS8700
FAS9000

Unsupported Models:

FAS2620
FAS3250
FAS3270
FAS500f
FAS6280
FAS/AFF 8020
FAS/AFF 8040
FAS/AFF 8060
FAS/AFF 8080

Troubleshooting

Some things to try if you encounter problems with this data collector:

Problem:	Try this:
When attempting to create an ONTAP REST data collector, an error like the following is seen: Configuration: 10.193.70.14: ONTAP rest API at 10.193.70.14 is not available: 10.193.70.14 failed to GET /api/cluster: 400 Bad Request	This is likely due to an oldeer ONTAP array) for example, ONTAP 9.6) which has no REST API capabilities. ONTAP 9.14.1 is the minimum ONTAP version supported by the ONTAP REST collector. "400 Bad Request" responses should be expected on pre-REST ONTAP releases. For ONTAP versions that do support REST but are not 9.14.1 or later, you may see the following simillar message: Configuration: 10.193.98.84: ONTAP rest API at 10.193.98.84 is not available: 10.193.98.84: ONTAP rest API at 10.193.98.84 is available: cheryl5-cluster-2 9.10.1 a3cb3247-3d3c-11ee-8ff3-005056b364a7 but is not of minimum version 9.14.1.

Problem:	Try this:
I see empty or "0" metrics where the ONTAP ontapi collector shows data.	ONTAP REST does not report metrics that are used internally on the ONTAP system only. For example, system aggregates will not be collected by ONTAP REST, only SVM's of type "data" will be collected. Other examples of ONTAP REST metrics that may report zero or empty data: InternalVolumes: REST no longer reports vol0. Aggregates: REST no longer reports aggr0. Storage: most metrics are a rollup of the Internal Volume metrics, and will be impacted by the above. Storage Virtual Machines: REST no longer reports SVM's of type other than 'data' (e.g. 'cluster', 'mgmt', 'node'). You may also notice a change in the appearance of graphs that do have data, due to the change in default performance polling period from 15 minutes to 5 minutes. More frequent polling means more data points to plot.

Additional information may be found from the Support page or in the Data Collector Support Matrix.

NetApp ONTAP ASA r2 (All-SAN Array) data collector

This data collector acquires inventory, EMS logs, and performance data from storage systems running ONTAP 9.16.0 and higher using REST API calls.

Requirements

The following are requirements to configure and use this data collector:

- You must have access to a user account with the required level of access. Note that Admin permissions are required if creating a new REST user/role.
  - Functionally, Data Infrastructure Insights primarily makes read requests, but some write permissions are required for Data Infrastructure Insights to register with the ONTAP array. See the *Note About Permissions* immediately below.
- ONTAP version 9.16.0 or higher.
- Port requirements: 443

ASA R2 denotes the latest generation models of the ONTAP ASA storage platform. This includes array models ASA A1K, A90, A70, A50, A30 and A20.

For all previous generation ASA systems, please use the ONTAP REST collector.

A Note About Permissions

Since a number of Data Infrastructure Insights' ONTAP dashboards rely on advanced ONTAP counters, you should keep **Enable Advanced Counter Data Collection** enabled in the data collector Advanced

Configuration section.

To create a local account for Data Infrastructure Insights at the cluster level, log in to ONTAP with the Cluster management Administrator username/password, and execute the following commands on the ONTAP server:

1. Before you begin, you must be signed in to ONTAP with an *Administrator* account, and *diagnostic-level commands* must be enabled.
2. Retrieve the name of the vserver that is of type *admin*. You will using this name in subsequent commands.

```
vserver show -type admin
```

3. Create a role using the following commands:

```
security login rest-role create -role {role name} -api /api -access
readonly
security login rest-role create -role {role name} -api
/api/cluster/agents -access all
vserver services web access create -name spi -role {role name} -vserver
{vserver name as retrieved above}
security login create -user-or-group-name {username} -application http
-authentication-method password -role {role name}
```

4. Create the read-only user using the following command. Once you have executed the create command, you will be prompted to enter a password for this user.

```
security login create -username ci_user -application http
-authentication-method password -role ci_readonly
```

If AD/LDAP account is used, the command should be

```
security login create -user-or-group-name DOMAIN\aduser/adgroup
-application http -authentication-method domain -role ci_readonly
```

The resulting role and user login will look something like the following. Your actual output may vary:

```
security login rest-role show -vserver <vserver name> -role restRole

               Role                                    Access
Vserver        Name            API                     Level
----------     -------------   --------------------    ------
<vserver name> restRole        /api                    readonly
                               /api/cluster/agents     all
2 entries were displayed.

security login show -vserver <vserver name> -user-or-group-name restUser

Vserver: <vserver name>
                                                                  Second
User/Group                 Authentication                 Acct
Authentication
Name           Application Method        Role Name        Locked Method
-------------- ----------- ------------- ---------------- ------
--------------
restUser       http        password      restRole         no     none
```

Migration

To migrate from a previous ONTAP (ontapi) data collector to the newer ONTAP REST collector, do the following:

1. Add the REST Collector. It is recommended to enter information for a different user than the one configured for the previous collector. For example, use the user noted in the Permissions section above.
2. Pause the previous collector, so it doesn't continue to collect data.
3. Let the new REST collector acquire data for at least 30 minutes. Ignore any data during this time that does not appear "normal".
4. After the rest period, you should see your data stabilize as the REST collector continues to to acquire.

You can use this same process to return to the previous collector, should you wish.

Configuration

Field	Description
ONTAP management IP Address	IP address or fully-qualified domain name of the NetApp cluster. Must be Cluster Management IP/FQDN.
ONTAP REST User Name	User name for NetApp cluster
ONTAP REST Password	Password for NetApp cluster

Advanced configuration

Field	Description
Inventory Poll Interval (min)	Default is 60 minutes.
Performance Poll Interval (sec)	Default is 60 seconds.
Advanced Counter Data Collection	Select this to include ONTAP Advanced Counter data in polls. Enabled by default.
Enable EMS Event Collection	Select this to include ONTAP EMS log event data. Enabled by default.
EMS Poll Interval (sec)	Default is 60 seconds.

Terminology

Data Infrastructure Insights acquires inventory, logs and performance data from the ONTAP data collector. For each asset type acquired, the most common terminology used for the asset is shown. When viewing or troubleshooting this data collector, keep the following terminology in mind:

Vendor/Model Term	Data Infrastructure Insights Term
Disk	Disk
Raid Group	Disk Group
Cluster	Storage
Node	Storage Node
Aggregate	Storage Pool
LUN	Volume
Volume	Internal Volume
Storage Virtual Machine/Vserver	Storage Virtual Machine

ONTAP Data Management Terminology

The following terms apply to objects or references that you might find on ONTAP Data Management storage asset landing pages. Many of these terms apply to other data collectors as well.

Storage

- Model – A comma-delimited list of the unique, discrete node model names within this cluster. If all the nodes in the clusters are the same model type, just one model name will appear.
- Vendor – same Vendor name you would see if you were configuring a new data source.
- Serial number – The array UUID
- IP – generally will be the IP(s) or hostname(s) as configured in the data source.
- Microcode version – firmware.
- Raw Capacity – base 2 summation of all the physical disks in the system, regardless of their role.
- Latency – a representation of what the host facing workloads are experiencing, across both reads and writes. Ideally, Data Infrastructure Insights is sourcing this value directly, but this is often not the case. In lieu of the array offering this up, Data Infrastructure Insights is generally performing an IOPs-weighted calculation derived from the individual internal volumes' statistics.

- Throughput – aggregated from internal volumes.
  Management – this may contain a hyperlink for the management interface of the device. Created programmatically by the Data Infrastructure Insights data source as part of inventory reporting.

Storage Pool

- Storage – what storage array this pool lives on. Mandatory.
- Type – a descriptive value from a list of an enumerated list of possibilities. Most commonly will be “Aggregate” or “RAID Group””.
- Node – if this storage array’s architecture is such that pools belong to a specific storage node, its name will be seen here as a hyperlink to its own landing page.
- Uses Flash Pool – Yes/No value – does this SATA/SAS based pool have SSDs used for caching acceleration?
- Redundancy – RAID level or protection scheme. RAID_DP is dual parity, RAID_TP is triple parity.
- Capacity – the values here are the logical used, usable capacity and the logical total capacity, and the percentage used across these.
- Over-committed capacity – If by using efficiency technologies you have allocated a sum total of volume or internal volume capacities larger than the logical capacity of the storage pool, the percentage value here will be greater than 0%.
- Snapshot – snapshot capacities used and total, if your storage pool architecture dedicates part of its capacity to segments areas exclusively for snapshots. ONTAP in MetroCluster configurations are likely to exhibit this, while other ONTAP configurations are less so.
- Utilization – a percentage value showing the highest disk busy percentage of any disk contributing capacity to this storage pool. Disk utilization does not necessarily have a strong correlation with array performance – utilization may be high due to disk rebuilds, deduplication activities, etc in the absence of host driven workloads. Also, many arrays’ replication implementations may drive disk utilization while not showing as internal volume or volume workload.
- IOPS – the sum IOPs of all the disks contributing capacity to this storage pool.
  Throughput – the sum throughput of all the disks contributing capacity to this storage pool.

Storage Node

- Storage – what storage array this node is part of. Mandatory.
- HA Partner – on platforms where a node will fail over to one and only one other node, it will generally be seen here.
- State – health of the node. Only available when the array is healthy enough to be inventoried by a data source.
- Model – model name of the node.
- Version – version name of the device.
- Serial number – The node serial number.
- Memory – base 2 memory if available.
- Utilization – On ONTAP, this is a controller stress index from a proprietary algorithm. With every performance poll, a number between 0 and 100% will be reported that is the higher of either WAFL disk contention, or average CPU utilization. If you observe sustained values > 50%, that is indicative of undersizing – potentially a controller/node not large enough or not enough spinning disks to absorb the write workload.

- IOPS – Derived directly from ONTAP REST calls on the node object.
- Latency – Derived directly from ONTAP REST calls on the node object.
- Throughput – Derived directly from ONTAP REST calls on the node object.
- Processors – CPU count.

ONTAP Power Metrics

Several ONTAP models provide power metrics for Data Infrastructure Insights that can be used for monitoring or alerting. The lists of supported and unsupported models below are not comprehensive but should provide some guidance; in general, if a model is in the same family as one on the list, the support should be the same.

Supported Models:

A200
A220
A250
A300
A320
A400
A700
A700s
A800
A900
C190
FAS2240-4
FAS2552
FAS2650
FAS2720
FAS2750
FAS8200
FAS8300
FAS8700
FAS9000

Unsupported Models:

FAS2620
FAS3250
FAS3270
FAS500f
FAS6280
FAS/AFF 8020
FAS/AFF 8040
FAS/AFF 8060
FAS/AFF 8080

Troubleshooting

Some things to try if you encounter problems with this data collector:

Problem:	Try this:
When attempting to create an ONTAP REST data collector, an error like the following is seen: Configuration: 10.193.70.14: ONTAP rest API at 10.193.70.14 is not available: 10.193.70.14 failed to GET /api/cluster: 400 Bad Request	This is likely due to an oldeer ONTAP array) for example, ONTAP 9.6) which has no REST API capabilities. ONTAP 9.14.1 is the minimum ONTAP version supported by the ONTAP REST collector. "400 Bad Request" responses should be expected on pre-REST ONTAP releases. For ONTAP versions that do support REST but are not 9.14.1 or later, you may see the following simillar message: Configuration: 10.193.98.84: ONTAP rest API at 10.193.98.84 is not available: 10.193.98.84: ONTAP rest API at 10.193.98.84 is available: cheryl5-cluster-2 9.10.1 a3cb3247-3d3c-11ee-8ff3-005056b364a7 but is not of minimum version 9.14.1.
I see empty or "0" metrics where the ONTAP ontapi collector shows data.	ONTAP REST does not report metrics that are used internally on the ONTAP system only. For example, system aggregates will not be collected by ONTAP REST, only SVM's of type "data" will be collected. Other examples of ONTAP REST metrics that may report zero or empty data: InternalVolumes: REST no longer reports vol0. Aggregates: REST no longer reports aggr0. Storage: most metrics are a rollup of the Internal Volume metrics, and will be impacted by the above. Storage Virtual Machines: REST no longer reports SVM's of type other than 'data' (e.g. 'cluster', 'mgmt', 'node'). You may also notice a change in the appearance of graphs that do have data, due to the change in default performance polling period from 15 minutes to 5 minutes. More frequent polling means more data points to plot.

Additional information may be found from the Support page or in the Data Collector Support Matrix.

NetApp ONTAP Data Management Software data collector

This data collector acquires inventory and performance data from storage systems running ONTAP using read-only API calls from an ONTAP account. This data collector also creates a record in the cluster application registry to accelerate support.

Terminology

Data Infrastructure Insights acquires inventory and performance data from the ONTAP data collector. For each asset type acquired, the most common terminology used for the asset is shown. When viewing or troubleshooting this data collector, keep the following terminology in mind:

Vendor/Model Term	Data Infrastructure Insights Term
Disk	Disk
Raid Group	Disk Group
Cluster	Storage
Node	Storage Node
Aggregate	Storage Pool
LUN	Volume
Volume	Internal Volume

ONTAP Data Management Terminology

The following terms apply to objects or references that you might find on ONTAP Data Management storage asset landing pages. Many of these terms apply to other data collectors as well.

Storage

- Model – A comma-delimited list of the unique, discrete node model names within this cluster. If all the nodes in the clusters are the same model type, just one model name will appear.
- Vendor – same Vendor name you would see if you were configuring a new data source.
- Serial number – The array serial number. On cluster architecture storage systems like ONTAP Data Management, this serial number may be less useful than the individual “Storage Nodes” serial numbers.
- IP – generally will be the IP(s) or hostname(s) as configured in the data source.
- Microcode version – firmware.
- Raw Capacity – base 2 summation of all the physical disks in the system, regardless of their role.
- Latency – a representation of what the host facing workloads are experiencing, across both reads and writes. Ideally, Data Infrastructure Insights is sourcing this value directly, but this is often not the case. In lieu of the array offering this up, Data Infrastructure Insights is generally performing an IOPs-weighted calculation derived from the individual internal volumes’ statistics.
- Throughput – aggregated from internal volumes.
  Management – this may contain a hyperlink for the management interface of the device. Created programmatically by the Data Infrastructure Insights data source as part of inventory reporting.

Storage Pool

- Storage – what storage array this pool lives on. Mandatory.
- Type – a descriptive value from a list of an enumerated list of possibilities. Most commonly will be “Aggregate” or “RAID Group””.
- Node – if this storage array’s architecture is such that pools belong to a specific storage node, its name will be seen here as a hyperlink to its own landing page.
- Uses Flash Pool – Yes/No value – does this SATA/SAS based pool have SSDs used for caching acceleration?
- Redundancy – RAID level or protection scheme. RAID_DP is dual parity, RAID_TP is triple parity.
- Capacity – the values here are the logical used, usable capacity and the logical total capacity, and the percentage used across these.

- Over-committed capacity – If by using efficiency technologies you have allocated a sum total of volume or internal volume capacities larger than the logical capacity of the storage pool, the percentage value here will be greater than 0%.
- Snapshot – snapshot capacities used and total, if your storage pool architecture dedicates part of its capacity to segments areas exclusively for snapshots. ONTAP in MetroCluster configurations are likely to exhibit this, while other ONTAP configurations are less so.
- Utilization – a percentage value showing the highest disk busy percentage of any disk contributing capacity to this storage pool. Disk utilization does not necessarily have a strong correlation with array performance – utilization may be high due to disk rebuilds, deduplication activities, etc in the absence of host driven workloads. Also, many arrays' replication implementations may drive disk utilization while not showing as internal volume or volume workload.
- IOPS – the sum IOPs of all the disks contributing capacity to this storage pool.
  Throughput – the sum throughput of all the disks contributing capacity to this storage pool.

Storage Node

- Storage – what storage array this node is part of. Mandatory.
- HA Partner – on platforms where a node will fail over to one and only one other node, it will generally be seen here.
- State – health of the node. Only available when the array is healthy enough to be inventoried by a data source.
- Model – model name of the node.
- Version – version name of the device.
- Serial number – The node serial number.
- Memory – base 2 memory if available.
- Utilization – On ONTAP, this is a controller stress index from a proprietary algorithm. With every performance poll, a number between 0 and 100% will be reported that is the higher of either WAFL disk contention, or average CPU utilization. If you observe sustained values > 50%, that is indicative of undersizing – potentially a controller/node not large enough or not enough spinning disks to absorb the write workload.
- IOPS – Derived directly from ONTAP ZAPI calls on the node object.
- Latency – Derived directly from ONTAP ZAPI calls on the node object.
- Throughput – Derived directly from ONTAP ZAPI calls on the node object.
- Processors – CPU count.

Requirements

The following are requirements to configure and use this data collector:

- You must have access to an Administrator account configured for read-only API calls.
- Account details include username and password.
- Port requirements: 80 or 443
- Account permissions:
  - Read only role name to ontapi application to the default Vserver
  - You may require additional optional write permissions. See the Note About Permissions below.

- ONTAP License requirements:
  - FCP license and mapped/masked volumes required for fibre-channel discovery

Permission Requirements for Collecting ONTAP Switch Metrics

Data Infrastructure Insights has the ability to collect ONTAP cluster switch data as an option in the collector's Advanced Configuration settings. In addition to enabling this on the Data Infrastructure Insights collector, you must also **configure the ONTAP system** itself to provide switch information, and ensure the correct permissions are set, in order to allow the switch data to be sent to Data Infrastructure Insights.

Configuration

Field	Description
NetApp Management IP	IP address or fully-qualified domain name of the NetApp cluster
User Name	User name for NetApp cluster
Password	Password for NetApp cluster

Advanced configuration

Field	Description
Connection type	Choose HTTP (default port 80) or HTTPS (default port 443). The default is HTTPS
Override Communication Port	Specify a different port if you do not want to use the default
Inventory Poll Interval (min)	Default is 60 minutes.
For TLS for HTTPS	Only allow TLS as protocol when using HTTPS
Automatically Lookup Netgroups	Enable the automatic netgroup lookups for export policy rules
Netgroup Expansion	Netgroup Expansion Strategy. Choose *file* or *shell*. The default is *shell*.
HTTP read timeout seconds	Default is 30
Force responses as UTF-8	Forces data collector code to interpret responses from the CLI as being in UTF-8
Performance Poll Interval (sec)	Default is 900 seconds.
Advanced Counter Data Collection	Enable ONTAP integration. Select this to include ONTAP Advanced Counter data in polls. Choose the desired counters from the list.

Field	Description
Cluster Switch Metrics	Allow Data Infrastructure Insights to collect cluster switch data. Note that in addition to enabling this on the Data Infrastructure Insights side, you must also configure the ONTAP system to provide switch information, and ensure the correct permissions are set, in order to allow the switch data to be sent to Data Infrastructure Insights. See "A Note About Permissions" below.

ONTAP Power Metrics

Several ONTAP models provide power metrics for Data Infrastructure Insights that can be used for monitoring or alerting.

> These lists are not comprehensive, and are subject to change. In general, if a model is in the same family as one on the list, the support should be the same but is not guaranteed to be. If you are unsure whether your model supports power metrics, contact ONTAP support.

Supported Models:

A200
A220
A250
A300
A320
A400
A700
A700s
A800
A900
C190
FAS2240-4
FAS2552
FAS2650
FAS2720
FAS2750
FAS8200
FAS8300
FAS8700
FAS9000

Unsupported Models:

FAS2620
FAS3250
FAS3270
FAS500f
FAS6280
FAS/AFF 8020
FAS/AFF 8040
FAS/AFF 8060
FAS/AFF 8080

A Note About Permissions

Since a number of Data Infrastructure Insights' ONTAP dashboards rely on advanced ONTAP counters, you must enable **Advanced Counter Data Collection** in the data collector Advanced Configuration section.

You should also ensure that write permission to the ONTAP API is enabled. This typically requires an account at the cluster level with the necessary permissions.

To create a local account for Data Infrastructure Insights at the cluster level, log in to ONTAP with the Cluster management Administrator username/password, and execute the following commands on the ONTAP server:

1. Before you begin, you must be signed in to ONTAP with an *Administrator* account, and *diagnostic-level commands* must be enabled.
2. Create a read-only role using the following commands.

```
security login role create -role ci_readonly -cmddirname DEFAULT -access
readonly
security login role create -role ci_readonly -cmddirname security
-access readonly
security login role create -role ci_readonly -access all -cmddirname
{cluster application-record create}
```

3. Create the read-only user using the following command. Once you have executed the create command, you will be prompted to enter a password for this user.

```
security login create -username ci_user -application ontapi
-authentication-method password -role ci_readonly
```

If AD/LDAP account is used, the command should be

```
security login create -user-or-group-name DOMAIN\aduser/adgroup
-application ontapi -authentication-method domain -role ci_readonly
```

If you are collecting cluster switch data:

```
security login rest-role create -role ci_readonly_rest -api
/api/network/ethernet -access readonly
security login create -user-or-group-name ci_user -application http
-authmethod password -role ci_readonly_rest
```

The resulting role and user login will look something like the following. Your actual output may vary:

```
Role Command/ Access
Vserver Name Directory Query Level
---------- ------------ --------- ------------------ --------
cluster1 ci_readonly DEFAULT read only
cluster1 ci_readonly security readonly
```

```
cluster1:security login> show
Vserver: cluster1
Authentication Acct
UserName    Application   Method      Role Name      Locked
---------   -------      ----------- -------------- --------
ci_user     ontapi      password    ci_readonly   no
```

If ONTAP access control is not set correctly, then Data Infrastructure Insights REST calls may fail, resulting in gaps in data for the device. For example, if you have enabled it on the Data Infrastructure Insights collector but have not configured the permissions on the ONTAP, acquisition will fail. Additionally, if the role is previously defined on the ONTAP and you are adding the Rest API abilities, ensure that *http* is added to the role.

Troubleshooting

Some things to try if you encounter problems with this data collector:

Inventory

Problem:	Try this:
Receive 401 HTTP response or 13003 ZAPI error code and ZAPI returns “Insufficient privileges” or “not authorized for this command”	Check username and password, and user privileges/permissions.
Cluster version is < 8.1	Cluster minimum supported version is 8.1. Upgrade to minimum supported version.
ZAPI returns "cluster role is not cluster_mgmt LIF"	AU needs to talk to cluster management IP. Check the IP and change to a different IP if necessary
Error: “7 Mode filers are not supported”	This can happen if you use this data collector to discover 7 mode filer. Change IP to point to cdot cluster instead.
ZAPI command fails after retry	AU has communication problem with the cluster. Check network, port number, and IP address. User should also try to run a command from command line from the AU machine.
AU failed to connect to ZAPI via HTTP	Check whether ZAPI port accepts plaintext. If AU tries to send plaintext to an SSL socket, the communication fails.

Problem:	Try this:
Communication fails with SSLException	AU is attempting to send SSL to a plaintext port on a filer. Check whether the ZAPI port accepts SSL, or use a different port.
Additional Connection errors: ZAPI response has error code 13001, “database is not open” ZAPI error code is 60 and response contains “API did not finish on time” ZAPI response contains “initialize_session() returned NULL environment” ZAPI error code is 14007 and response contains “Node is not healthy”	Check network, port number, and IP address. User should also try to run a command from command line from the AU machine.

Performance

Problem:	Try this:
“Failed to collect performance from ZAPI” error	This is usually due to perf stat not running. Try the following command on each node: *> system node systemshell -node * -command “spmctl -h cmd –stop; spmctl -h cmd –exec”*

Additional information may be found from the Support page or in the Data Collector Support Matrix.

NetApp ONTAP REST data collector

This data collector acquires inventory, EMS logs, and performance data from storage systems running ONTAP 9.14.1 and higher using REST API calls. For ONTAP systems on earlier releases, use the ZAPI-based "NetApp ONTAP Data Management Software" collector type.

The ONTAP REST collector may be used as a replacement for the previous ONTAPI-based collector. As such, there may be differences in the metrics that are collected or reported. For more information about the differences between ONTAPI and REST, see the ONTAP 9.14.1 ONTAPI-to-REST mapping documentation.

Requirements

The following are requirements to configure and use this data collector:

- You must have access to a user account with the required level of access. Note that Admin permissions are required if creating a new REST user/role.
  - Functionally, Data Infrastructure Insights primarily makes read requests, but some write permissions are required for Data Infrastructure Insights to register with the ONTAP array. See the *Note About Permissions* immediately below.

- ONTAP version 9.14.1 or higher.
- Port requirements: 443
- Note that Fpolicy metrics are not supported in the ONTAP REST collector (Workload Security data collectors use FPolicy servers, and FPolicy metrics allow you to monitor Workload Security data collection activity). To support Fpolicy metrics, use the ONTAP Data Management collector.

A Note About Permissions

Since a number of Data Infrastructure Insights' ONTAP dashboards rely on advanced ONTAP counters, you should keep **Enable Advanced Counter Data Collection** enabled in the data collector Advanced Configuration section.

To create a local account for Data Infrastructure Insights at the cluster level, log in to ONTAP with the Cluster management Administrator username/password, and execute the following commands on the ONTAP server:

1. Before you begin, you must be signed in to ONTAP with an *Administrator* account, and *diagnostic-level commands* must be enabled.
2. Retrieve the name of the vserver that is of type *admin*. You will using this name in subsequent commands.

```
vserver show -type admin
```

3. Create a role using the following commands:

```
security login rest-role create -role {role name} -api /api -access
readonly
security login rest-role create -role {role name} -api
/api/cluster/agents -access all
vserver services web access create -name spi -role {role name} -vserver
{vserver name as retrieved above}
```

4. Create the read-only user using the following command. Once you have executed the create command, you will be prompted to enter a password for this user. Note that in the following command we show the role set to *ci_readonly*. If you create a role with different name in step 3 above, use that custom role name instead.

```
security login create -user-or-group-name {username} -application http
-authentication-method password -role {role name}
```

If AD/LDAP account is used, the command should be

```
security login create -user-or-group-name DOMAIN\aduser/adgroup
-application http -authentication-method domain -role ci_readonly
```

The resulting role and user login will look something like the following. Your actual output may vary:

```
security login rest-role show -vserver <vserver name> -role restRole

               Role                                    Access
Vserver        Name            API                     Level
----------     -------------   --------------------    ------
<vserver name> restRole        /api                    readonly
                               /api/cluster/agents     all
2 entries were displayed.

security login show -vserver <vserver name> -user-or-group-name restUser

Vserver: <vserver name>
                                                                Second
User/Group                 Authentication                 Acct
Authentication
Name           Application Method        Role Name        Locked Method
-------------- ----------- ------------- ---------------- ------
--------------
restUser       http        password      restRole         no     none
```

You can check for spi access if needed:

```
**Vserver:> vserver services web access show -name spi
 Vserver                      Type     Service Name     Role
-------------- -------- ---------------- ----------------
<vserver name >     admin             spi           admin
<vserver name >     admin             spi           csrestrole

2 entries were displayed.**
```

Migration

To migrate from a previous ONTAP (ontapi) data collector to the newer ONTAP REST collector, do the following:

1. Add the REST Collector. It is recommended to enter information for a different user than the one configured for the previous collector. For example, use the user noted in the Permissions section above.
2. Pause the previous collector, so it doesn’t continue to collect data.
3. Let the new REST collector acquire data for at least 30 minutes. Ignore any data during this time that does not appear "normal".
4. After the rest period, you should see your data stabilize as the REST collector continues to to acquire.

You can use this same process to return to the previous collector, should you wish.

Configuration

Field	Description
ONTAP management IP Address	IP address or fully-qualified domain name of the NetApp cluster. Must be Cluster Management IP/FQDN.
ONTAP REST User Name	User name for NetApp cluster
ONTAP REST Password	Password for NetApp cluster

Advanced configuration

Field	Description
Inventory Poll Interval (min)	Default is 60 minutes.
Performance Poll Interval (sec)	Default is 60 seconds.
Advanced Counter Data Collection	Select this to include ONTAP Advanced Counter data in polls. Enabled by default.
Enable EMS Event Collection	Select this to include ONTAP EMS log event data. Enabled by default.
EMS Poll Interval (sec)	Default is 60 seconds.

Terminology

Data Infrastructure Insights acquires inventory, logs and performance data from the ONTAP data collector. For each asset type acquired, the most common terminology used for the asset is shown. When viewing or troubleshooting this data collector, keep the following terminology in mind:

Vendor/Model Term	Data Infrastructure Insights Term
Disk	Disk
Raid Group	Disk Group
Cluster	Storage
Node	Storage Node
Aggregate	Storage Pool
LUN	Volume
Volume	Internal Volume
Storage Virtual Machine/Vserver	Storage Virtual Machine

ONTAP Data Management Terminology

The following terms apply to objects or references that you might find on ONTAP Data Management storage asset landing pages. Many of these terms apply to other data collectors as well.

Storage

- Model – A comma-delimited list of the unique, discrete node model names within this cluster. If all the

nodes in the clusters are the same model type, just one model name will appear.

- Vendor – same Vendor name you would see if you were configuring a new data source.
- Serial number – The array UUID
- IP – generally will be the IP(s) or hostname(s) as configured in the data source.
- Microcode version – firmware.
- Raw Capacity – base 2 summation of all the physical disks in the system, regardless of their role.
- Latency – a representation of what the host facing workloads are experiencing, across both reads and writes. Ideally, Data Infrastructure Insights is sourcing this value directly, but this is often not the case. In lieu of the array offering this up, Data Infrastructure Insights is generally performing an IOPs-weighted calculation derived from the individual internal volumes' statistics.
- Throughput – aggregated from internal volumes.
  Management – this may contain a hyperlink for the management interface of the device. Created programmatically by the Data Infrastructure Insights data source as part of inventory reporting.

Storage Pool

- Storage – what storage array this pool lives on. Mandatory.
- Type – a descriptive value from a list of an enumerated list of possibilities. Most commonly will be "Aggregate" or "RAID Group"".
- Node – if this storage array's architecture is such that pools belong to a specific storage node, its name will be seen here as a hyperlink to its own landing page.
- Uses Flash Pool – Yes/No value – does this SATA/SAS based pool have SSDs used for caching acceleration?
- Redundancy – RAID level or protection scheme. RAID_DP is dual parity, RAID_TP is triple parity.
- Capacity – the values here are the logical used, usable capacity and the logical total capacity, and the percentage used across these.
- Over-committed capacity – If by using efficiency technologies you have allocated a sum total of volume or internal volume capacities larger than the logical capacity of the storage pool, the percentage value here will be greater than 0%.
- Snapshot – snapshot capacities used and total, if your storage pool architecture dedicates part of its capacity to segments areas exclusively for snapshots. ONTAP in MetroCluster configurations are likely to exhibit this, while other ONTAP configurations are less so.
- Utilization – a percentage value showing the highest disk busy percentage of any disk contributing capacity to this storage pool. Disk utilization does not necessarily have a strong correlation with array performance – utilization may be high due to disk rebuilds, deduplication activities, etc in the absence of host driven workloads. Also, many arrays' replication implementations may drive disk utilization while not showing as internal volume or volume workload.
- IOPS – the sum IOPs of all the disks contributing capacity to this storage pool.
  Throughput – the sum throughput of all the disks contributing capacity to this storage pool.

Storage Node

- Storage – what storage array this node is part of. Mandatory.
- HA Partner – on platforms where a node will fail over to one and only one other node, it will generally be seen here.
- State – health of the node. Only available when the array is healthy enough to be inventoried by a data

source.

- Model – model name of the node.
- Version – version name of the device.
- Serial number – The node serial number.
- Memory – base 2 memory if available.
- Utilization – On ONTAP, this is a controller stress index from a proprietary algorithm. With every performance poll, a number between 0 and 100% will be reported that is the higher of either WAFL disk contention, or average CPU utilization. If you observe sustained values > 50%, that is indicative of undersizing – potentially a controller/node not large enough or not enough spinning disks to absorb the write workload.
- IOPS – Derived directly from ONTAP REST calls on the node object.
- Latency – Derived directly from ONTAP REST calls on the node object.
- Throughput – Derived directly from ONTAP REST calls on the node object.
- Processors – CPU count.

ONTAP Power Metrics

Several ONTAP models provide power metrics for Data Infrastructure Insights that can be used for monitoring or alerting. The lists of supported and unsupported models below are not comprehensive but should provide some guidance; in general, if a model is in the same family as one on the list, the support should be the same.

Supported Models:

A200
A220
A250
A300
A320
A400
A700
A700s
A800
A900
C190
FAS2240-4
FAS2552
FAS2650
FAS2720
FAS2750
FAS8200
FAS8300
FAS8700
FAS9000

Unsupported Models:

FAS2620
FAS3250
FAS3270
FAS500f

FAS6280
FAS/AFF 8020
FAS/AFF 8040
FAS/AFF 8060
FAS/AFF 8080

Troubleshooting

Some things to try if you encounter problems with this data collector:

Problem:	Try this:
When attempting to create an ONTAP REST data collector, an error like the following is seen: Configuration: 10.193.70.14: ONTAP rest API at 10.193.70.14 is not available: 10.193.70.14 failed to GET /api/cluster: 400 Bad Request	This is likely due to an oldeer ONTAP array) for example, ONTAP 9.6) which has no REST API capabilities. ONTAP 9.14.1 is the minimum ONTAP version supported by the ONTAP REST collector. "400 Bad Request" responses should be expected on pre-REST ONTAP releases. For ONTAP versions that do support REST but are not 9.14.1 or later, you may see the following simillar message: Configuration: 10.193.98.84: ONTAP rest API at 10.193.98.84 is not available: 10.193.98.84: ONTAP rest API at 10.193.98.84 is available: cheryl5-cluster-2 9.10.1 a3cb3247-3d3c-11ee-8ff3-005056b364a7 but is not of minimum version 9.14.1.
I see empty or "0" metrics where the ONTAP ontapi collector shows data.	ONTAP REST does not report metrics that are used internally on the ONTAP system only. For example, system aggregates will not be collected by ONTAP REST, only SVM's of type "data" will be collected. Other examples of ONTAP REST metrics that may report zero or empty data: InternalVolumes: REST no longer reports vol0. Aggregates: REST no longer reports aggr0. Storage: most metrics are a rollup of the Internal Volume metrics, and will be impacted by the above. Storage Virtual Machines: REST no longer reports SVM's of type other than 'data' (e.g. 'cluster', 'mgmt', 'node'). You may also notice a change in the appearance of graphs that do have data, due to the change in default performance polling period from 15 minutes to 5 minutes. More frequent polling means more data points to plot.

Additional information may be found from the Support page or in the Data Collector Support Matrix.

NetApp Data ONTAP operating in 7-Mode data collector

For storage systems using Data ONTAP software operating in 7-Mode, you use the 7-mode data collector, which uses the CLI to obtain capacity and performance data.

Terminology

Data Infrastructure Insights acquires the following inventory information from the NetApp 7-mode data collector. For each asset type acquired, the most common terminology used for this asset is shown. When viewing or troubleshooting this data collector, keep the following terminology in mind:

This data collector is deprecated.

Vendor/Model Term	Data Infrastructure Insights Term
Disk	Disk
Raid Group	Disk Group
Filer	Storage
Filer	Storage Node
Aggregate	Storage Pool
LUN	Volume
Volume	Internal Volume

Note: These are common terminology mappings only and might not represent every case for this data collector.

Requirements

You need the following to configure and use this data collector:

- IP addresses of the FAS storage controller and partner.
- Port 443
- A custom admin level username and password for controller and partner controller with the following role capabilities for 7-Mode:
  - "api-*": Use this to allow OnCommand Insight to execute all NetApp storage API commands.
  - "login-http-admin": Use this to allow OnCommand Insight to connect to the NetApp storage via HTTP.
  - "security-api-vfiler": Use this to allow OnCommand Insight to execute NetApp storage API commands to retrieve vFiler unit information.
  - "cli-options": Use this to read storage system options.
  - "cli-lun": Access these commands for managing LUNs. Displays the status (LUN path, size, online/offline state, and shared state) of the given LUN or class of LUNs.
  - "cli-df": Use this to display free disk space.
  - "cli-ifconfig": Use this to display interfaces and IP addresses.

Configuration

Field	Description
Address of storage system	IP address or fully-qualified domain name for the NetApp storage system
User Name	User name for the NetApp storage system
Password	Password for the NetApp storage system
Address of HA Partner in Cluster	IP address or fully-qualified domain name for the HA Partner
User Name of HA Partner in Cluster	User name for the HA partner
Password of HA Partner Filer in Cluster	Password for the HA Partner

Advanced configuration

Field	Description
Inventory Poll Interval (min)	Interval between inventory polls. The default is 20 minutes.
Connection Type	HTTPS or HTTP, also displays the default port
Override Connection Port	If blank, use the default port in the Connection Type field, otherwise enter the connection port to use
Performance Poll Interval (sec)	Interval between performance polls. The default is 300 seconds.

Storage systems connection

As an alternative to using the default administrative user for this data collector, you can configure a user with administrative rights directly on the NetApp storage systems so that this data collector can acquire data from NetApp storage systems.

Connecting to NetApp storage systems requires that the user, who is specified when acquiring the main pfiler (on which the storage system exist), meet the following conditions:

- The user must be on vfiler0 (root filer/pfiler).

  Storage systems are acquired when acquiring the main pfiler.

- The following commands define the user role capabilities:
  - "api-*": Use this to allow Data Infrastructure Insights to execute all NetApp storage API commands.

    This command is required to use the ZAPI.

  - "login-http-admin": Use this to allow Data Infrastructure Insights to connect to the NetApp storage via HTTP. This command is required to use the ZAPI.
  - "security-api-vfiler": Use this to allow Data Infrastructure Insights to execute NetApp storage API commands to retrieve vFiler unit information.
  - "cli-options": For "options" command and used for partner IP and enabled licenses.

- "cli-lun": Access these command for managing LUNs. Displays the status (LUN path, size, online/offline state, and shared state) of the given LUN or class of LUNs.
- "cli-df": For "df -s", "df -r", "df -A -r" commands and used to display free space.
- "cli-ifconfig": For "ifconfig -a" command and used for getting filer IP address.
- "cli-rdfile": For "rdfile /etc/netgroup" command and used for getting netgroups.
- "cli-date": For "date" command and used to get full date for getting Snapshot copies.
- "cli-snap": For "snap list" command and used for getting Snapshot copies.

If cli-date or cli-snap permissions are not provided, acquisition can finish, but Snapshot copies are not reported.

To acquire a 7-Mode data source successfully and generate no warnings on the storage system, you should use one of the following command strings to define your user roles. The second string listed here is a streamlined version of the first:

- login-http-admin,api-*,security-api-vfile,cli-rdfile,cli-options,cli-df,cli-lun,cli-ifconfig,cli-date,cli-snap,_
- login-http-admin,api-* ,security-api-vfile,cli-

Troubleshooting

Some things to try if you encounter problems with this data collector:

Inventory

Problem:	Try this:
Receive 401 HTTP response or 13003 ZAPI error code and ZAPI returns “Insufficient privileges” or “not authorized for this command”	Check username and password, and user privileges/permissions.
“Failed to execute command” error	Check whether the user has the following permission on the device: • api-* • cli-date • cli-df • cli-ifconfig • cli-lun • cli-operations • cli-rdfile • cli-snap • login-http-admin • security-api-vfiler Also check if the ONTAP version is supported by Data Infrastructure Insights and verify if the credentials used match device credentials
Cluster version is < 8.1	Cluster minimum supported version is 8.1. Upgrade to minimum supported version.
ZAPI returns "cluster role is not cluster_mgmt LIF"	AU needs to talk to cluster management IP. Check the IP and change to a different IP if necessary

Problem:	Try this:
Error: “7 Mode filers are not supported”	This can happen if you use this data collector to discover 7 mode filer. Change IP to point to cdot filer instead.
ZAPI command fails after retry	AU has communication problem with the cluster. Check network, port number, and IP address. User should also try to run a command from command line from the AU machine.
AU failed to connect to ZAPI	Check IP/port connectivity and assert ZAPI configuration.
AU failed to connect to ZAPI via HTTP	Check whether ZAPI port accepts plaintext. If AU tries to send plaintext to an SSL socket, the communication fails.
Communication fails with SSLException	AU is attempting to send SSL to a plaintext port on a filer. Check whether the ZAPI port accepts SSL, or use a different port.
Additional Connection errors: ZAPI response has error code 13001, “database is not open” ZAPI error code is 60 and response contains “API did not finish on time” ZAPI response contains “initialize_session() returned NULL environment” ZAPI error code is 14007 and response contains “Node is not healthy”	Check network, port number, and IP address. User should also try to run a command from command line from the AU machine.
Socket timeout error with ZAPI	Check filer connectivity and/or increase timeout.
“C Mode clusters are not supported by the 7 Mode data source” error	Check IP and change the IP to a 7 Mode cluster.
“Failed to connect to vFiler” error	Check that the acquiring user capabilities include the following at a minimum: api-* security-api-vfiler login-http-admin Confirm that filer is running minimum ONTAPI version 1.7.

Additional information may be found from the Support page or in the Data Collector Support Matrix.

NetApp E-Series Legacy Santricity API data collector

The NetApp E-Series Legacy Santricity API data collector gathers inventory and performance data. The collector supports firmware 7.x+ using the same configurations and reporting the same data.

Terminology

Cloud insight acquires the following inventory information from the NetApp E-Series data collector. For each asset type acquired, the most common terminology used for this asset is shown. When viewing or troubleshooting this data collector, keep the following terminology in mind:

Vendor/Model Term	Data Infrastructure Insights Term
Disk	Disk
Volume Group	Disk Group
Storage Array	Storage
Controller	Storage Node
Volume Group	Storage Pool
Volume	Volume

Note: These are common terminology mappings only and might not represent every case for this data collector.

E-Series Terminology (Landing Page)

The following terms apply to objects or references that you might find on NetApp E-Series asset landing pages. Many of these terms apply to other data collectors as well.

Storage

- Model – model name of the device.
- Vendor – same Vendor name you would see if you were configuring a new datasource
- Serial number – The array serial number. On cluster architecture storage systems like NetApp Clustered Data Ontap, this serial number may be less useful than the individual “Storage Nodes” serial numbers
- IP – generally will be the IP(s) or hostname(s) as configured in the data source
- Microcode version – firmware
- Raw Capacity – base 2 summation of all the physical disks in the system, regardless of their role
- Latency – a representation of what the host facing workloads are experiencing, across both reads and writes. Ideally, Data Infrastructure Insights is sourcing this value directly, but this is often not the case. In lieu of the array offering this up, Data Infrastructure Insights is generally performing an IOPs-weighted calculation derived from the individual volumes’ statistics.
- Throughput – the array’s total host facing throughput. Ideally sourced directly from the array, if unavailable, Data Infrastructure Insights is summing the volumes’ throughput to derive this value
- Management – this may contain a hyperlink for the management interface of the device. Created programmatically by the Data Infrastructure Insights datasource as part of inventory reporting

Storage Pool

- Storage – what storage array this pool lives on. Mandatory
- Type – a descriptive value from a list of an enumerated list of possibilities. Most commonly will be “Thin Provisioning” or “RAID Group”

- Node – if this storage array's architecture is such that pools belong to a specific storage node, its name will be seen here as a hyperlink to its own landing page
- Uses Flash Pool – Yes/No value
- Redundancy – RAID level or protection scheme. E-Series reports "RAID 7" for DDP pools
- Capacity – the values here are the logical used, usable capacity and the logical total capacity, and the percentage used across these. These value both include E-Series "preservation" capacity, resulting both in numbers and the percentage being higher than what the E-Series own user interface may show
- Over-committed capacity – If via efficiency technologies you have allocated a sum total of volume or internal volume capacities larger than the logical capacity of the storage pool, the percentage value here will be greater than 0%.
- Snapshot – snapshot capacities used and total, if your storage pool architecture dedicates part of its capacity to segments areas exclusively for snapshots
- Utilization – a percentage value showing the highest disk busy percentage of any disk contributing capacity to this storage pool. Disk utilization does not necessarily have a strong correlation with array performance – utilization may be high due to disk rebuilds, deduplication activities, etc in the absence of host driven workloads. Also, many arrays' replication implementations may drive disk utilization while not showing as volume workload.
- IOPS – the sum IOPs of all the disks contributing capacity to this storage pool. If disk IOPs is not available on a given platform, this value will be sourced from the sum of volume IOPs for all the volumes sitting on this storage pool
- Throughput – the sum throughput of all the disks contributing capacity to this storage pool. If disk throughput is not available on a given platform, this value will be sourced from the sum of volume throughout for all the volumes sitting on this storage pool

Storage Node

- Storage – what storage array this node is part of. Mandatory
- HA Partner – on platforms where a node will fail over to one and only one other node, it will generally be seen here
- State – health of the node. Only available when the array is healthy enough to be inventoried by a data source
- Model – model name of the node
- Version – version name of the device.
- Serial number – The node serial number
- Memory – base 2 memory if available
- Utilization – Generally a CPU utilization number, or in the case of NetApp Ontap, a controller stress index. Utilization is not currently available for NetApp E-Series
- IOPS – a number representing the host driven IOPs on this controller. Ideally sourced directly from the array, if unavailable, it will be calculated by summing all the IOPs for volumes that belong exclusively to this node.
- Latency – a number representing the typical host latency or response time on this controller. Ideally sourced directly from the array, if unavailable, it will be calculated by performing an IOPs weighted calculation from volumes that belong exclusively to this node.
- Throughput – a number representing the host driven throughput on this controller. Ideally sourced directly from the array, if unavailable, it will be calculated by summing all the throughput for volumes that belong exclusively to this node.

- Processors – CPU count

Requirements

- The IP address of each controller on the array
- Port requirement 2463

Configuration

Field	Description
Comma-separated list of Array SANtricity Controller IPs	IP addresses and/or fully-qualified domain names for the array controllers

Advanced configuration

Field	Description
Inventory Poll Interval (min)	Default is 30 minutes
Performance Poll Interval up to 3600 seconds	Default is 300 seconds

Troubleshooting

Additional information on this data collector may be found from the Support page or in the Data Collector Support Matrix.

NetApp E-Series REST data collector

The NetApp E-Series REST data collector gathers inventory, logs and performance data. The collector supports firmware 7.x+ using the same configurations and reporting the same data. The REST collector monitors the encryption status of storage pools, disks and volumes, storage node CPU Utilization as well as logs functionality not provided in the legacy SANtricity E-Series collector.

Terminology

Data Infrastructure Insights acquires the following inventory information from the NetApp E-Series, using REST. For each asset type acquired, the most common terminology used for this asset is shown. When viewing or troubleshooting this data collector, keep the following terminology in mind:

Vendor/Model Term	Data Infrastructure Insights Term
Disk	Disk
Volume Group	Disk Group
Storage Array	Storage
Controller	Storage Node
Volume Group	Storage Pool
Volume	Volume

Note: These are common terminology mappings only and might not represent every case for this data collector.

Requirements

- The IP address of each controller on the array
- This collector supports only E-Series model arrays with **native REST API capabilities**. The E-Series org ships an off-array, installable REST API distribution for older E-Series arrays: this collector does not support that scenario. Users with older arrays should continue to use Data Infrastructure Insights' E-Series SANtricity API collector.
- The "E-Series Controller IP Addresses" field supports a comma delimited string of 2 IP/hostnames; the collector will intelligently try the second IP/hostname if the first is inaccessible.
- HTTPS port: default is 8443.

Configuration

Field	Description
E-Series Controller IP Addresses	comma-separated IP addresses and/or fully-qualified domain names for the array controllers

Advanced configuration

Field	Description
Inventory Poll Interval (min)	Default is 30 minutes
Performance Poll Interval up to 3600 seconds	Default is 300 seconds

E-Series Terminology (Landing Page)

The following terms apply to objects or references that you might find on NetApp E-Series asset landing pages. Many of these terms apply to other data collectors as well.

Storage

- Model – model name of the device.
- Vendor – same Vendor name you would see if you were configuring a new datasource
- Serial number – The array serial number. On cluster architecture storage systems like NetApp Clustered Data Ontap, this serial number may be less useful than the individual “Storage Nodes” serial numbers
- IP – generally will be the IP(s) or hostname(s) as configured in the data source
- Microcode version – firmware
- Raw Capacity – base 2 summation of all the physical disks in the system, regardless of their role
- Latency – a representation of what the host facing workloads are experiencing, across both reads and writes. Ideally, Data Infrastructure Insights is sourcing this value directly, but this is often not the case. In lieu of the array offering this up, Data Infrastructure Insights is generally performing an IOPs-weighted calculation derived from the individual volumes’ statistics.
- Throughput – the array’s total host facing throughput. Ideally sourced directly from the array, if unavailable, Data Infrastructure Insights is summing the volumes’ throughput to derive this value
- Management – this may contain a hyperlink for the management interface of the device. Created

programmatically by the Data Infrastructure Insights datasource as part of inventory reporting

Storage Pool

- Storage – what storage array this pool lives on. Mandatory
- Type – a descriptive value from a list of an enumerated list of possibilities. Most commonly will be “Thin Provisioning” or “RAID Group”
- Node – if this storage array’s architecture is such that pools belong to a specific storage node, its name will be seen here as a hyperlink to its own landing page
- Uses Flash Pool – Yes/No value
- Redundancy – RAID level or protection scheme. E-Series reports “RAID 7” for DDP pools
- Capacity – the values here are the logical used, usable capacity and the logical total capacity, and the percentage used across these. These value both include E-Series “preservation” capacity, resulting both in numbers and the percentage being higher than what the E-Series own user interface may show
- Over-committed capacity – If via efficiency technologies you have allocated a sum total of volume or internal volume capacities larger than the logical capacity of the storage pool, the percentage value here will be greater than 0%.
- Snapshot – snapshot capacities used and total, if your storage pool architecture dedicates part of its capacity to segments areas exclusively for snapshots
- Utilization – a percentage value showing the highest disk busy percentage of any disk contributing capacity to this storage pool. Disk utilization does not necessarily have a strong correlation with array performance – utilization may be high due to disk rebuilds, deduplication activities, etc in the absence of host driven workloads. Also, many arrays’ replication implementations may drive disk utilization while not showing as volume workload.
- IOPS – the sum IOPs of all the disks contributing capacity to this storage pool. If disk IOPs is not available on a given platform, this value will be sourced from the sum of volume IOPs for all the volumes sitting on this storage pool
- Throughput – the sum throughput of all the disks contributing capacity to this storage pool. If disk throughput is not available on a given platform, this value will be sourced from the sum of volume throughout for all the volumes sitting on this storage pool

Storage Node

- Storage – what storage array this node is part of. Mandatory
- HA Partner – on platforms where a node will fail over to one and only one other node, it will generally be seen here
- State – health of the node. Only available when the array is healthy enough to be inventoried by a data source
- Model – model name of the node
- Version – version name of the device.
- Serial number – The node serial number
- Memory – base 2 memory if available
- Utilization – Generally a CPU utilization number, or in the case of NetApp Ontap, a controller stress index. Utilization is not currently available for NetApp E-Series
- IOPS – a number representing the host driven IOPs on this controller. Ideally sourced directly from the

array, if unavailable, it will be calculated by summing all the IOPs for volumes that belong exclusively to this node.

- Latency – a number representing the typical host latency or response time on this controller. Ideally sourced directly from the array, if unavailable, it will be calculated by performing an IOPs weighted calculation from volumes that belong exclusively to this node.
- Throughput – a number representing the host driven throughput on this controller. Ideally sourced directly from the array, if unavailable, it will be calculated by summing all the throughput for volumes that belong exclusively to this node.
- Processors – CPU count

Troubleshooting

Additional information on this data collector may be found from the Support page or in the Data Collector Support Matrix.

Configuring the NetApp HCI Management server data collector

The NetApp HCI Management server data collector collects NetApp HCI Host information and requires read-only privileges on all objects within the Management server.

This data collector acquires from the **NetApp HCI Management server only**. To collect data from the storage system, you must also configure the NetApp SolidFire data collector.

Terminology

Data Infrastructure Insights acquires the following inventory information from this data collector. For each asset type acquired, the most common terminology used for the asset is shown. When viewing or troubleshooting this data collector, keep the following terminology in mind:

Vendor/Model Term	Data Infrastructure Insights Term
Virtual disk	Disk
Host	Host
Virtual machine	Virtual machine
Data store	Data store
LUN	Volume
Fibre channel port	Port

These are common terminology mappings only and might not represent every case for this data collector.

Requirements

The following information is required to configure this data collector:

- IP address of the NetApp HCI Management server
- Read-only username and password for the NetApp HCI Management server
- Read only privileges on all objects in the NetApp HCI Management server.
- SDK access on the NetApp HCI Management server – normally already set up.

- Port requirements: http-80 https-443
- Validate access:
  - Log into the NetApp HCI Management server using above username and password
  - Verify SDK enabled: telnet <vc_ip> 443

Setup and connection

Field	Description
Name	Unique name for the data collector
Acquisition unit	Name of acquisition unit

Configuration

Field	Description
NetApp HCI Storage Cluster MVIP	Management Virtual IP Address
SolidFire Management Node (mNode)	Management Node IP Address
User name	User name used to access the NetApp HCI Management server
Password	Password used to access the NetApp HCI Management server
VCenter User Name	User name for VCenter
VCenter Password	Password for VCenter

Advanced configuration

In the advanced configuration screen, check the **VM Performance** box to collect performance data. Inventory collection is enabled by default.
The following fields can be configured:

Field	Description
Inventory poll interval (min)	Deafult is 20
Filter VMs by	Select CLUSTER, DATACENTER, or ESX HOST
Choose 'Exclude' or 'Include' to Specify a List	Specify Whether to Include or Exclude VMs
Filter Device List	List of VMs to filter (comma separated, or semicolon separated if comma is used in the value) for for Filtering by ESX_HOST, CLUSTER, and DATACENTER Only
Performance poll interval (sec)	Default is 300

Troubleshooting

Some things to try if you encounter problems with this data collector:

Inventory

Problem:	Try this:
Error: Include list to filter VMs cannot be empty	If Include List is selected, please list valid DataCenter, Cluster, or Host names to filter VMs
Error: Failed to instantiate a connection to VirtualCenter at IP	Possible solutions: * Verify credentials and IP address entered. * Try to communicate with Virtual Center using Infrastructure Client. * Try to communicate with Virtual Center using Managed Object Browser (e.g MOB).
Error: VirtualCenter at IP has non-conform certificate that JVM requires	Possible solutions: * Recommended: Re-generate certificate for Virtual Center by using stronger (e.g. 1024-bit) RSA key. * Not Recommended: Modify the JVM java.security configuration to leverage the constraint jdk.certpath.disabledAlgorithms to allow 512-bit RSA key. See JDK 7 update 40 release notes at "http://www.oracle.com/technetwork/java/javase/7u40-relnotes-2004172.html"

Additional information may be found from the Support page or in the Data Collector Support Matrix.

NetApp SolidFire All-Flash Array data collector

The NetApp SolidFire All-Flash Array data collector supports inventory and performance collection from both iSCSI and Fibre Channel SolidFire configurations.

The SolidFire data collector utilizes the SolidFire REST API. The acquisition unit where the data collector resides needs to be able to initiate HTTPS connections to TCP port 443 on the SolidFire cluster management IP address. The data collector needs credentials capable of making REST API queries on the SolidFire cluster.

Terminology

Data Infrastructure Insights acquires the following inventory information from the NetApp SolidFire All-Flash Array data collector. For each asset type acquired by Data Infrastructure Insights, the most common terminology used for this asset is shown. When viewing or troubleshooting this data collector, keep the following terminology in mind:

Vendor/Model Term	Data Infrastructure Insights Term
Drive	Disk
Cluster	Storage
Node	Storage Node
Volume	Volume
Fibre channel port	Port
Volume Access Group, LUN Assignment	Volume Map

Vendor/Model Term	Data Infrastructure Insights Term
iSCSI Session	Volume Mask

Note: These are common terminology mappings only and might not represent every case for this data collector.

Requirements

The following are requirements for configuring this data collector:

- Management Virtual IP Address
- Read-only username and credentials
- Port 443

Configuration

Field	Description
Management Virtual IP Address (MVIP)	Management Virtual IP address of the SolidFire Cluster
User Name	Name used to log into the SolidFire cluster
Password	Password used to log into the SolidFire cluster

Advanced configuration

Field	Description
Connection Type	Choose connection type
Communication Port	Port used for NetApp API
Inventory Poll Interval (min)	Default is 20 minutes
Performance Poll Interval (sec)	Default is 300 seconds

Troubleshooting

When SolidFire reports an error it is displayed in Data Infrastructure Insights as follows:

An error message was received from a SolidFire device while trying to retrieve data. The call was <method> (<parameterString>). The error message from the device was (check the device manual): <message>

Where:

- The <method> is an HTTP method, such as GET or PUT.
- The <parameterString> is a comma separated list of parameters that were included in the REST call.
- The <message> is whatever the device returned as the error message.

Additional information on this Data Collector may be found from the Support page or in the Data Collector Support Matrix.

NetApp StorageGRID data collector

The NetApp StorageGRID data collector supports inventory and performance collection from StorageGRID configurations.

> To provide consistent metering against DII entitlements across all StorageGRID systems regardless of underlying hardware topology and configuration, Data Infrastructure Insights uses the total available capacity (storagegrid_storage_utilization_total_space_bytes), instead of RAW capacity based on physical disk layout.
>
> For customers using the Capacity-based licensing model, StorageGRID is metered as “Object” storage.
>
> For customers using the legacy (MU) licensing model, StorageGRID is metered as secondary storage, at a rate of 40TiB to 1 MU.

Terminology

Data Infrastructure Insights acquires the following inventory information from the NetApp StorageGRID collector. For each asset type acquired, the most common terminology used for this asset is shown. When viewing or troubleshooting this data collector, keep the following terminology in mind:

Vendor/Model Term	Data Infrastructure Insights Term
StorageGRID	Storage
Node	Node
Tenant	Storage Pool
Bucket	Internal Volume

Requirements

The following are requirements for configuring this data source:

- StorageGRID Host IP Address
- A username and password for a user that has had the Metric Query and Tenant Access roles assigned
- Port 443

Configuration

Field	Description
StorageGRID Host IP Address	Management Virtual IP address of the StorageGRID appliance
User Name	Name used to log into the StorageGRID appliance
Password	Password used to log into the StorageGRID appliance

Advanced configuration

Field	Description
Inventory Poll Interval (min)	Default is 60 minutes
performance Poll Interval (sec)	Default is 900 seconds

Single Sign-On (SSO)

The StorageGRID firmware versions have corresponding API versions; 3.0 API and newer versions support single sign-on (SSO) login.

Firmware version	API version	Support single sign on (SSO)
11.1	2	No
11.2	3.0	Yes
11.5	3.3	Yes

Troubleshooting

Additional information on this Data Collector may be found from the Support page or in the Data Collector Support Matrix.

Nutanix NX data collector

Data Infrastructure Insights uses the Nutanix data collector to discover inventory and performance data for Nutanix NX storage systems.

Terminology

Data Infrastructure Insights acquires the following inventory information from the Nutanix data collector. For each asset type acquired by Data Infrastructure Insights, the most common terminology used for this asset is shown. When viewing or troubleshooting this data collector, keep the following terminology in mind:

Vendor/Model Term	Data Infrastructure Insights Term
Storage Pool	Storage Pool
Nutanix Container	Internal Volume
Nutanix Container	File Share
NFS Share	Share

Note: These are common terminology mappings only and might not represent every case for this data collector.

Requirements

- The external data services IP address for the cluster
- Read-only user name and password, unless volume_groups are in use, in which case, Admin user name and password are required
- Port requirement: HTTPS 443

Configuration

Field	Description
Prism External IP Address	The external data services IP address for the cluster
User name	User name for the Admin account
Password	Password for the Admin account

Advanced configuration

Field	Description
TCP port	TCP Port used to connect to Nutanix array. The default is 9440.
Inventory poll interval (min)	Interval between inventory polls. The default is 60 minutes.
Performance poll interval(sec)	Interval between performance polls. The default is 300 seconds.

Troubleshooting

Additional information on this Data Collector may be found from the Support page or in the Data Collector Support Matrix.

Oracle ZFS Storage Appliance data collector

Data Infrastructure Insights uses the Oracle ZFS Storage Appliance data collector to gather inventory and performance data.

Terminology

Data Infrastructure Insights acquires inventory information with the Oracle ZFS data collector. For each asset type acquired by Data Infrastructure Insights, the most common terminology used for this asset is shown. When viewing or troubleshooting this data collector, keep the following terminology in mind:

Vendor/Model Term	Data Infrastructure Insights Term
Disk (SSD)	Disk
Cluster	Storage
Controller	Storage Node
LUN	Volume
LUN Map	Volume Map
Initiator,Target	Volume Mask
Share	Internal Volume

Note: These are common terminology mappings only and might not represent every case for this data source.

Requirements

- Host names for the ZFS Controller-1 and the ZFS Controller-2
- Administrator user name and password
- Port requirement: 215 HTTP/HTTPS

Required Performance metrics

Oracle ZFS appliances give storage administators large amounts of flexibility to capture performance statistics. Data Infrastructure Insights expects you to have *each* controller in a high availability pair configured to capture the following metrics:

- smb2.ops[share]
- nfs3.ops[share]
- nfs4.ops[share]
- nfs4-1.ops[share]

Failure to have the controller capture any or all of these will likely result in Data Infrastructure Insights not having, or underreporting, the workload on the "Internal Volumes".

Configuration

Field	Description
ZFS Controller-1 Hostname	Host name for storage controller 1
ZFS Controller-2 Hostname	Host name for storage controller 2
User name	User name for the storage system administrator user account
Password	Password for the administrator user account

Advanced configuration

Field	Description
Connection Type	HTTPS or HTTP, also displays the default port
Override Connection Port	If blank, use the default port in the Connection Type field, otherwise enter the connection port to use
Inventory poll interval	The default is 60 seconds
Performance Poll Interval (sec)	The default is 300.

Troubleshooting

Some things to try if you encounter problems with this data collector:

Inventory

Problem:	Try this:
"Invalid login credentials"	validate Zfs user account and password
"Request failed with status 404 https://… ..:215/api/access/v1"	Your ZFS array may be too old to have REST API support. AK 2013.1.3.0 was the first REST API-capable ZFS OS release, and not all ZFS appliances can be upgraded to it.
"Configuration error" with error message “REST Service is disabled”	Verify REST service is enabled on this device.
"Configuration error " with error message “User unauthorized for command”	This error is likely due to certain roles (for example, 'advanced_analytics') not being included for the configured user. Applying the Analytics scope for the user with read-only role may remove the error. Follow these steps: 1. On the ZFs system, from the Configuration → Users screen, move your mouse over the role and double click to allow editing 2. Select "Analytics" from the Scope drop down menu. A list of the possible properties appears. 3. Click the top most check box and it will select all three properties. 4. Click the Add button on the right side. 5. Click the Apply button at the top right of the pop-up window. The pop-up window will close.

Additional information may be found from the Support page or in the Data Collector Support Matrix.

Pure Storage FlashArray data collector

Data Infrastructure Insights uses the Pure Storage FlashArray data collector to gather inventory and performance data.

Terminology

For each asset type acquired by Data Infrastructure Insights, the most common terminology used for the asset is shown. When viewing or troubleshooting this data collector, keep the following terminology in mind:

Vendor/Model Term	Data Infrastructure Insights Term
Drive (SSD)	Disk
Array	Storage
Controller	Storage Node
Volume	Volume

Vendor/Model Term	Data Infrastructure Insights Term
LUN Map	Volume Map
Initiator,Target	Volume Mask

Note: These are common terminology mappings only and might not represent every case for this data collector.

Requirements

- Storage system IP address
- User name and password for the Administrator account of the Pure storage system.
- Port requirement: HTTP/HTTPS 80/443

Configuration

Field	Description
FlashArray Host IP Address	IP address of the storage system
User name	User name with admin privileges
Password for the admin privileged account	Password

Advanced configuration

Field	Description
Connection type	Choose HTTP or HTTPS. Also displays the default port.
Override TCP port	If blank, use the default port in the Connection Type field, otherwise enter the connection port to use
Inventory poll interval (min)	The default is 60 minutes
Performance Poll Interval (sec)	The default is 300

Troubleshooting

Some things to try if you encounter problems with this data collector:

Inventory

Problem:	Try this:
"Invalid login credentials" with error messages “Policy doesn’t allow” or “You are not authorized”	Validate Pure user account and password via Pure http interface

Additional information may be found from the Support page or in the Data Collector Support Matrix.

Red Hat Virtualization data collector

Data Infrastructure Insights uses the Red Hat Virtualization data collector to gather

inventory data from virtualized Linux and Microsoft Windows workloads.

Terminology

For each asset type acquired by Data Infrastructure Insights, the most common terminology used for the asset is shown. When viewing or troubleshooting this data collector, keep the following terminology in mind:

Vendor/Model Term	Data Infrastructure Insights Term
Disk	Virtual Disk
Host	Host
Virtual Machine	Virtual Machine
Storage Domain	Data Store
Logical Unit	LUN

Note: These are common terminology mappings only and might not represent every case for this data collector.

Requirements

- IP address of the RHEV server over port 443 via REST API
- Read-only username and password
- RHEV Version 3.0+

Configuration

Field	Description
RHEV Server IP Address	IP address of the storage system
User name	User name with admin privileges
Password for the admin privileged account	Password

Advanced configuration

Field	Description
HTTPS Communication Port	Port used for HTTPS communication to RHEV
Inventory poll interval (min)	The default is 20 minutes.

Troubleshooting

Additional information on this Data Collector may be found from the Support page or in the Data Collector Support Matrix.

Rubrik CDM Data Collector

Data Infrastructure Insights uses the Rubrik data collector to acquire inventory and performance data from Rubrik storage appliances.

Terminology

Data Infrastructure Insights acquires the following inventory information from the Rubrik data collector. For each asset type acquired by Data Infrastructure Insights, the most common terminology used for this asset is shown. When viewing or troubleshooting this data collector, keep the following terminology in mind:

Vendor/Model Term	Data Infrastructure Insights Term
Cluster	Storage, Storage Pool
Node	Storage Node
Disk	Disk

Note: These are common terminology mappings only and might not represent every case for this data source.

Requirements

The following are required to configure this data collector:

- The Data Infrastructure Insights Acquisition Unit will initiate connections to TCP port 443 to Rubrik cluster. One collector per cluster.
- Rubrik cluster IP address.
- User name and password to the cluster.
- Rubrik cluster IP address or hostname.
- For Basic Authentication, a user name and password to the cluster. If you prefer to use Service Account based authentication, you need a Service Account, Secret, and an Organization ID
- Port requirement: HTTPS 443

Configuration

Field	Description
IP	IP address of the Rubrik cluster
Username / Service account	User name for the cluster
Password / Secret	Password for the cluster
Organization ID for the service-account	This needs to be the full string like "Organization:::nnnnnn-nnnn….."

Advanced configuration

Inventory poll interval (min)	The default is 60
Performance Poll Interval (sec)	The default is 300

Troubleshooting

Some things to try if you encounter problems with this data collector:

Inventory

Problem:	Try this:
I received a message that more than one storage is created.	Check that the cluster is configured correctly, and the collector is pointing to a single cluster.
Poll fails with 400 [Bad Request]….Invalid ManagedId….	You have populated the Organization ID field with a value, but the Rubrik cluster does NOT believe that is a valid Organization ID, despite the error message from Rubrik referring to it as a "ManagedId"

Additional information may be found from the Support page or in the Data Collector Support Matrix.

Configuring the VMware VSphere data collector

The data collector for VMware vSphere collects VM guest and ESXi Host performance and configuration information and requires read-only privileges on all objects within vSphere. As of August 2024, the vSphere collector additionally brings in log messages from vSphere environments, and some VMware specific metrics. Please note that Data Infrastructure Insights can only retrieve VMware logs information from vSphere 8.0.1 or higher environments. Similarly, the Vendor Specific Metrics is only supported for vSphere 7+ environments. As such, you may want to disable the logs and/or Vendor Specific Metrics checkbox on a given collector if it is pointed at an older vSphere instance.

Terminology

Data Infrastructure Insights acquires the following inventory information from the VMware vSphere data collector. For each asset type acquired, the most common terminology used for the asset is shown. When viewing or troubleshooting this data collector, keep the following terminology in mind:

Vendor/Model Term	Data Infrastructure Insights Term
Virtual disk	Disk
Host	Host
Virtual machine	Virtual machine
Data store	Data store
LUN	Volume
Fibre channel port	Port

These are common terminology mappings only and might not represent every case for this data collector.

Requirements

The following information is required to configure this data collector:

- IP address of the Virtual Center server
- Read-only username and password in Virtual Center
- We require read only privileges on all objects within Virtual Center.

- SDK access on the Virtual Center server – normally already setup.
- Port requirements: http-80 https-443
- Validate access:
  - Log into Virtual Center Client using above username and password
  - Verify SDK enabled: telnet <vc_ip> 443

Setup and connection

Field	Description
Name	Unique name for the data collector
Acquisition unit	Name of acquisition unit

Configuration

Field	Description
Virtual center IP Address	IP address of the Virtual Center
User name	User name used to access the Virtual Center
Password	Password used to access the Virtual Center

Advanced configuration

In the advanced configuration screen, check the **VM Performance** box to collect performance data. Inventory collection is enabled by default.
The following fields can be configured:

Field	Description
Inventory poll interval (min)	Default is 20
Filter VMs	Select CLUSTER, DATACENTER, or ESX HOST
Choose 'Exclude' or 'Include' to Specify a List	Create a filter list (CLUSTER, DATACENTER, and/or ESX_HOST)
Number of retries	Default is 3
Communication port	Default is 443

Filter Device List…	This list must consist of exact string matches - if you intend to filter by ESX_HOST, you must build a comma delimited list of the exact "names" of your ESX hosts as reported in both Data Infrastructure Insights and vSphere. These "names" may be either IP addresses, simple hostnames, or fully qualified domain names (FQDNs) - this is determined by how these hosts were named when they were originally added to vSphere. When filtering by CLUSTER, use the Data Infrastructure Insights-style cluster names as reported by CI on hypervisors - Data Infrastructure Insights prepends the vSphere cluster name with the vSphere datacenter name and a forward slash - "DC1/clusterA" is the cluster name Data Infrastructure Insights would report on a hypervisor in clusterA within data center DC1.
Performance poll interval (sec)	Default is 300

Mapping VMware tags to Data Infrastructure Insights annotations

The VMware data collector allows you to populate Data Infrastructure Insights annotations with tags configured on VMware. The DII annotation name must be identical to the VMware **category** name; the tag is then applied as an annotation value of a DII annotation of the same name as the category. Data Infrastructure Insights will always populate same-named text-type annotations, and will make a "best attempt" to populate annotations of other types (number, boolean, etc). If your annotation is of a different type and the data collector fails to populate it, it may be necessary to remove the annotation and re-create it as a text type.

Note that VMware tags may be case-sensitive, while Data Infrastructure Insights tags are case-insensitive. So if you create an annotation named "OWNER" in Data Infrastructure Insights, and tags named "OWNER", "Owner", and "owner" in VMware, all of those variations of "owner" would map to Cloud Insight's "OWNER" annotation.

Keep the following in mind:

- Data Infrastructure Insights currently only auto-publishes support information for NetApp devices.
- Since this support information is held in annotation form, you can query it or use it in dashboards.
- If a user overwrites or empties the annotation value, the value is autofilled again when Data Infrastructure Insights updates annotations, which it does once a day.

Troubleshooting

Some things to try if you encounter problems with this data collector:

Inventory

Problem:	Try this:
Error: Include list to filter VMs cannot be empty	If Include List is selected, please list valid DataCenter, Cluster, or Host names to filter VMs

Problem:	Try this:
Error: Failed to instantiate a connection to VirtualCenter at IP	Possible solutions: * Verify credentials and IP address entered. * Try to communicate with Virtual Center using VMware Infrastructure Client. * Try to communicate with Virtual Center using Managed Object Browser (e.g MOB).
Error: VirtualCenter at IP has non-conform certificate that JVM requires	Possible solutions: * Recommended: Re-generate certificate for Virtual Center by using stronger (e.g. 1024-bit) RSA key. * Not Recommended: Modify the JVM java.security configuration to leverage the constraint jdk.certpath.disabledAlgorithms to allow 512-bit RSA key. See JDK 7 update 40 release notes.
I'm seeing the message: "VMware Logs package is not supported on VMware below version 8.0.1"	Logs collection is not supported on VMware versions prior to 8.0.1. Upgrade your VI Center Infrastructure to version 8.0.1 or later if you wish to use the Logs Collections feature within Data Infrastructure Insights. For more information, see this KB Article.

Additional information may be found from the Support page or in the Data Collector Support Matrix.

Data Collector Reference - Services

Node Data Collection

Data Infrastructure Insights gathers metrics from the node on which you install an agent.

Installation

1. From **Observability > Collectors**, choose an operating system/platform. Note that installing any integration data collector (Kubernetes, Docker, Apache, etc.) will also configure node data collection.
2. Follow the instructions to configure the agent. The instructions vary depending on the type of Operating System or Platform you are using to collect data.

Objects and Counters

The following objects and their counters are collected as Node metrics:

Object:	Identifiers:	Attributes:	Datapoints:
Node Filesystem	Node UUID Device Path Type	Node IP Node Name Node OS Mode	Free Inodes Free Inodes Total Inodes Used Total Used Total Used
Node Disk	Node UUID Disk	Node IP Node Name Node OS	IO Time Total IOPS In Progress Read Bytes (per sec) Read Time Total Reads (per sec) Weighted IO Time Total Write Bytes (per sec) Write Time Total Writes (per sec) Current Disk Queue Length Write Time Read Time IO Time
Node CPU	Node UUID CPU	Node IP Node Name Node OS	System CPU Usage User CPU Usage Idle CPU Usage Processor CPU Usage Interrupt CPU Usage DPC CPU Usage

Object:	Identifiers:	Attributes:	Datapoints:
Node	Node UUID	Node IP Node Name Node OS	Kernel Boot Time Kernel Context Switches (per sec) Kernel Entropy Available Kernel Interrupts (per sec) Kernel Processes Forked (per sec) Memory Active Memory Available Total Memory Available Memory Buffered Memory Cached Memory Commit Limit Memory Committed As Memory Dirty Memory Free Memory High Free Memory High Total Memory Huge Page Size Memory Huge Pages Free Memory Huge Pages Total Memory Low Free Memory Low Total Memory Mapped Memory Page Tables Memory Shared Memory Slab Memory Swap Cached Memory Swap Free Memory Swap Total Memory Total Memory Used Total Memory Used Memory Vmalloc Chunk Memory Vmalloc Total Memory Vmalloc Used Memory Wired Memory Writeback Total Memory Writeback Tmp Memory Cache Faults Memory Demand Zero Faults Memory Page Faults Memory Pages Memory Nonpaged Memory Paged Memory Cache Core Memory Standby Cache Normal Memory Standby Cache Reserve Memory Transition Faults Processes Blocked Processes Dead

Object:	Identifiers:	Attributes:	Datapoints:
Node Network	Network Interface Node UUID	Node Name Node IP Node OS	Bytes Received Bytes Sent Packets Outboud Discarded Packets Outboud Errors Packets Received Discarded Packets Received Errors Packets Received Packets Sent

Setup

Setup and Troubleshooting information can be found on the Configuring an Agent page.

ActiveMQ Data Collector

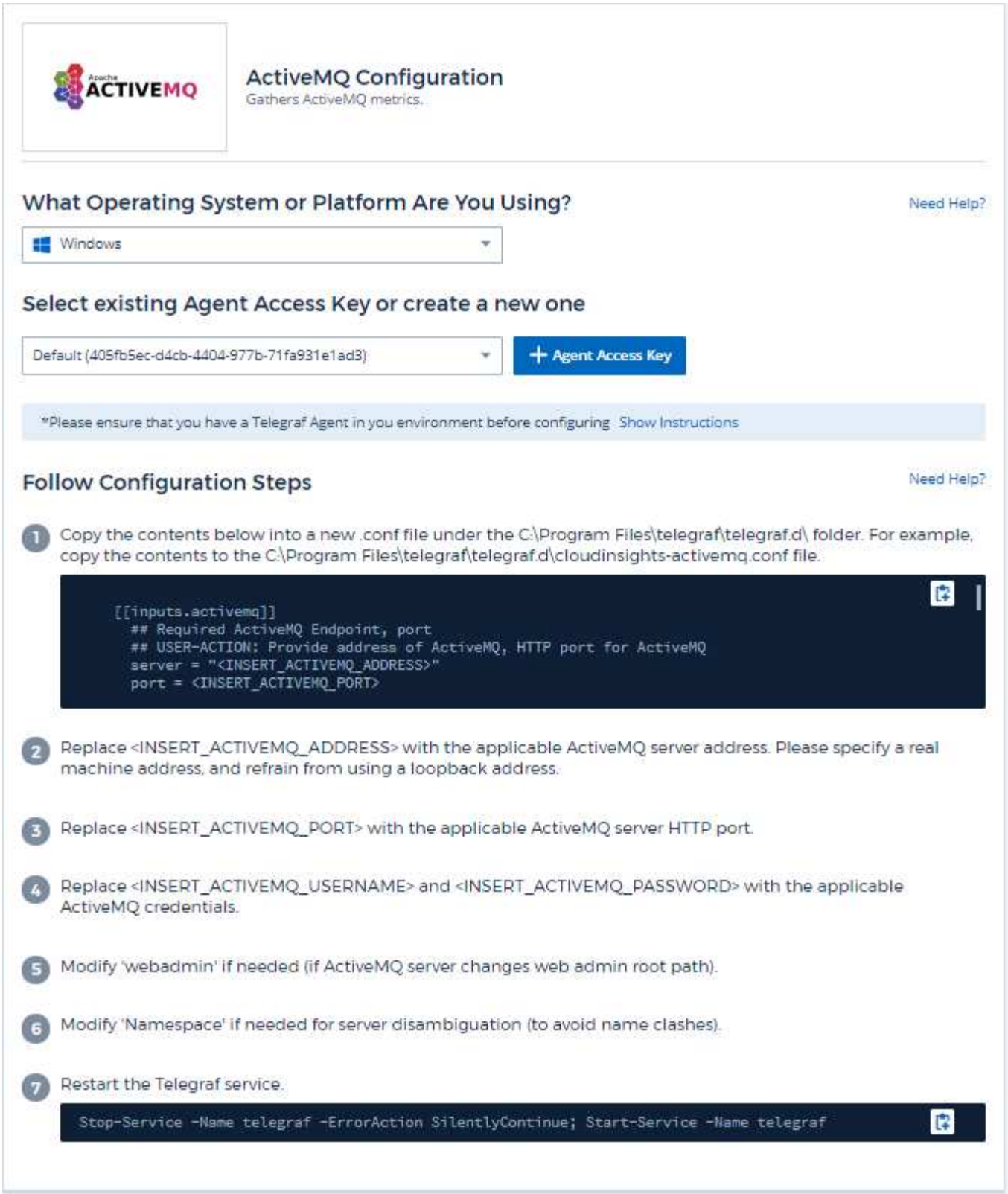
Data Infrastructure Insights uses this data collector to gather metrics from ActiveMQ.

Installation

1. From **Observability > Collectors**, click **+Data Collector**. Choose ActiveMQ.

   Select the Operating System or Platform on which the Telegraf agent is installed.

2. If you haven't already installed an Agent for collection, or you wish to install an Agent for a different Operating System or Platform, click *Show Instructions* to expand the Agent installation instructions.
3. Select the Agent Access Key for use with this data collector. You can add a new Agent Access Key by clicking the **+ Agent Access Key** button. Best practice: Use a different Agent Access Key only when you want to group data collectors, for example, by OS/Platform.
4. Follow the configuration steps to configure the data collector. The instructions vary depending on the type of Operating System or Platform you are using to collect data.



Setup

Information may be found in the ActiveMQ documentation

Objects and Counters

The following objects and their counters are collected:

Object:	Identifiers:	Attributes:	Datapoints:
ActiveMQ Queue	Namespace Queue Port Server	Node Name Node IP Node UUID	Consumer Count Dequeue Count Enqueue Count Queue Size
ActiveMQ Subscriber	Client ID Connection ID Port Server Namespace	Is Active Destination Node Name Node IP Node UUID Node OS Selector Subscription	Dequeue Count Dispatched Count Dispatched Queue Size Enqueue Count Pending Queue Size
ActiveMQ Topic	Topic Port Server Namespace	Node Name Node IP Node UUID Node OS	Consumer Count Dequeue Count Enqueue Count Size

Troubleshooting

Additional information may be found from the Support page.

Apache Data Collector

This data collector allows collection of data from Apache servers on your tenant.

Pre-requisites

- You must have your Apache HTTP Server set up and properly running
- You must have sudo or administrator permissions on your agent host/VM
- Typically, the Apache *mod_status* module is configured to expose a page at the '/server-status?auto' location of the Apache server. The *ExtendedStatus* option must be enabled in order to collect all available fields. For information about how to configure your server, see the Apache module documentation: https://httpd.apache.org/docs/2.4/mod/mod_status.html#enable

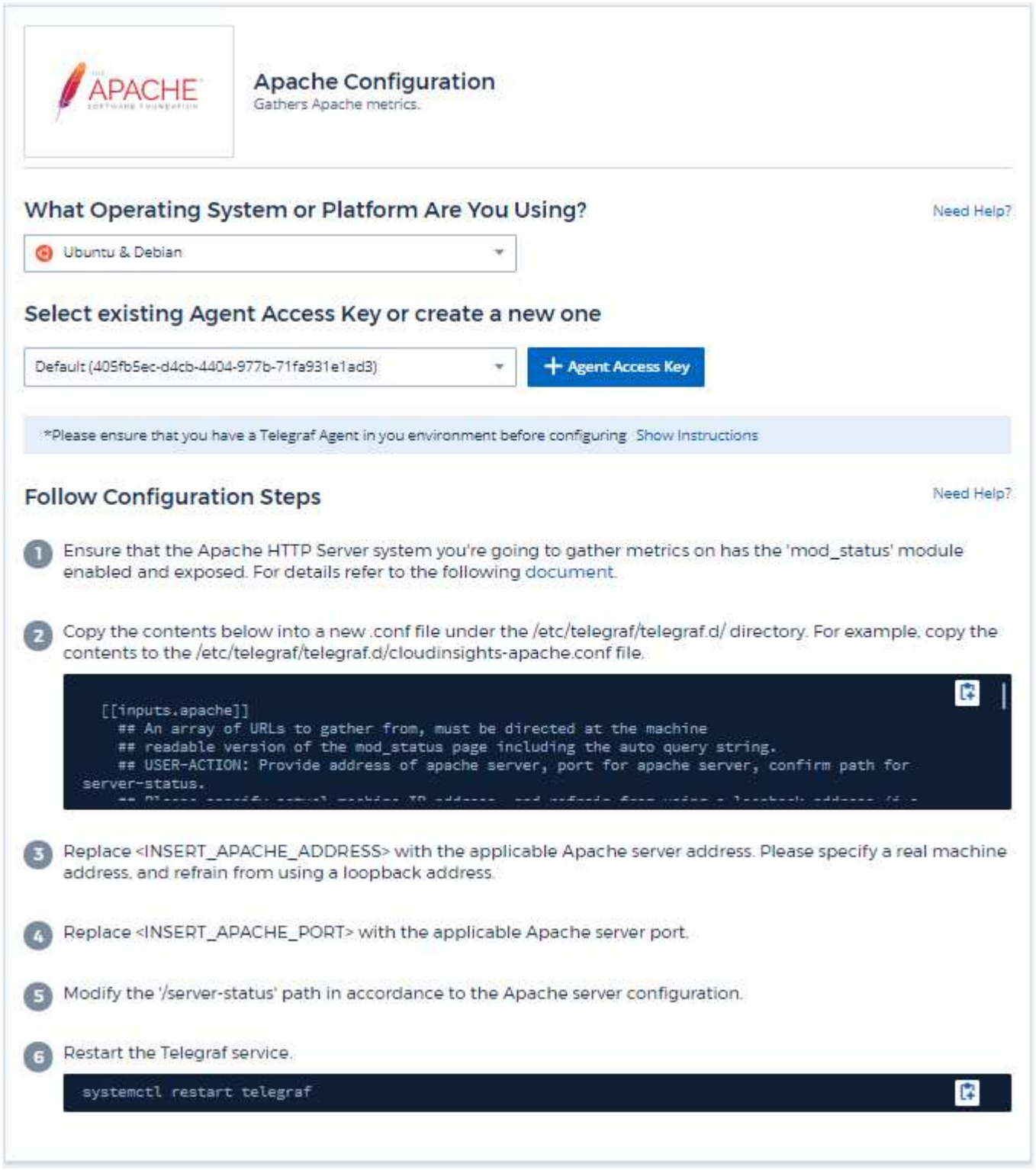
Installation

1. From **Observability > Collectors**, click **+Data Collector**. Choose Apache.

   Select the Operating System or Platform on which the Telegraf agent is installed.

2. If you haven't already installed an Agent for collection, or you wish to install an Agent for a different Operating System or Platform, click *Show Instructions* to expand the Agent installation instructions.
3. Select the Agent Access Key for use with this data collector. You can add a new Agent Access Key by clicking the **+ Agent Access Key** button. Best practice: Use a different Agent Access Key only when you want to group data collectors, for example, by OS/Platform.

4. Follow the configuration steps to configure the data collector. The instructions vary depending on the type of Operating System or Platform you are using to collect data.



Setup

Telegraf's plugin for Apache's HTTP Server relies on the 'mod_status' module to be enabled. When this is enabled, Apache's HTTP Server will expose an HTML endpoint that can be viewed on your browser or scraped for extraction of status of all Apache's HTTP Server configuration.

Compatibility:

Configuration was developed against Apache's HTTP Server version 2.4.38.

Enabling mod_status:

Enabling and exposing the 'mod_status' modules involves two steps:

- Enabling module
- Exposing stats from module

Enabling module:

The loading of modules is controlled by the config file under '/usr/local/apache/conf/httpd.conf'. Edit the config file and uncomment the following lines:

```
LoadModule status_module modules/mod_status.so
```

```
Include conf/extra/httpd-info.conf
```

Exposing stats from module:

The exposing of 'mod_status' is controlled by the config file under '/usr/local/apache2/conf/extra/httpd-info.conf'. Make sure you have the following in that configuration file (at least, other directives will be there):

```
# Allow server status reports generated by mod_status,
# with the URL of http://servername/server-status
<Location /server-status>
    SetHandler server-status
</Location>

#
# ExtendedStatus controls whether Apache will generate "full" status
# information (ExtendedStatus On) or just basic information
(ExtendedStatus
# Off) when the "server-status" handler is called. The default is Off.
#
ExtendedStatus On
```

For detailed instructions on the 'mod_status' module, see the Apache documentation

Objects and Counters

The following objects and their counters are collected:

Object:	Identifiers:	Attributes:	Datapoints:
Apache	Namespace Server	Node IP Node Name Port Parent Server Config Generation Parent Server MPM Generation Server Uptime Is Stopping	Busy Workers Bytes per Request Bytes per Second CPU Children System CPU Children User CPU Load CPU System CPU User Asynchronous Connections Closing Asynchronous Connections Keep Alive Asynchronous Connections Writing Connections Total Duration per Request Idle Workers Load Average (last 1m) Load Average (last 15m) Load Average (last 5m) Processes Requests per Second Total Accesses Total Duration Total KBytes Scoreboard Closing Scoreboard DNS Lookups Scoreboard Finishing Scoreboard Idle Cleanup Scoreboard Keep Alive Scoreboard Logging Scoreboard Open Scoreboard Reading Scoreboard Sending Scoreboard Starting Scoreboard Waiting

Troubleshooting

Additional information may be found from the Support page.

Consul Data Collector

Data Infrastructure Insights uses this data collector to gather metrics from Consul.

Installation

1. From **Observability > Collectors**, click **+Data Collector**. Choose Consul.

   If you haven't configured an Agent for collection, you are prompted to install an agent on your tenant.

   If you have an agent already configured, select the appropriate Operating System or Platform and click

Continue.

2. Follow the instructions in the Consul Configuration screen to configure the data collector. The instructions vary depending on the type of Operating System or Platform you are using to collect data.

Setup

Information may be found in the Consul documentation.

Objects and Counters for consul

The following objects and their counters are collected:

Object:	Identifiers:	Attributes:	Datapoints:
Consul	Namespace Check ID Service Node	Node IP Node OS Node UUID Node Name Service Name Check Name Service ID Status	Critical Passing Warning

Troubleshooting

Additional information may be found from the Support page.

Couchbase Data Collector

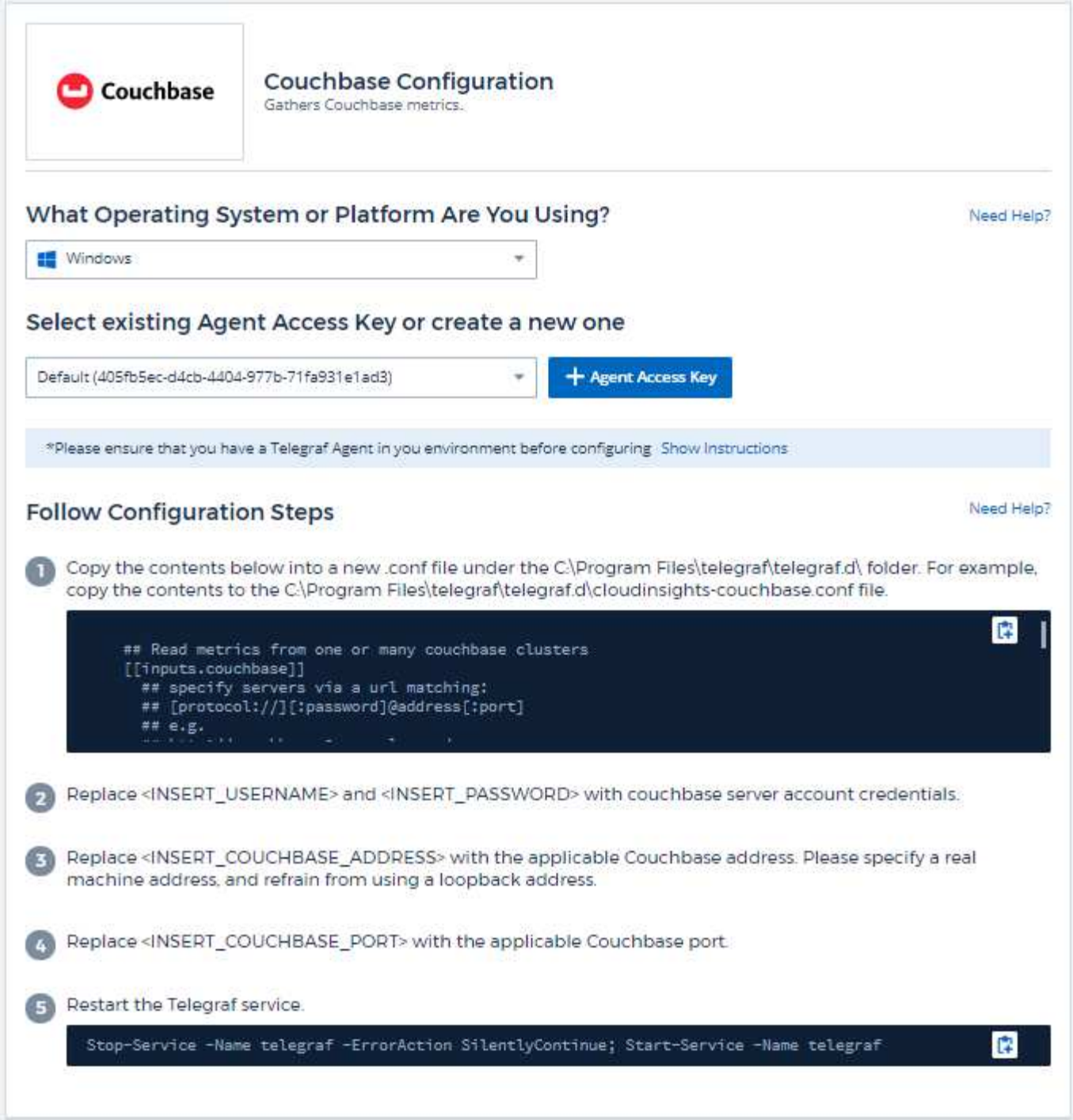
Data Infrastructure Insights uses this data collector to gather metrics from Couchbase.

Installation

1. From **Observability > Collectors**, click **+Data Collector**. Choose Couchbase.

   Select the Operating System or Platform on which the Telegraf agent is installed.

2. If you haven't already installed an Agent for collection, or you wish to install an Agent for a different Operating System or Platform, click *Show Instructions* to expand the Agent installation instructions.
3. Select the Agent Access Key for use with this data collector. You can add a new Agent Access Key by clicking the **+ Agent Access Key** button. Best practice: Use a different Agent Access Key only when you want to group data collectors, for example, by OS/Platform.
4. Follow the configuration steps to configure the data collector. The instructions vary depending on the type of Operating System or Platform you are using to collect data.



Setup

Information may be found in the Couchbase documentation.

Objects and Counters

The following objects and their counters are collected:

Object:	Identifiers:	Attributes:	Datapoints:
Couchbase Node	Namespace Cluster Couchbase Node Hostname	Node Name Node IP	Memory Free Memory Total
Couchbase Bucket	Namespace Bucket Cluster	Node Name Node IP	Data Used Data Fetches Disk Used Item Count Memory Used Operations Per Second Quota Used

Troubleshooting

Additional information may be found from the Support page.

CouchDB Data Collector

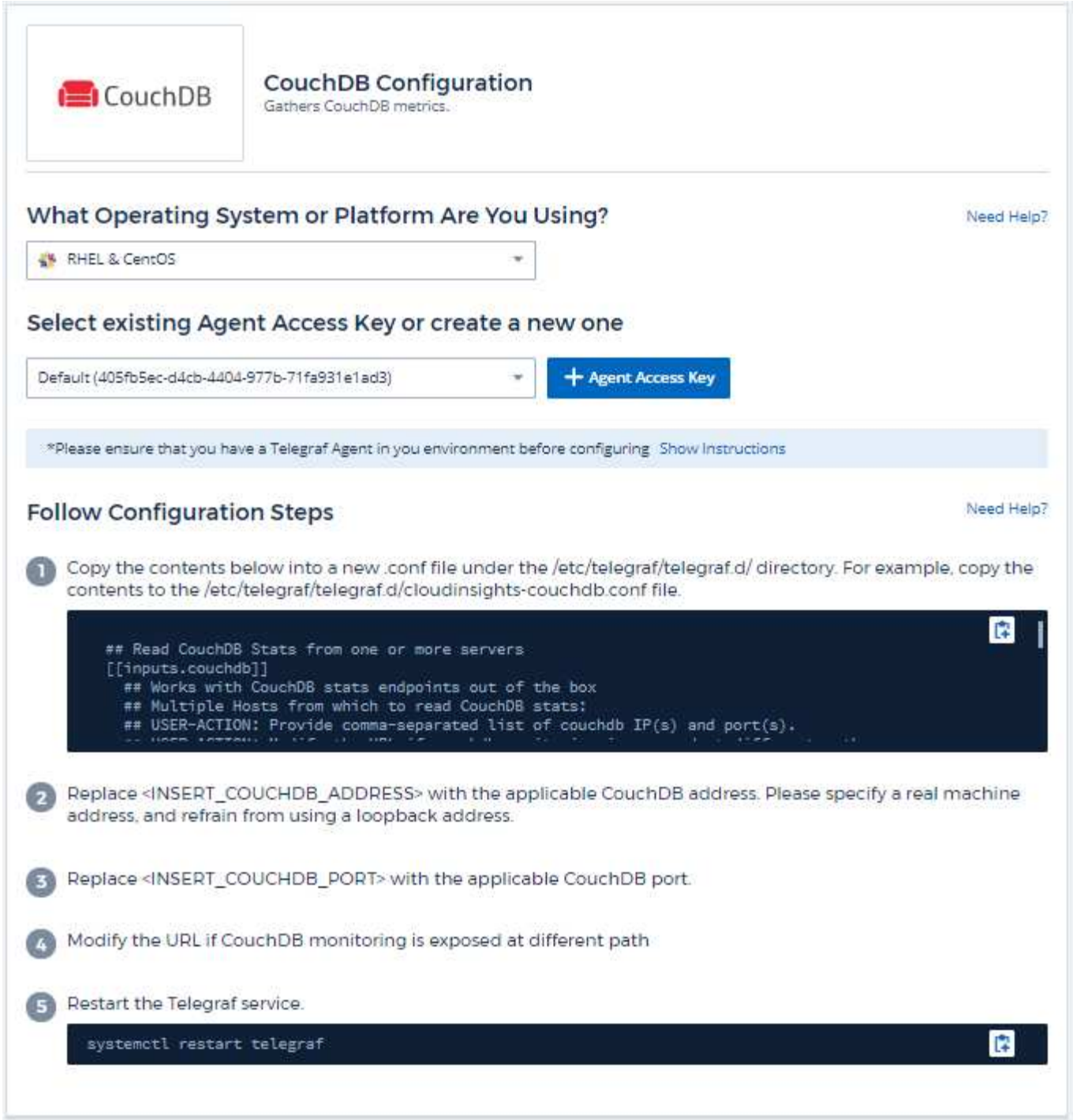
Data Infrastructure Insights uses this data collector to gather metrics from CouchDB.

Installation

1. From **Observability > Collectors**, click **+Data Collector**. Choose CouchDB.

   Select the Operating System or Platform on which the Telegraf agent is installed.

2. If you haven't already installed an Agent for collection, or you wish to install an Agent for a different Operating System or Platform, click *Show Instructions* to expand the Agent installation instructions.
3. Select the Agent Access Key for use with this data collector. You can add a new Agent Access Key by clicking the **+ Agent Access Key** button. Best practice: Use a different Agent Access Key only when you want to group data collectors, for example, by OS/Platform.
4. Follow the configuration steps to configure the data collector. The instructions vary depending on the type of Operating System or Platform you are using to collect data.



Setup

Information may be found in the CouchDB documentation.

Objects and Counters

The following objects and their counters are collected:

Object:	Identifiers:	Attributes:	Datapoints:
CouchDB	Namespace Server	Node Name Node IP	Authentication Cache Hits Authentication Cache Miss Database Reads Database Writes Databases Open Open OS Files Max Request Time Min Request Time Httpd Request Methods Copy Httpd Request Methods Delete Httpd Request Methods Get Httpd Request Methods Head Httpd Request Methods Post Httpd Request Methods Put Status Codes 200 Status Codes 201 Status Codes 202 Status Codes 301 Status Codes 304 Status Codes 400 Status Codes 401 Status Codes 403 Status Codes 404 Status Codes 405 Status Codes 409 Status Codes 412 Status Codes 500

Troubleshooting

Additional information may be found from the Support page.

Docker Data Collector

Data Infrastructure Insights uses this data collector to gather metrics from Docker.

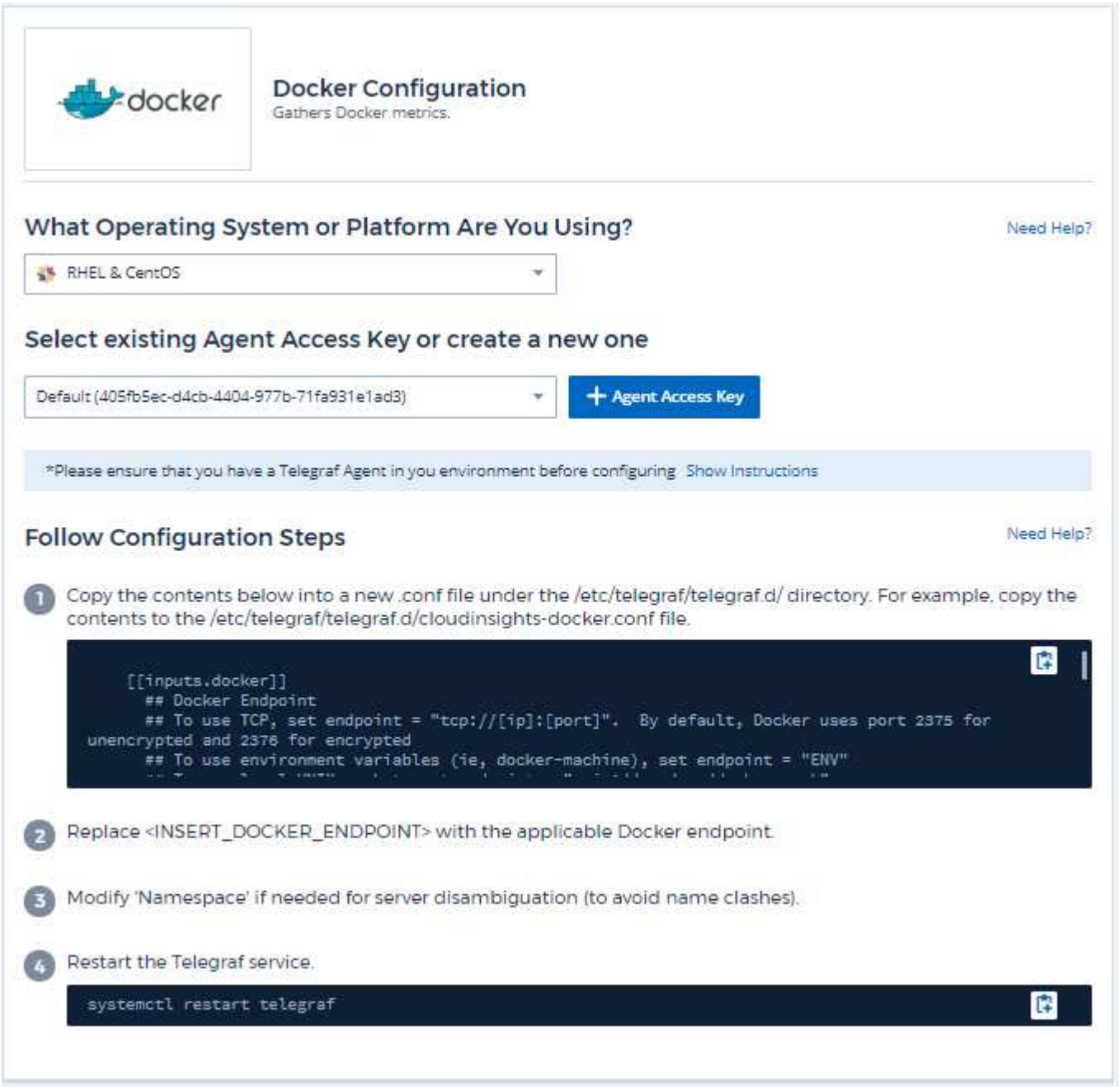
Installation

1. From **Observability > Collectors**, click **+Data Collector**. Choose Docker.

   If you haven't configured an Agent for collection, you are prompted to install an agent on your tenant.

   If you have an agent already configured, select the appropriate Operating System or Platform and click **Continue**.

2. Follow the instructions in the Docker Configuration screen to configure the data collector. The instructions vary depending on the type of Operating System or Platform you are using to collect data.



Setup

The Telegraf input plugin for Docker collects metrics through a specified UNIX socket or a TCP endpoint.

Compatibility

Configuration was developed against Docker version 1.12.6.

Setting Up

Accessing Docker through a UNIX socket

If the Telegraf agent is running on baremetal, add the telegraf Unix user to the docker Unix group by running the following:

```
sudo usermod -aG docker telegraf
```

If the Telegraf agent is running within a Kubernetes pod, expose the Docker Unix socket by mapping the socket into the pod as a volume and then mounting that volume to /var/run/docker.sock. For example, add the following to the PodSpec:

```
volumes:
...
- name: docker-sock
hostPath:
path: /var/run/docker.sock
type: File
```

Then, add the following to the Container:

```
volumeMounts:
...
- name: docker-sock
mountPath: /var/run/docker.sock
```

Note that the Data Infrastructure Insights installer provided for the Kubernetes platform takes care of this mapping automatically.

Access Docker through a TCP endpoint

By default, Docker uses port 2375 for unencrypted access and port 2376 for encrypted access.

Objects and Counters

The following objects and their counters are collected:

Object:	Identifiers:	Attributes:	Datapoints:
Docker Engine	Namespace Docker Engine	Node Name Node IP Node UUID Node OS Kubernetes Cluster Docker Version Unit	Memory Containers Containers Paused Containers Running Containers Stopped CPUs Go Routines Images Listener Events Used File Descriptors Data Available Data Total Data Used Metadata Available Metadata Total Metadata Used Pool Blocksize

Object:	Identifiers:	Attributes:	Datapoints:
Docker Container	Namespace Container Name Docker Engine	Kubernetes Container Hash Kubernetes Container Ports Kubernetes Container Restart Count Kubernetes Container Termination Message Path Kubernetes Container Termination Message Policy Kubernetes Pod Termination Grace Period Container Image Container Status Container Version Node Name Kubernetes Container Log Path Kubernetes Container Name Kubernetes Docker Type Kubernetes Pod Name Kubernetes Pod Namespace Kubernetes Pod UID Kubernetes Sandbox ID Node IP Node UUID Docker Version Kubernetes IO Config Seen Kubernetes IO Config Source OpenShift IO SCC Kubernetes Description Kubernetes Display Name OpenShift Tags Kompose Service Pod Template Hash Controller Revision Hash Pod Template Generation License Schema Build Date Schema License Schema Name Schema URL Schema VCS URL Schema Vendor Schema Version Schema Schema Version Maintainer Customer Pod	Memory Active Anonymous Memory Active File Memory Cache Memory Hierarchical Limit Memory Inactive Anonymous Memory Inactive File Memory Limit Memory Mapped File Memory Max Usage Memory Page Fault Memory Page Major Fault Memory Paged In Memory Paged Out Memory Resident Set Size Memory Resident Set Size Huge Memory Total Active Anonymous Memory Total Active File Memory Total Cache Memory Total Inactive Anonymous Memory Total Inactive File Memory Total Mapped File Memory Total Page Fault Memory Total Page Major Fault Memory Total Paged In Memory Total Paged Out Memory Total Resident Set Size Memory Total Resident Set Size Huge Memory Total Unevictable Memory Unevictable Memory Usage Memory Usage Percent Exit Code OOM Killed PID Started At Failing Streak

Object:	Identifiers:	Attributes:	Datapoints:
Docker Container Block IO	Namespace Container Name Device Docker Engine	Kubernetes Container Hash Kubernetes Container Ports Kubernetes Container Restart Count Kubernetes Container Termination Message Path Kubernetes Container Termination Message Policy Kubernetes Pod Termination Grace Period Container Image Container Status Container Version Node Name Kubernetes Container Log Path Kubernetes Container Name Kubernetes Docker Type Kubernetes Pod Name Kubernetes Pod Namespace Kubernetes Pod UID Kubernetes Sandbox ID Node IP Node UUID Docker Version Kubernetes Config Seen Kubernetes Config Source OpenShift SCC Kubernetes Description Kubernetes Display Name OpenShift Tags Schema Schema Version Pod Template Hash Controller Revision Hash Pod Template Generation Kompose Service Schema Build Date Schema License Schema Name Schema Vendor Customer Pod Kubernetes StatefulSet Pod Name Tenant Webconsole Build Date License Vendor	IO Service Bytes Recursive Async IO Service Bytes Recursive Read IO Service Bytes Recursive Sync IO Service Bytes Recursive Total IO Service Bytes Recursive Write IO Serviced Recursive Async IO Serviced Recursive Read IO Serviced Recursive Sync IO Serviced Recursive Total IO Serviced Recursive Write

Object:	Identifiers:	Attributes:	Datapoints:
Docker Container Network	Namespace Container Name Network Docker Engine	Container Image Container Status Container Version Node Name Node IP Node UUID Node OS K8s Cluster Docker Version Container ID	RX Dropped RX Bytes RX Errors RX Packets TX Dropped TX Bytes TX Errors TX Packets

Object:	Identifiers:	Attributes:	Datapoints:
Docker Container CPU	Namespace Container Name CPU Docker Engine	Kubernetes Container Hash Kubernetes Container Ports Kubernetes Container Restart Count Kubernetes Container Termination Message Path Kubernetes Container Termination Message Policy Kubernetes Pod Termination Grace Period Kubernetes Config Seen Kubernetes Config Source OpenShift SCC Container Image Container Status Container Version Node Name Kubernetes Container Log Path Kubernetes Container name Kubernetes Docker Type Kubernetes Pod Name Kubernetes Pod Namespace Kubernetes Pod UID Kubernetes Sandbox ID Node IP Node UUID Node OS Kubernetes Cluster Docker Version Kubernetes Description Kubernetes Display Name OpenShift Tags Schema Version Pod Template Hash Controller Revision Hash Pod Template Generation Kompose Service Schema Build Date Schema License Schema Name Schema Vendor Customer Pod Kubernetes StatefulSet Pod Name Tenant Webconsole Build Date	Throttling Periods Throttling Throttled Periods Throttling Throttled Time Usage In Kernel Mode Usage In User Mode Usage Percent Usage System Usage Total

Troubleshooting

Problem:	Try this:
I do not see my Docker metrics in Data Infrastructure Insights after following the instructions on the configuration page.	Check the Telegraf agent logs to see if it reports the following error: E! Error in plugin [inputs.docker]: Got permission denied while trying to connect to the Docker daemon socket If it does, take the necessary steps to provide the Telegraf agent access to the Docker Unix socket as specified above.

Additional information may be found from the Support page.

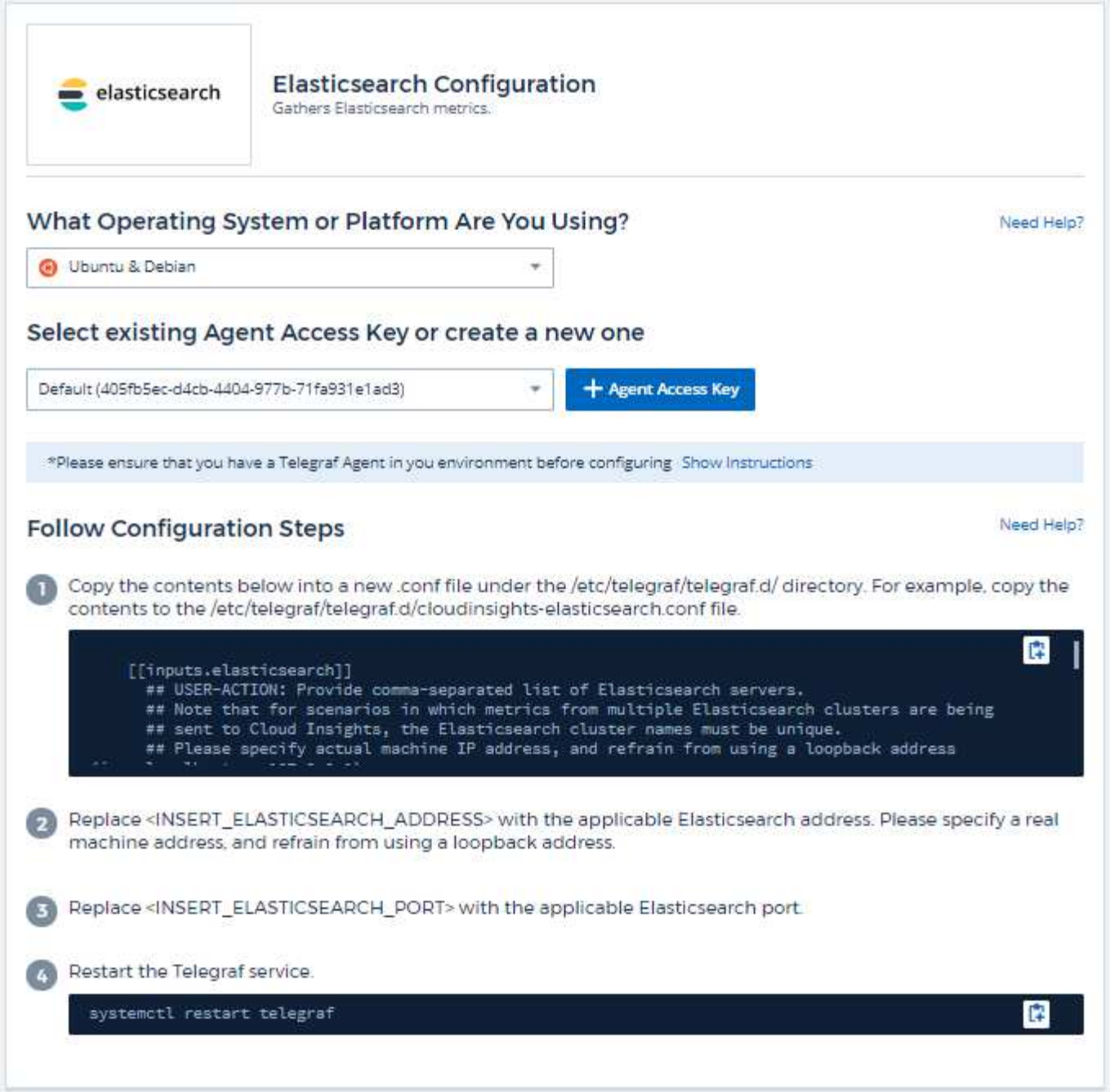
Elasticsearch Data Collector

Data Infrastructure Insights uses this data collector to gather metrics from Elasticsearch.

1. From **Observability > Collectors**, click **+Data Collector**. Choose Elasticsearch.

   Select the Operating System or Platform on which the Telegraf agent is installed.

2. If you haven't already installed an Agent for collection, or you wish to install an Agent for a different Operating System or Platform, click *Show Instructions* to expand the Agent installation instructions.
3. Select the Agent Access Key for use with this data collector. You can add a new Agent Access Key by clicking the **+ Agent Access Key** button. Best practice: Use a different Agent Access Key only when you want to group data collectors, for example, by OS/Platform.
4. Follow the configuration steps to configure the data collector. The instructions vary depending on the type of Operating System or Platform you are using to collect data.



Setup

Information may be found in the Elasticsearch documentation.

Objects and Counters

The following objects and their counters are collected:

Object:	Identifiers:	Attributes:
Elasticsearch Cluster	Namespace Cluster	Node IP Node Name Cluster Status

Object:	Identifiers:	Attributes:
Elasticsearch Node	Namespace Cluster ES Node ID ES Node IP ES Node	Zone ID

Troubleshooting

Additional information may be found from the Support page.

Flink Data Collector

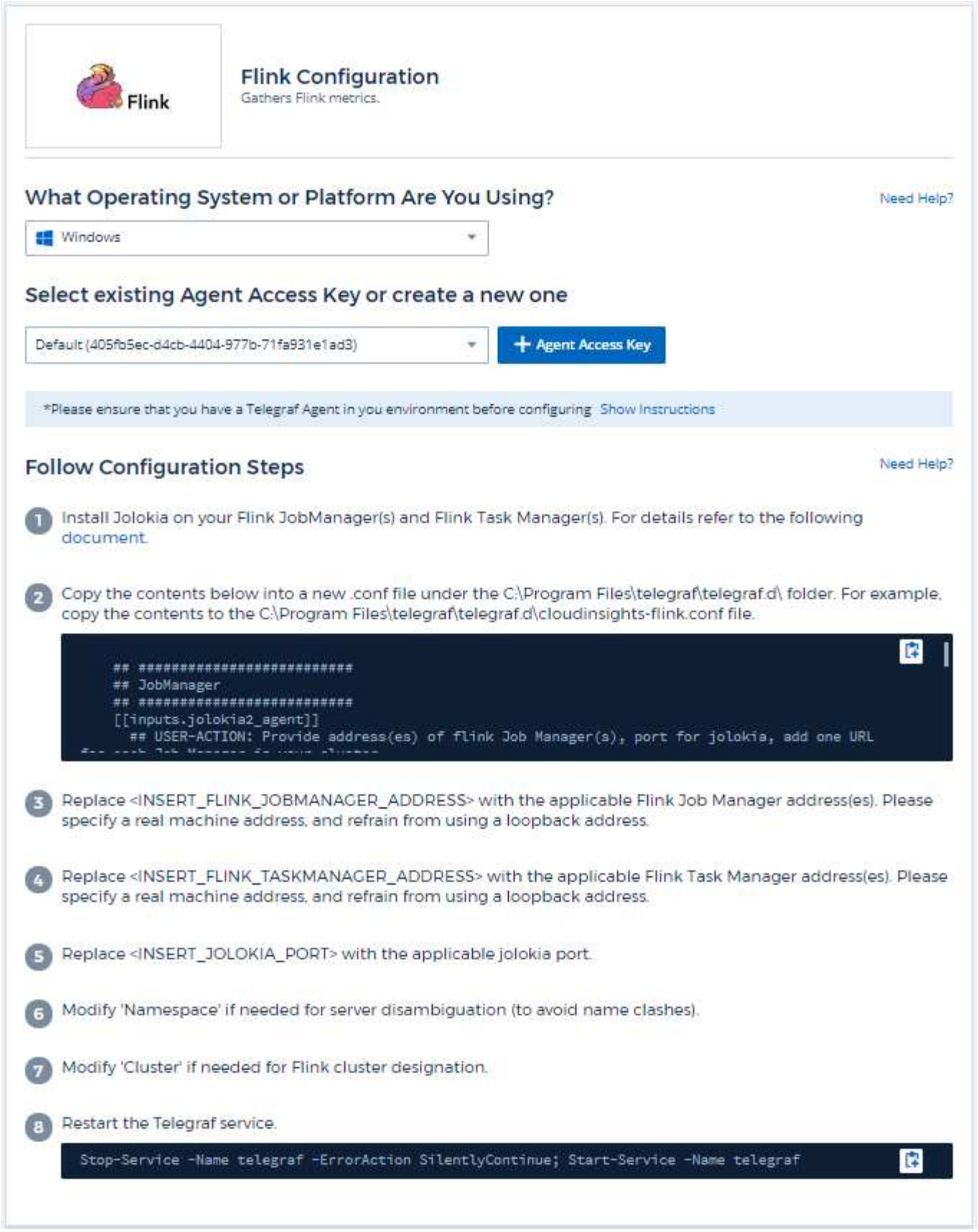
Data Infrastructure Insights uses this data collector to gather metrics from Flink.

Installation

1. From **Observability > Collectors**, click **+Data Collector**. Choose Flink.

   Select the Operating System or Platform on which the Telegraf agent is installed.

2. If you haven't already installed an Agent for collection, or you wish to install an Agent for a different Operating System or Platform, click *Show Instructions* to expand the Agent installation instructions.
3. Select the Agent Access Key for use with this data collector. You can add a new Agent Access Key by clicking the **+ Agent Access Key** button. Best practice: Use a different Agent Access Key only when you want to group data collectors, for example, by OS/Platform.
4. Follow the configuration steps to configure the data collector. The instructions vary depending on the type of Operating System or Platform you are using to collect data.



Setup

A full Flink deployment involves the following components:

JobManager: The Flink primary system. Coordinates a series of TaskManagers. In a High Availability setup, system will have more than one JobManager.
TaskManager: This is where Flink operators are executed.
The Flink plugin is based on the telegraf's Jolokia plugin. As such as a requirement to gather info from all Flink components, JMX needs to be configured and exposed via Jolokia on all components.

Compatibility

Configuration was developed against Flink version 1.7.0.

Setting Up

Jolokia Agent Jar

For all individual components, a version the Jolokia agent jar file must be downloaded. The version tested against was Jolokia agent 1.6.0.

Instructions below assume that downloaded jar file (jolokia-jvm-1.6.0-agent.jar) is placed under location '/opt/flink/lib/'.

JobManager

To configure JobManager to expose the Jolokia API, you can setup the following environment variable on your nodes then restart the JobManager:

```
export FLINK_ENV_JAVA_OPTS="-javaagent:/opt/flink/lib/jolokia-jvm-1.6.0-agent.jar=port=8778,host=0.0.0.0"
```

You can choose a different port for Jolokia (8778). If you have an internal IP to lock Jolokia onto you can replace the "catch all" 0.0.0.0 by your own IP. Notice this IP needs to be accessible from the telegraf plugin.

TaskManager

To configure TaskManager(s) to expose the Jolokia API, you can setup the following environment variable on your nodes then restart the TaskManager:

```
export FLINK_ENV_JAVA_OPTS="-javaagent:/opt/flink/lib/jolokia-jvm-1.6.0-agent.jar=port=8778,host=0.0.0.0"
```

You can choose a different port for Jolokia (8778). If you have an internal IP to lock Jolokia onto you can replace the "catch all" 0.0.0.0 by your own IP. Notice this IP needs to be accessible from the telegraf plugin.

Objects and Counters

The following objects and their counters are collected:

Object:	Identifiers:	Attributes:	Datapoints:
Flink Task Manager	Cluster Namespace Server	Node Name Task Manager ID Node IP	Network Available Memory Segments Network Total Memory Segments Garbage Collection PS MarkSweep Count Garbage Collection PS MarkSweep Time Garbage Collection PS Scavenge Count Garbage Collection PS Scavenge Time Heap Memory Committed Heap Memory Init Heap Memory Max Heap Memory Used Thread Count Daemon Thread Count Peak Thread Count Thread Count Total Started
Flink Job	Cluster Namespace server Job ID	Node Name Job Name Node IP Last Checkpoint External Path Restarting Time	Downtime Full Restarts Last Checkpoint Alignment Buffered Last Checkpoint Duration Last Checkpoint Size Number of Completed Checkpoints Number of Failed Checkpoints Number of in Progress Checkpoints Number of Checkpoints Uptime

Object:	Identifiers:	Attributes:	Datapoints:
Flink Job Manager	Cluster Namespace Server	Node Name Node IP	Garbage Collection PS MarkSweep Count Garbage Collection PS MarkSweep Time Garbage Collection PS Scavenge Count Garbage Collection PS Scavenge Time Heap Memory Committed Heap Memory Init Heap Memory Max Heap Memory Used Number Registered Task Managers Number Running Jobs Task Slots Available Task Slots Total Thread Count Daemon Thread Count Peak Thread Count Thread Count Total Started

Object:	Identifiers:	Attributes:	Datapoints:
Flink Task	Cluster Namespace Job ID Task ID	Server Node Name Job Name Sub Task Index Task Attempt ID Task Attempt Number Task Name Task Manager ID Node IP Current Input Watermark	Buffers In Pool Usage Buffers In Queue Length Buffers Out Pool Usage Buffers Out Queue Length Number Buffers In Local Number Bufffers In Local Per Second Count Number Buffers in Local Per Second Rate Number Buffers In Remote Number Buffers In Remote Per Second Count Number Buffers In Remote Per Second Rate Number Buffers Out Number Buffers Out Per Second Count Number Buffers Out Per Second Rate Number Bytes In Local Number Bytes In Local Per Second Count Number Bytes In Local Per Second Rate Number Bytes In Remote Number Bytes In Remote Per Second Count Number Bytes In Remote Per Second Rate Number Bytes Out Number Bytes Out Per Second Count Number Bytes Out Per Second Rate Number Records In Number Records In Per Second Count Number Records In Per Second Rate Number Records Out Number Records Out Per Second Count Number Records Out Per Second Rate

Object:	Identifiers:	Attributes:	Datapoints:
Flink Task Operator	Cluster Namespace Job ID Operator ID Task ID	Server Node Name Job Name Operator Name Sub Task Index Task Attempt ID Task Attempt Number Task Name Task Manager ID Node IP	Current Input Watermark Current Output Watermark Number Records In Number Records In Per Second Count Number Records In Per Second Rate Number Records Out Number Records Out Per Second Count Number Records Out Per Second Rate Number Late Records Dropped Assigned Partitions Bytes Consumed Rate Commit Latency Avg Commit Latency Max Commit Rate Commits Failed Commits Succeeded Connection Close Rate Connection Count Connection Creation Rate Count Fetch Latency Avg Fetch Latency Max Fetch Rate Fetch Size Avg Fetch Size Max Fetch Throttle Time Avg Fetch Throttle Time Max Heartbeat Rate Incoming Byte Rate IO Ratio IO Time Avg (ns) IO Wait Ratio IO Wait Time Avg (ns) Join Rate Join Time Avg Last Heartbeat Ago Network IO Rate Outgoing Byte Rate Records Consumed Rate Records Lag Max Records per Request Avg Request Rate Request Size Avg Request Size Max Response Rate Select Rate Sync Rate Sync Time Avg Heartbeat Response Time

Troubleshooting

Additional information may be found from the Support page.

Hadoop Data Collector

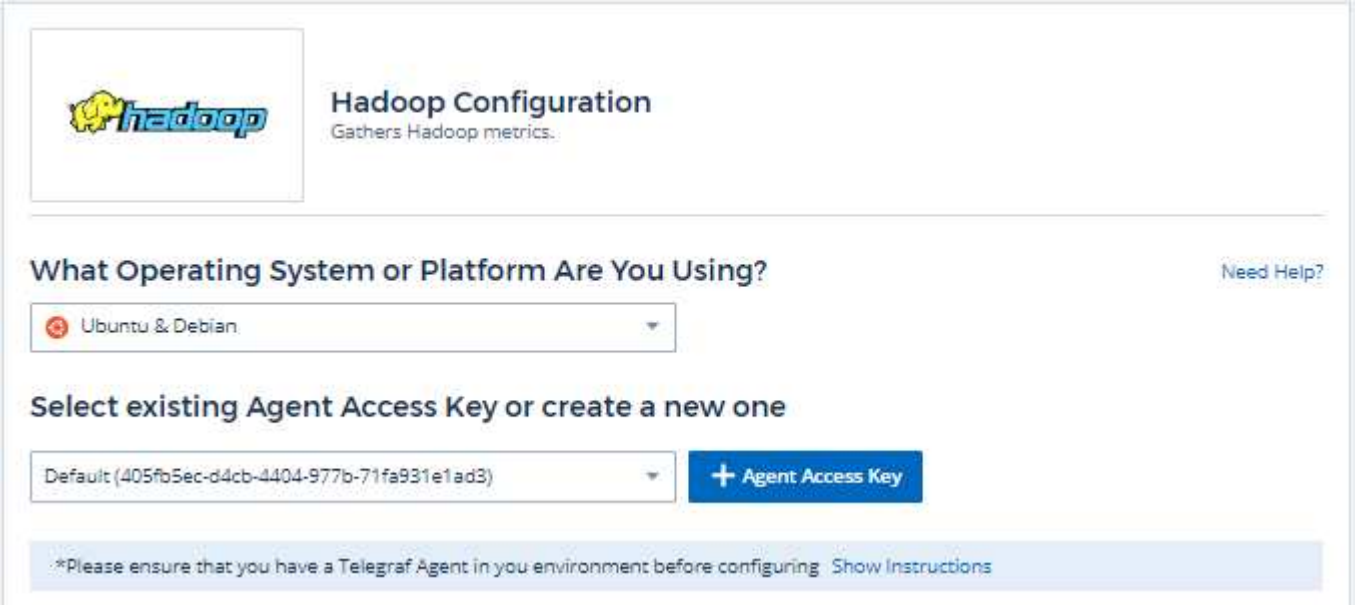
Data Infrastructure Insights uses this data collector to gather metrics from Hadoop.

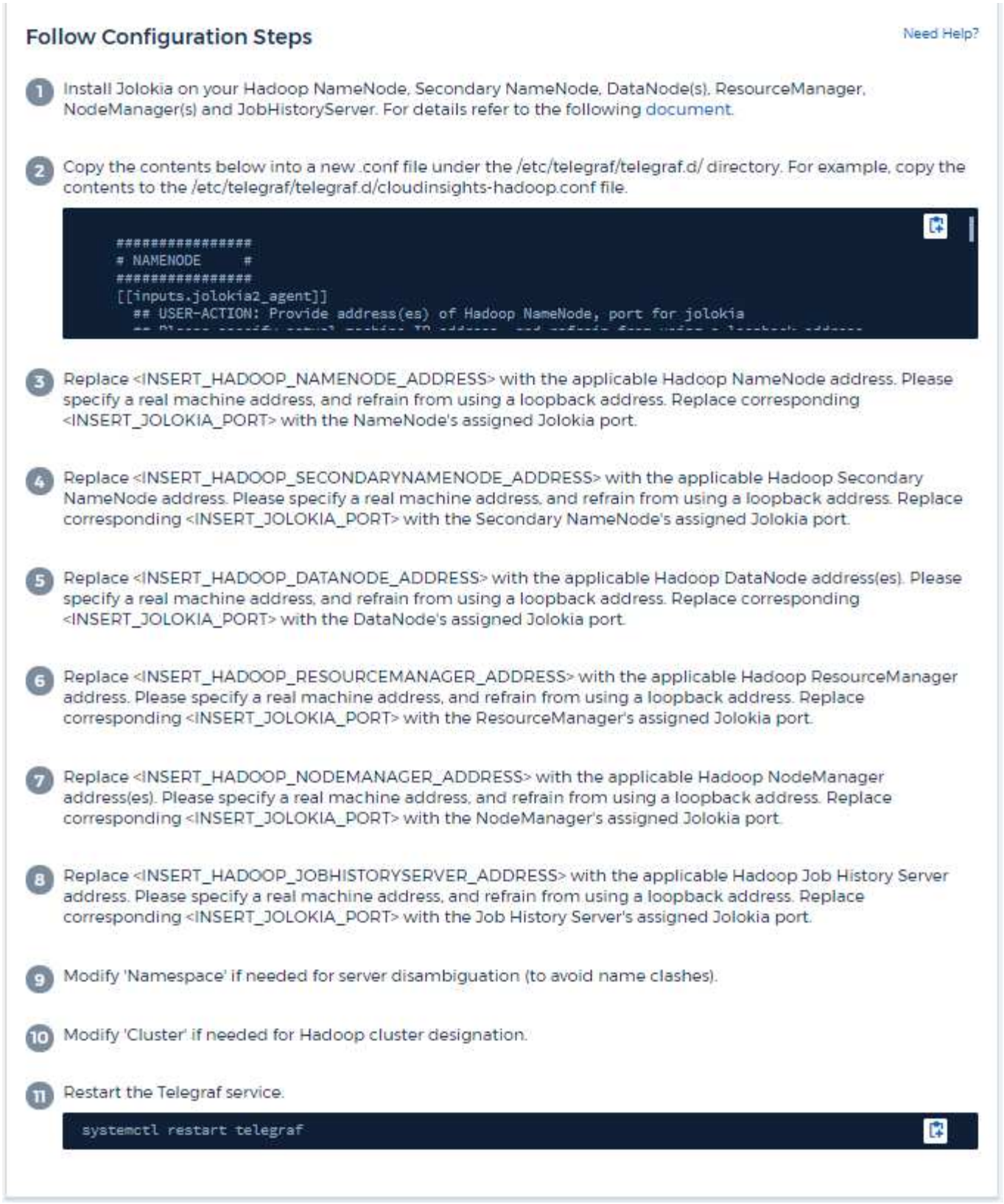
Installation

1. From **Observability > Collectors**, click **+Data Collector**. Choose Hadoop.

   Select the Operating System or Platform on which the Telegraf agent is installed.

2. If you haven't already installed an Agent for collection, or you wish to install an Agent for a different Operating System or Platform, click *Show Instructions* to expand the Agent installation instructions.
3. Select the Agent Access Key for use with this data collector. You can add a new Agent Access Key by clicking the **+ Agent Access Key** button. Best practice: Use a different Agent Access Key only when you want to group data collectors, for example, by OS/Platform.
4. Follow the configuration steps to configure the data collector. The instructions vary depending on the type of Operating System or Platform you are using to collect data.





Setup

A full Hadoop deployment involves the following components:

- NameNode: The Hadoop Distributed File System (HDFS) primary system. Coordinates a series of DataNodes.

- Secondary NameNode: a warm failover for the main NameNode. In Hadoop the promotion to NameNode does not occur automatically. Secondary NameNode gathers information from NameNode to be ready to be promoted when needed.
- DataNode: Actual owner for data.
- ResourceManager: The compute primary system (Yarn). Coordinates a series of NodeManagers.
- NodeManager: The resource for compute. Actual location for running of applications.
- JobHistoryServer: Responsible for servicing all job history related requests.

The Hadoop plugin is based on the telegraf's Jolokia plugin. As such as a requirement to gather info from all Hadoop components, JMX needs to be configured and exposed via Jolokia on all components.

Compatibility

Configuration was developed against Hadoop version 2.9.2.

Setting Up

Jolokia Agent Jar

For all individual components, a version the Jolokia agent jar file must be downloaded. The version tested against was Jolokia agent 1.6.0.

Instructions below assume that downloaded jar file (jolokia-jvm-1.6.0-agent.jar) is placed under location '/opt/hadoop/lib/'.

NameNode

To configure NameNode to expose the Jolokia API, you can setup the following in <HADOOP_HOME>/etc/hadoop/hadoop-env.sh:

```
export HADOOP_NAMENODE_OPTS="$HADOOP_NAMENODE_OPTS
-javaagent:/opt/hadoop/lib/jolokia-jvm-1.6.0
-agent.jar=port=7800,host=0.0.0.0 -Dcom.sun.management.jmxremote
-Dcom.sun.management.jmxremote.port=8000
-Dcom.sun.management.jmxremote.ssl=false
-Dcom.sun.management.jmxremote.password.file=$HADOOP_HOME/conf/jmxremote.p
assword"
You can choose a different port for JMX (8000 above) and Jolokia (7800).
If you have an internal IP to lock Jolokia onto you can replace the "catch
all" 0.0.0.0 by your own IP. Notice this IP needs to be accessible from
the telegraf plugin. You can use the option '-
Dcom.sun.management.jmxremote.authenticate=false' if you don't want to
authenticate. Use at your own risk.
```

Secondary NameNode

To configure the Secondary NameNode to expose the Jolokia API, you can setup the following in <HADOOP_HOME>/etc/hadoop/hadoop-env.sh:

```
export HADOOP_SECONDARYNAMENODE_OPTS="$HADOOP_SECONDARYNAMENODE_OPTS
-javaagent:/opt/hadoop/lib/jolokia-jvm-1.6.0
-agent.jar=port=7802,host=0.0.0.0 -Dcom.sun.management.jmxremote
-Dcom.sun.management.jmxremote.port=8002
-Dcom.sun.management.jmxremote.ssl=false
-Dcom.sun.management.jmxremote.password.file=$HADOOP_HOME/conf/jmxremote.p
assword"
You can choose a different port for JMX (8002 above) and Jolokia (7802).
If you have an internal IP to lock Jolokia onto you can replace the "catch
all" 0.0.0.0 by your own IP. Notice this IP needs to be accessible from
the telegraf plugin. You can use the option '-
Dcom.sun.management.jmxremote.authenticate=false' if you don't want to
authenticate. Use at your own risk.
```

DataNode

To configure the DataNodes to expose the Jolokia API, you can setup the following in <HADOOP_HOME>/etc/hadoop/hadoop-env.sh:

```
export HADOOP_DATANODE_OPTS="$HADOOP_DATANODE_OPTS
-javaagent:/opt/hadoop/lib/jolokia-jvm-1.6.0
-agent.jar=port=7801,host=0.0.0.0 -Dcom.sun.management.jmxremote
-Dcom.sun.management.jmxremote.port=8001
-Dcom.sun.management.jmxremote.ssl=false
-Dcom.sun.management.jmxremote.password.file=$HADOOP_HOME/conf/jmxremote.p
assword"
You can choose a different port for JMX (8001 above) and Jolokia (7801).
If you have an internal IP to lock Jolokia onto you can replace the "catch
all" 0.0.0.0 by your own IP. Notice this IP needs to be accessible from
the telegraf plugin. You can use the option '-
Dcom.sun.management.jmxremote.authenticate=false' if you don't want to
authenticate. Use at your own risk.
```

ResourceManager

To configure the ResourceManager to expose the Jolokia API, you can setup the following in <HADOOP_HOME>/etc/hadoop/hadoop-env.sh:

```
export YARN_RESOURCEMANAGER_OPTS="$YARN_RESOURCEMANAGER_OPTS
-javaagent:/opt/hadoop/lib/jolokia-jvm-1.6.0
-agent.jar=port=7803,host=0.0.0.0 -Dcom.sun.management.jmxremote
-Dcom.sun.management.jmxremote.port=8003
-Dcom.sun.management.jmxremote.ssl=false
-Dcom.sun.management.jmxremote.password.file=$HADOOP_HOME/conf/jmxremote.p
assword"
You can choose a different port for JMX (8003 above) and Jolokia (7803).
If you have an internal IP to lock Jolokia onto you can replace the "catch
all" 0.0.0.0 by your own IP. Notice this IP needs to be accessible from
the telegraf plugin. You can use the option '-
Dcom.sun.management.jmxremote.authenticate=false' if you don't want to
authenticate. Use at your own risk.
```

NodeManager

To configure the NodeManagers to expose the Jolokia API, you can setup the following in <HADOOP_HOME>/etc/hadoop/hadoop-env.sh:

```
export YARN_NODEMANAGER_OPTS="$YARN_NODEMANAGER_OPTS
-javaagent:/opt/hadoop/lib/jolokia-jvm-1.6.0
-agent.jar=port=7804,host=0.0.0.0 -Dcom.sun.management.jmxremote
-Dcom.sun.management.jmxremote.port=8004
-Dcom.sun.management.jmxremote.ssl=false
-Dcom.sun.management.jmxremote.password.file=$HADOOP_HOME/conf/jmxremote.p
assword"
You can choose a different port for JMX (8004 above) and Jolokia (7804).
If you have an internal IP to lock Jolokia onto you can replace the "catch
all" 0.0.0.0 by your own IP. Notice this IP needs to be accessible from
the telegraf plugin. You can use the option '-
Dcom.sun.management.jmxremote.authenticate=false' if you don't want to
authenticate. Use at your own risk.
```

JobHistoryServer

To configure the JobHistoryServer to expose the Jolokia API, you can setup the following in <HADOOP_HOME>/etc/hadoop/hadoop-env.sh:

```
export HADOOP_JOB_HISTORYSERVER_OPTS="$HADOOP_JOB_HISTORYSERVER_OPTS
-javaagent:/opt/hadoop/lib/jolokia-jvm-1.6.0
-agent.jar=port=7805,host=0.0.0.0 -Dcom.sun.management.jmxremote
-Dcom.sun.management.jmxremote.port=8005
-Dcom.sun.management.jmxremote.password.file=$HADOOP_HOME/conf/jmxremote.p
assword"
You can choose a different port for JMX (8005 above) and Jolokia (7805).
If you have an internal IP to lock Jolokia onto you can replace the "catch
all" 0.0.0.0 by your own IP. Notice this IP needs to be accessible from
the telegraf plugin. You can use the option '-
Dcom.sun.management.jmxremote.authenticate=false' if you don't want to
authenticate. Use at your own risk.
```

Objects and Counters

The following objects and their counters are collected:

Object:	Identifiers:	Attributes:
Hadoop Secondary NameNode	Cluster Namespace Server	Node Name Node IP Compile Info Version
Hadoop NodeManager	Cluster Namespace Server	Node Name Node IP
Hadoop ResourceManager	Cluster Namespace Server	Node Name Node IP
Hadoop DataNode	Cluster Namespace Server	Node Name Node IP Cluster ID Version
Hadoop NameNode	Cluster Namespace Server	Node Name Node IP Transaction ID Last Written Time Since Last Loaded Edits HA State File System State Block Pool ID Cluster ID Compile Info Distinct Version Count Version
Hadoop JobHistoryServer	Cluster Namespace Server	Node Name Node IP

Troubleshooting

Additional information may be found from the Support page.

HAProxy Data Collector

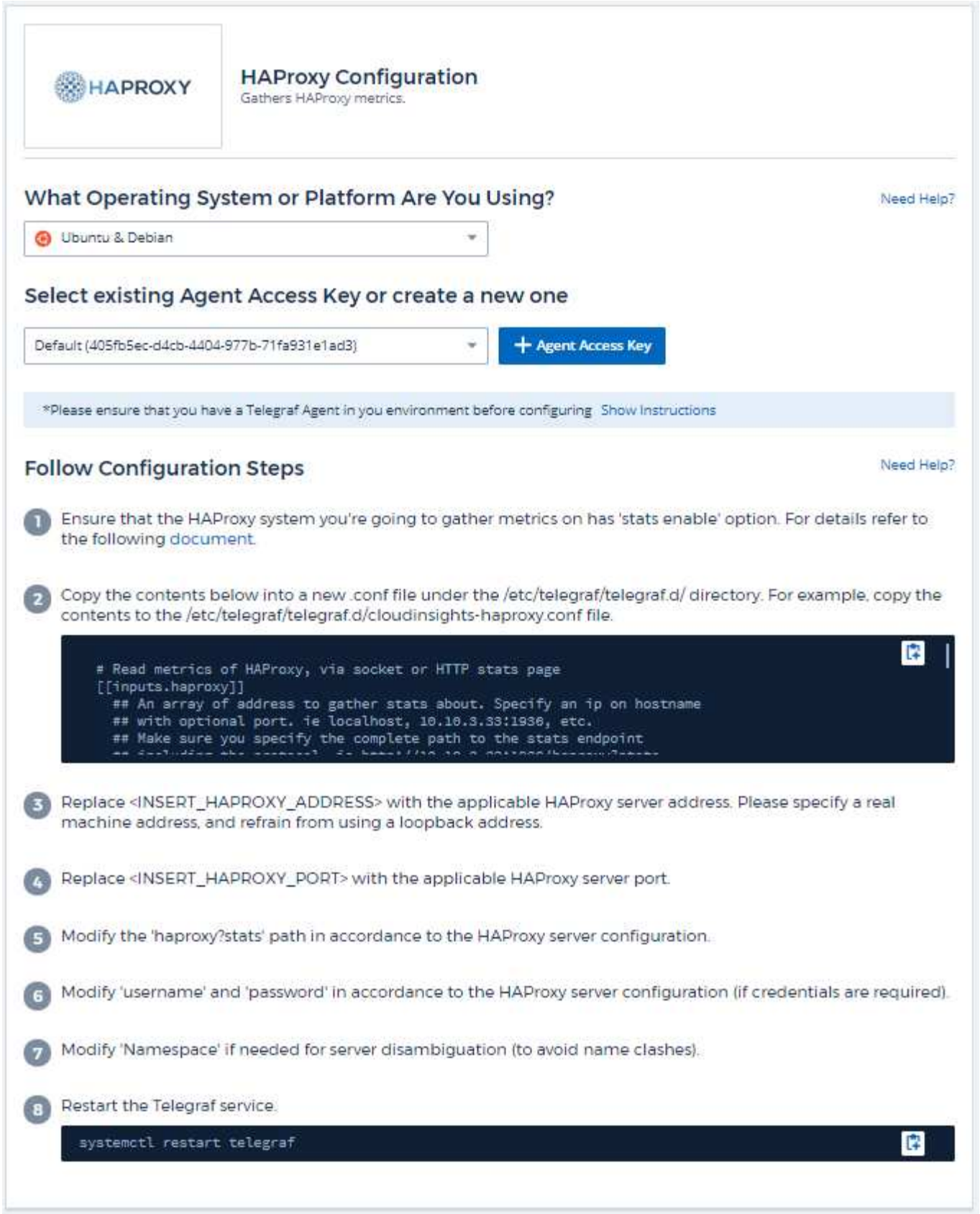
Data Infrastructure Insights uses this data collector to gather metrics from HAProxy.

Installation

1. From **Observability > Collectors**, click **+Data Collector**. Choose HAProxy.

   Select the Operating System or Platform on which the Telegraf agent is installed.

2. If you haven't already installed an Agent for collection, or you wish to install an Agent for a different Operating System or Platform, click *Show Instructions* to expand the Agent installation instructions.
3. Select the Agent Access Key for use with this data collector. You can add a new Agent Access Key by clicking the **+ Agent Access Key** button. Best practice: Use a different Agent Access Key only when you want to group data collectors, for example, by OS/Platform.
4. Follow the configuration steps to configure the data collector. The instructions vary depending on the type of Operating System or Platform you are using to collect data.



Setup

Telegraf's plugin for HAProxy relies on HAProxy Stats enablement. This is a configuration built into HAProxy but it is not enabled out of the box. When enabled, HAProxy will expose an HTML endpoint that can be viewed

on your browser or scraped for extraction of status of all HAProxy configurations.

Compatibility:

Configuration was developed against HAProxy version 1.9.4.

Setting Up:

To enable stats, edit your haproxy configuration file and add the the following lines after the 'defaults' section, using your own user/password and/or haproxy URL:

```
stats enable
stats auth myuser:mypassword
stats uri  /haproxy?stats
```

The following is a simplified example configuration file with stats enabled:

```
global
  daemon
  maxconn 256

defaults
  mode http
  stats enable
  stats uri /haproxy?stats
  stats auth myuser:mypassword
  timeout connect 5000ms
  timeout client 50000ms
  timeout server 50000ms

frontend http-in
  bind *:80
  default_backend servers

frontend http-in9080
  bind *:9080
  default_backend servers_2

backend servers
  server server1 10.128.0.55:8080 check ssl verify none
  server server2 10.128.0.56:8080 check ssl verify none

backend servers_2
  server server3 10.128.0.57:8080 check ssl verify none
  server server4 10.128.0.58:8080 check ssl verify none
```

For complete and up to date instructions, see the HAProxy documentation.

Objects and Counters

The following objects and their counters are collected:

Object:	Identifiers:	Attributes:	Datapoints:
HAProxy Frontend	Namespace Address Proxy	Node IP Node Name Proxy ID Mode Process id Sessions Rate Limit Server id Sessions Limit Status	Bytes In Bytes Out Cache Hits Cache Lookups Compression Bytes Bypassed Compression Bytes In Compression Bytes Out Compression Responses Connection Rate Connection Rate Max Connections Total Requests Denied by Connection Rule Requests Denied by Security Concerns Responses Denied by Security Concerns Requests Denied by Session Rule Requests Errors Responses 1xx Responses 2xx Responses 3xx Responses 4xx Responses 5xx Responses Other Requests Intercepted Sessions Rate Sessions Rate Max Requests Rate Requests Rate Max Requests Total Sessions Sessions Max Sessions Total Requests Rewrites

Object:	Identifiers:	Attributes:	Datapoints:
HAProxy Server	Namespace Address Proxy Server	Node IP Node Name Check Time to Finish Check Fall Configuration Check Health Value Check Rise Configuration Check Status Proxy ID Last Change Time Last Session Time Mode Process id Server id Status Weight	Active Servers Backup Servers Bytes In Bytes Out Check Downs Check Fails Client Aborts Connections Connection Average Time Downtime Total Denied Responses Connection Errors Response Errors Responses 1xx Responses 2xx Responses 3xx Responses 4xx Responses 5xx Responses Other Server Selected Total Queue Current Queue Max Queue Average Time Sessions per Second Sessions per Second Max Connection Reuse Response Time Average Sessions Sessions Max Server Transfer Aborts Sessions Total Sessions Total Time Average Requests Redispatches Requests Retries Requests Rewrites

Object:	Identifiers:	Attributes:	Datapoints:
HAProxy Backend	Namespace Address Proxy	Node IP Node Name Proxy ID Last Change Time Last Session Time Mode Process id Server id Sessions Limit Status Weight	Active Servers Backup Servers Bytes In Bytes Out Cache Hits Cache Lookups Check Downs Client Aborts Compression Bytes Bypassed Compression Bytes In Compression Bytes Out Compression Responses Connections Connection Average Time Downtime Total Requests Denied by Security Concerns Responses Denied by Security Concerns Connection Errors Response Errors Responses 1xx Responses 2xx Responses 3xx Responses 4xx Responses 5xx Responses Other Server Selected Total Queue Current Queue Max Queue Average Time Sessions per Second Sessions per Second Max Requests Total Connection Reuse Response Time Average Sessions Sessions Max Server Transfer Aborts Sessions Total Sessions Total Time Average Requests Redispatches Requests Retries Requests Rewrites

Troubleshooting

Additional information may be found from the Support page.

JVM Data Collector

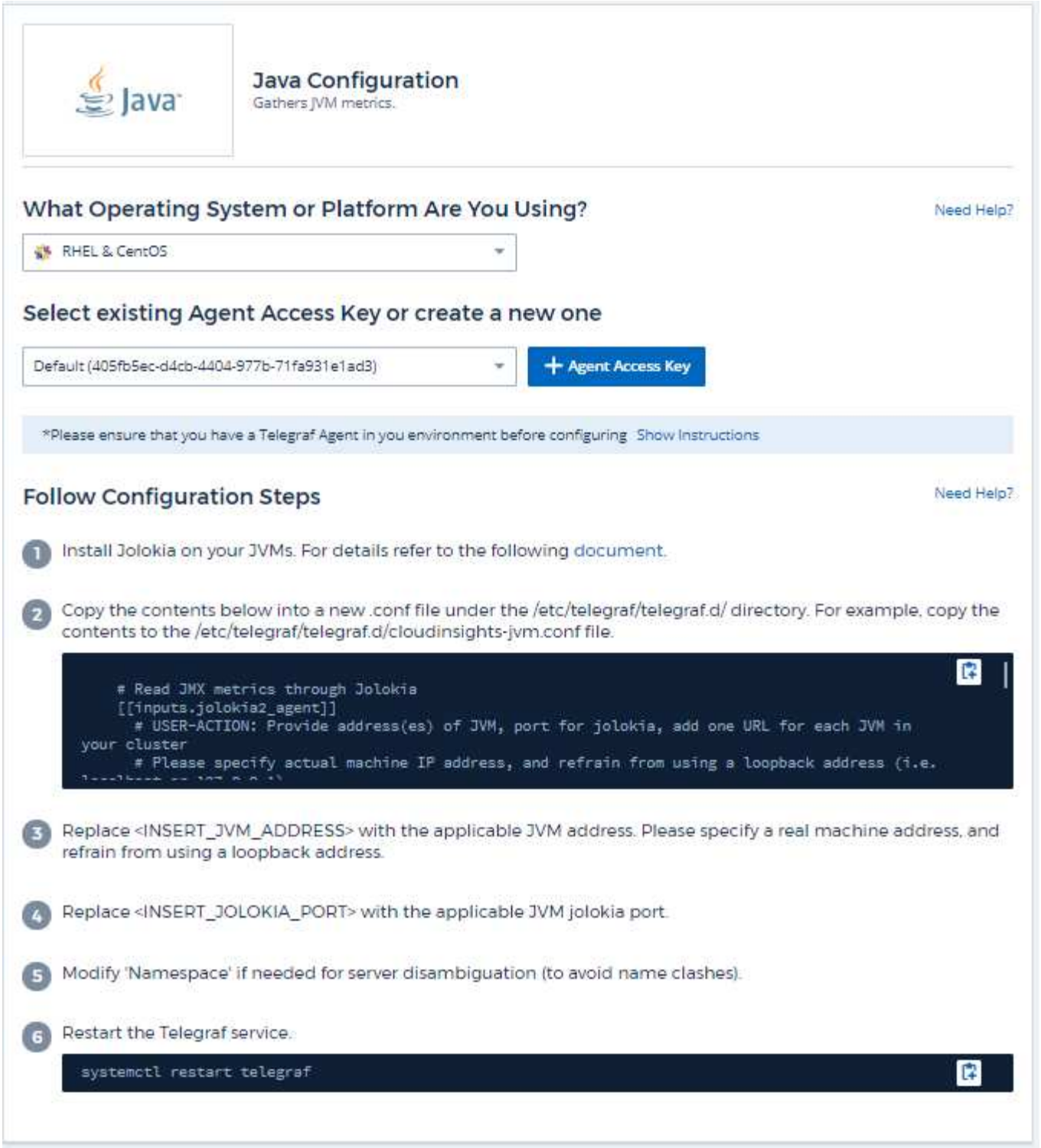
Data Infrastructure Insights uses this data collector to gather metrics from JVM.

Installation

1. From **Observability > Collectors**, click **+Data Collector**. Choose JVM.

   Select the Operating System or Platform on which the Telegraf agent is installed.

2. If you haven't already installed an Agent for collection, or you wish to install an Agent for a different Operating System or Platform, click *Show Instructions* to expand the Agent installation instructions.
3. Select the Agent Access Key for use with this data collector. You can add a new Agent Access Key by clicking the **+ Agent Access Key** button. Best practice: Use a different Agent Access Key only when you want to group data collectors, for example, by OS/Platform.
4. Follow the configuration steps to configure the data collector. The instructions vary depending on the type of Operating System or Platform you are using to collect data.



Setup

Information may be found in JVM documentation.

Objects and Counters

The following objects and their counters are collected:

Object:	Identifiers:	Attributes:	Datapoints:
JVM	Namespace JVM	OS Architecture OS Name OS Version Runtime Specification Runtime Specification Vendor Runtime Specification Version Uptime Runtime VM Name Runtime VM Vendor Runtime VM Version Node Name Node IP	Class Loaded Class Loaded Total Class Unloaded Memory Heap Committed Memory Heap Init Memory Heap Used Max Memory Heap Used Memory Non Heap Committed Memory Non Heap Init Memory Non Heap Max Memory Non Heap Used Memory Objects Pending Finalization OS Processors Available OS Committed Virtual Memory Size OS Free Physical Memory Size OS Free Swap Space Size OS Max File Descriptor Count OS Open File Descriptors Count OS Processor CPU Load OS Processor CPU Time OS System CPU Load OS System Load Average OS Total Physical Memory Size OS Total Swap Space Size Thread Daemon Count Thread Peak Count Thread Count Thread Total Started Count Garbage Collector Copy Collection Count Garbage Collector Copy Collection Time Garbage Collector Mark-sweep Collection Count Garbage Collector Mark-sweep Collection Time Garbage Collector G1 Old Generation Collection Count Garbage Collector G1 Old Generation Collection Time Garbage Collector G1 Young Generation

Troubleshooting

Additional information may be found from the Support page.

Kafka Data Collector

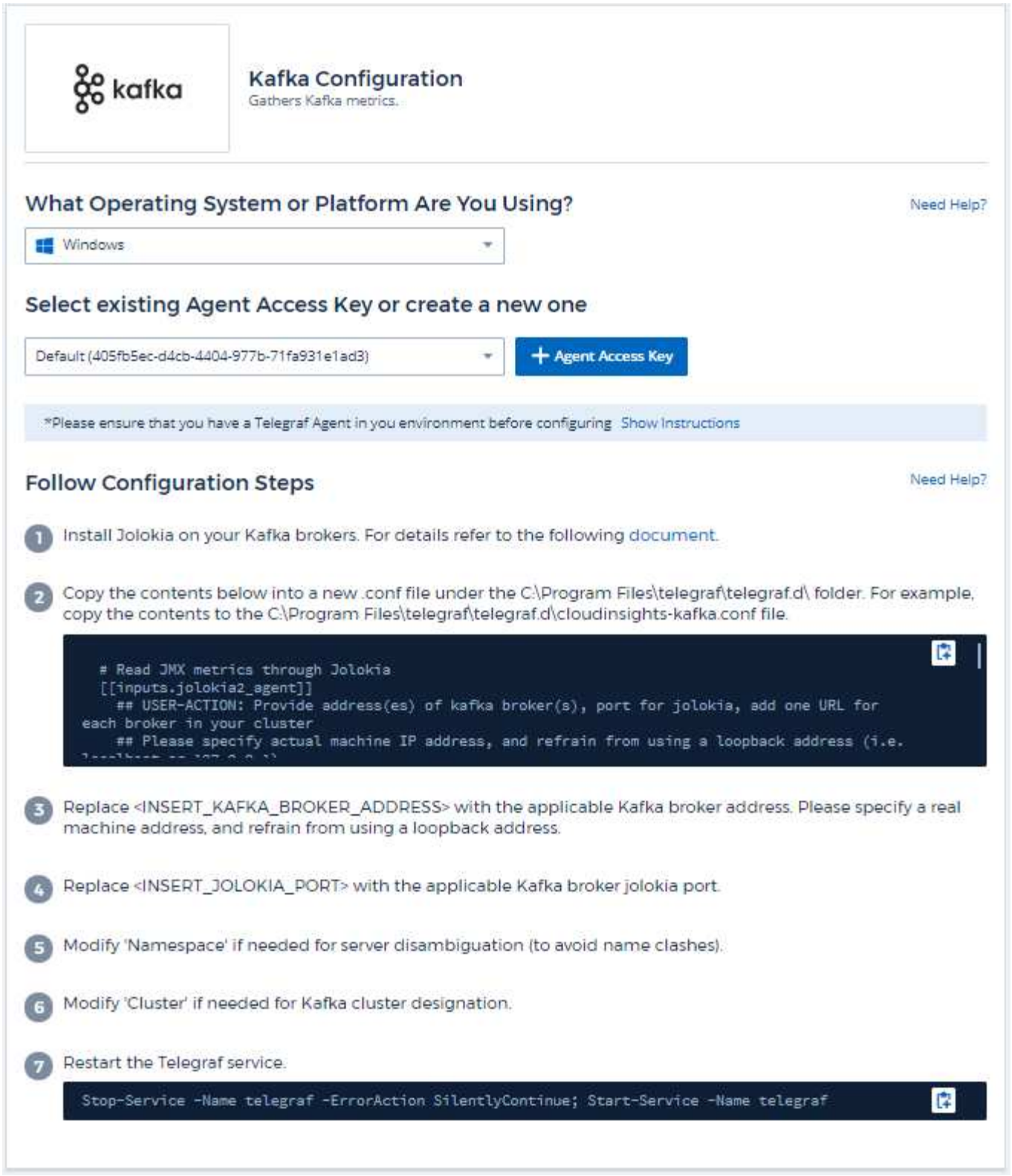
Data Infrastructure Insights uses this data collector to gather metrics from Kafka.

Installation

1. From **Observability > Collectors**, click **+Data Collector**. Choose Kafka.

   Select the Operating System or Platform on which the Telegraf agent is installed.

2. If you haven't already installed an Agent for collection, or you wish to install an Agent for a different Operating System or Platform, click *Show Instructions* to expand the Agent installation instructions.
3. Select the Agent Access Key for use with this data collector. You can add a new Agent Access Key by clicking the **+ Agent Access Key** button. Best practice: Use a different Agent Access Key only when you want to group data collectors, for example, by OS/Platform.
4. Follow the configuration steps to configure the data collector. The instructions vary depending on the type of Operating System or Platform you are using to collect data.



Setup

The Kafka plugin is based on the telegraf's Jolokia plugin. As such as a requirement to gather info from all Kafka brokers, JMX needs to be configured and exposed via Jolokia on all components.

Compatibility

Configuration was developed against Kafka version 0.11.0.2.

Setting up

All the instructions below assume your install location for kafka is '/opt/kafka'. You can adapt instructions below to reflect your install location.

Jolokia Agent Jar

A version the Jolokia agent jar file must be downloaded. The version tested against was Jolokia agent 1.6.0.

Instructions below assume that the downloaded jar file (jolokia-jvm-1.6.0-agent.jar) is placed under the location '/opt/kafka/libs/'.

Kafka Brokers

To configure Kafka Brokers to expose the Jolokia API, you can add the following in <KAFKA_HOME>/bin/kafka-server-start.sh, just before the 'kafka-run-class.sh' call:

```
export JMX_PORT=9999
export RMI_HOSTNAME=`hostname -I`
export KAFKA_JMX_OPTS="-javaagent:/opt/kafka/libs/jolokia-jvm-1.6.0-
agent.jar=port=8778,host=0.0.0.0
-Dcom.sun.management.jmxremote.password.file=/opt/kafka/config/jmxremote.p
assword -Dcom.sun.management.jmxremote.ssl=false
-Djava.rmi.server.hostname=$RMI_HOSTNAME
-Dcom.sun.management.jmxremote.rmi.port=$JMX_PORT"
```

Note that example above is using 'hostname -I' to setup the 'RMI_HOSTNAME' environment variable. In multiple IP machines, this will need to be tweaked to gather the IP you care about for RMI connections.

You can choose a different port for JMX (9999 above) and Jolokia (8778). If you have an internal IP to lock Jolokia onto you can replace the "catch all" 0.0.0.0 by your own IP. Notice this IP needs to be accessible from the telegraf plugin. You can use the option '-Dcom.sun.management.jmxremote.authenticate=false' if you don't want to authenticate. Use at your own risk.

Objects and Counters

The following objects and their counters are collected:

Object:	Identifiers:	Attributes:
Kafka Broker	Cluster Namespace Broker	Node Name Node IP

Troubleshooting

Additional information may be found from the Support page.

Kibana Data Collector

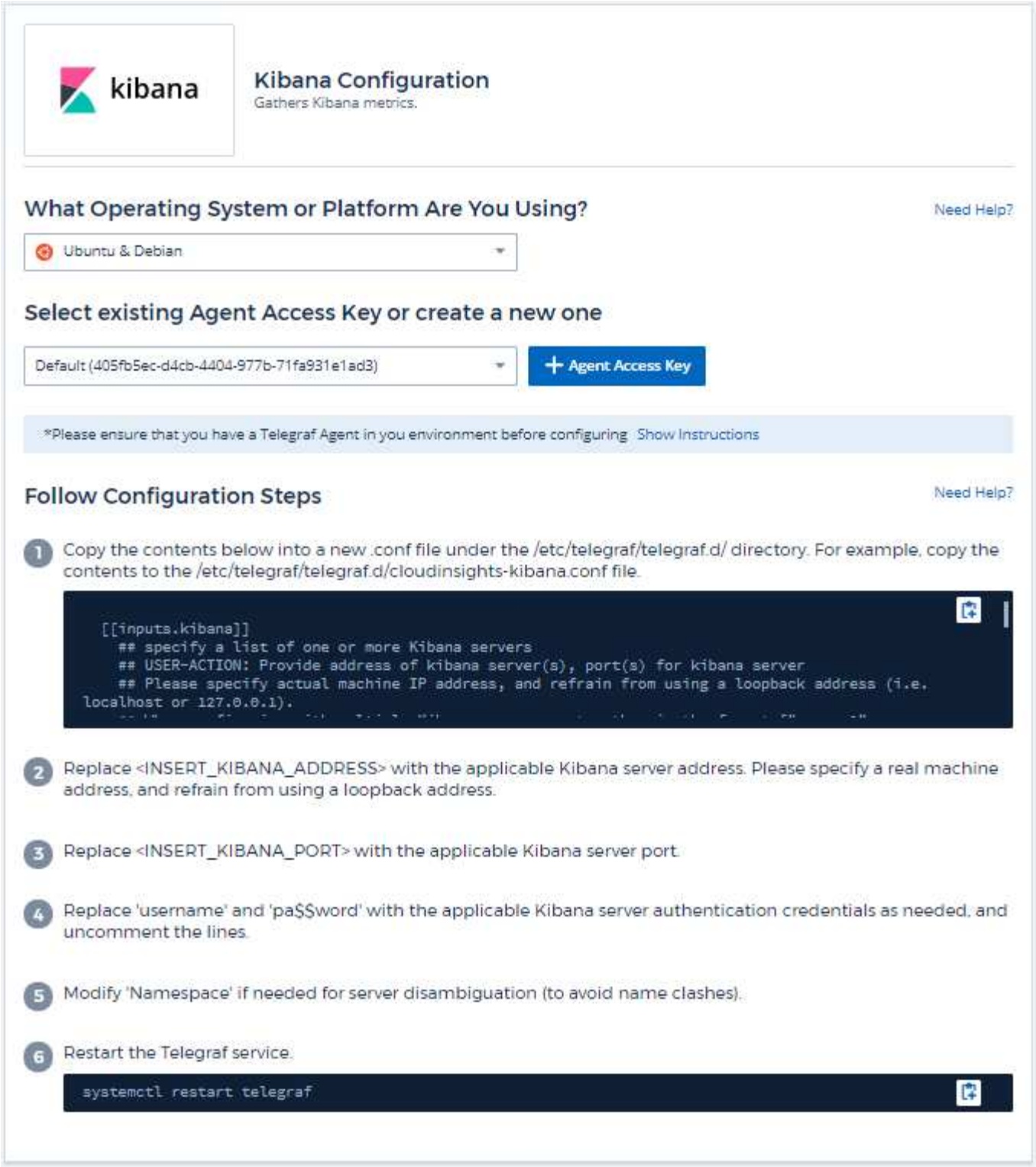
Data Infrastructure Insights uses this data collector to gather metrics from Kibana.

Installation

1. From **Observability > Collectors**, click **+Data Collector**. Choose Kibana.

   Select the Operating System or Platform on which the Telegraf agent is installed.

2. If you haven't already installed an Agent for collection, or you wish to install an Agent for a different Operating System or Platform, click *Show Instructions* to expand the Agent installation instructions.
3. Select the Agent Access Key for use with this data collector. You can add a new Agent Access Key by clicking the **+ Agent Access Key** button. Best practice: Use a different Agent Access Key only when you want to group data collectors, for example, by OS/Platform.
4. Follow the configuration steps to configure the data collector. The instructions vary depending on the type of Operating System or Platform you are using to collect data.



Setup

Information may be found in the Kibana documentation.

Objects and Counters

The following objects and their counters are collected:

Object:	Identifiers:	Attributes:	Datapoints:
Kibana	Namespace Address	Node IP Node Name Version Status	Concurrent Connections Heap Max Heap Used Requests per Second Response Time Average Response Time Max Uptime

Troubleshooting

Additional information may be found from the Support page.

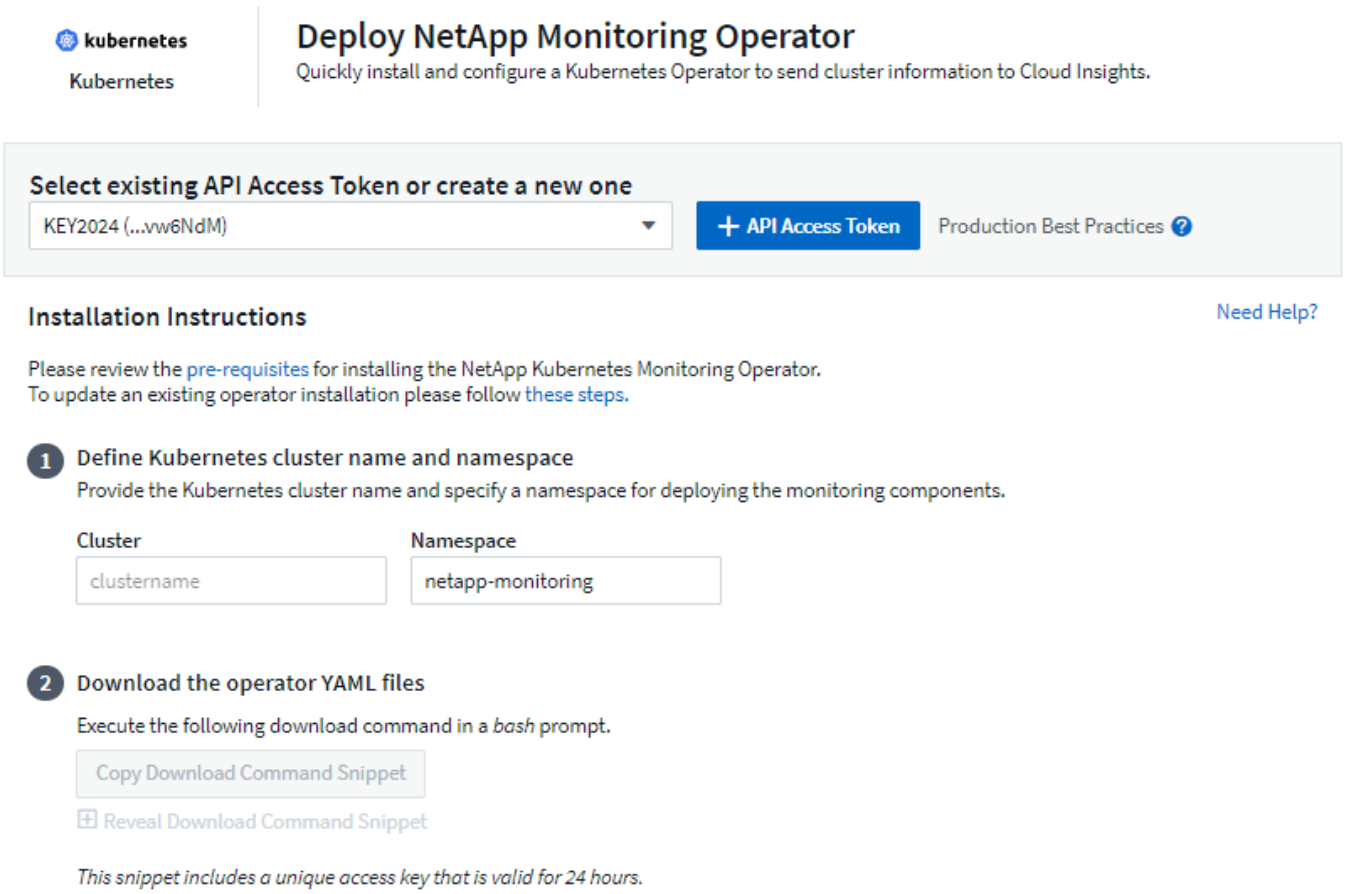
Kubernetes Monitoring Operator Installation and Configuration

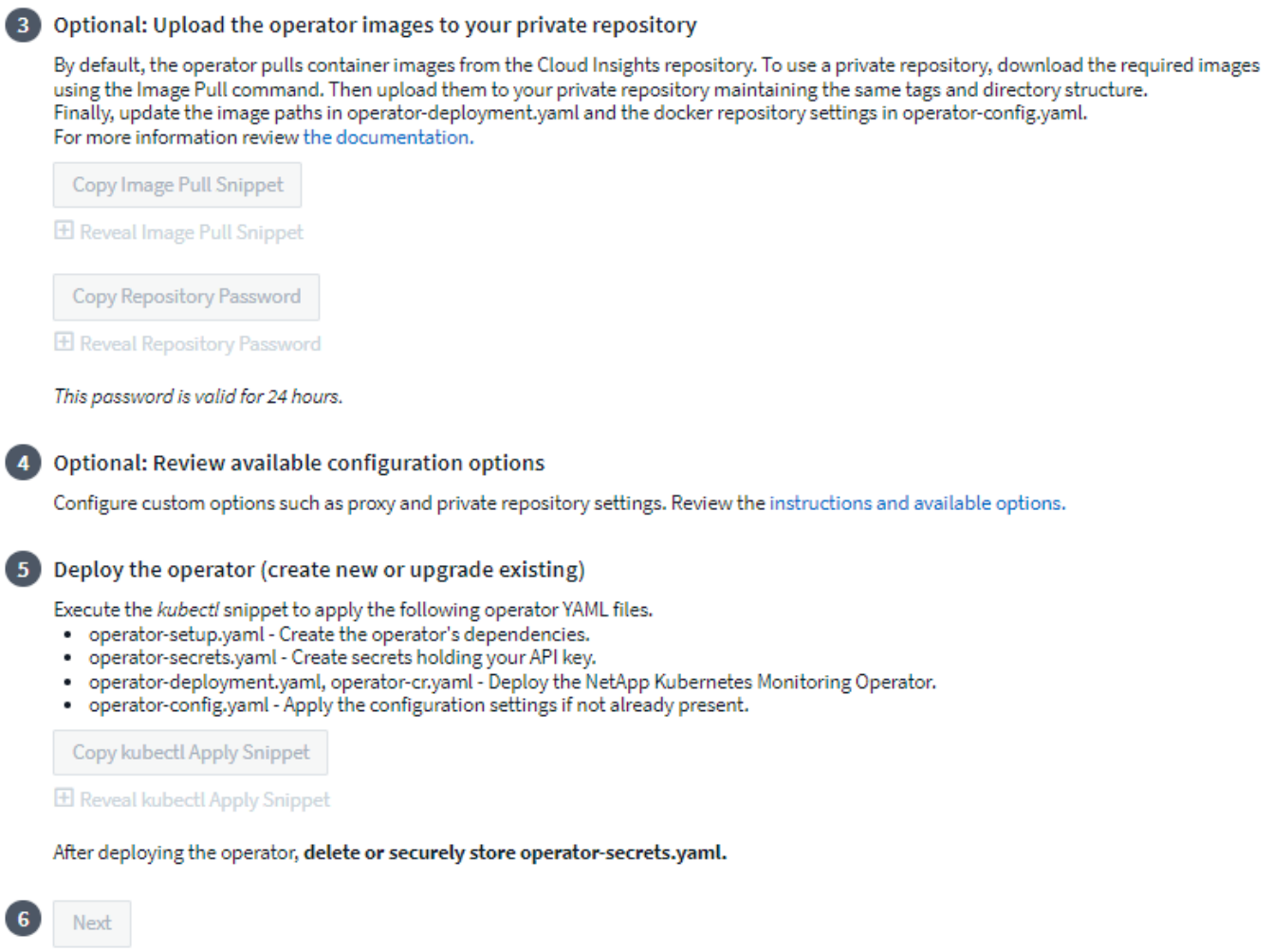
Data Infrastructure Insights offers the **Kubernetes Monitoring Operator** for Kubernetes collection. Navigate to **Kubernetes > Collectors > +Kubernetes Collector** to deploy a new operator.

Before installing the Kubernetes Monitoring Operator

See the Pre-requisites documentation before installing or upgrading the Kubernetes Monitoring Operator.

Installing the Kubernetes Monitoring Operator





Steps to install Kubernetes Monitoring Operator agent on Kubernetes:

1. Enter a unique cluster name and namespace. If you are upgrading from a previous Kubernetes Operator, use the same cluster name and namespace.
2. Once these are entered, you can copy the Download Command snippet to the clipboard.
3. Paste the snippet into a *bash* window and execute it. The Operator installation files will be downloaded. Note that the snippet has a unique key and is valid for 24 hours.
4. If you have a custom or private repository, copy the optional Image Pull snippet, paste it into a *bash* shell and execute it. Once the images have been pulled, copy them to your private repository. Be sure to maintain the same tags and folder structure. Update the paths in *operator-deployment.yaml* as well as the docker repository settings in *operator-config.yaml*.
5. If desired, review available configuration options such as proxy or private repository settings. You can read more about configuration options.
6. When you are ready, deploy the Operator by copying the kubectl Apply snippet, downloading it, and executing it.
7. The installation proceeds automatically. When it is complete, click the *Next* button.
8. When installation is complete, click the *Next* button. Be sure to also delete or securely store the *operator-secrets.yaml* file.

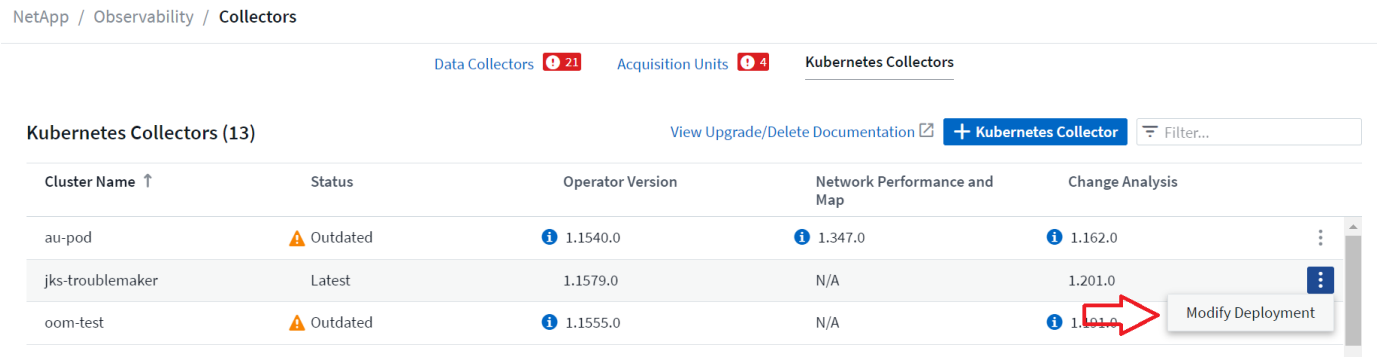
If you are have a custom repository, read about using a custom/private docker repository.

Kubernetes Monitoring Components

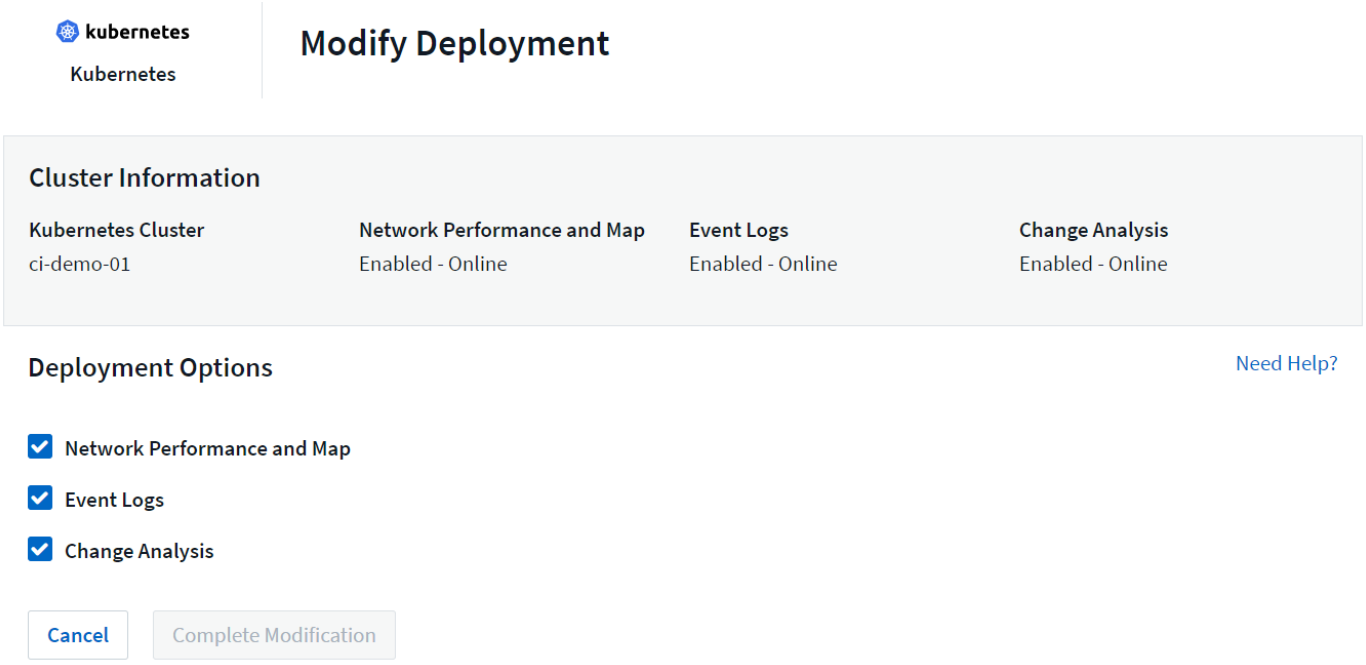
Data Infrastructure Insights Kubernetes Monitoring is comprised of four monitoring components:

- Cluster Metrics
- Network Performance and Map (optional)
- Event Logs (optional)
- Change Analysis (optional)

The optional components above are enabled by default for each Kubernetes collector; if you decide you don't need a component for a particular collector, you can disable it by navigating to **Kubernetes > Collectors** and selecting *Modify Deployment* from the collector's "three dots" menu on the right of the screen.



The screen shows the current state of each component and allows you to disable or enable components for that collector as needed.



Upgrading to the latest Kubernetes Monitoring Operator

DII Push-Button upgrades

You can upgrade the Kubernetes Monitoring Operator through the DII Kubernetes Collectors page. Click on the

menu next to the cluster you would like to upgrade and select *Upgrade*. The operator will verify the image signatures, perform a snapshot of your current installation and perform the upgrade. Within a few minutes you should see the operator Status progress through Upgrade In Progress to Latest. If you encounter an error you can select the Error status for more details and refer to the Push-Button Upgrades Troubleshooting table below.

Push-Button upgrades with private repositories

If your operator is configured to use a private repository please ensure all images required to run the operator and their signatures are available in your repository. If you encounter an error during the upgrade process for missing images simply add them to your repository and retry the upgrade. To upload the image signatures to your repository please use the cosign tool as follows, making sure to upload signatures for all images specified under 3 Optional: Upload the operator images to your private repository > Image Pull Snippet

```
cosign copy example.com/src:v1 example.com/dest:v1
#Example
cosign copy <DII container registry>/netapp-monitoring:<image version>
<private repository>/netapp-monitoring:<image version>
```

Rolling back to a previously running version

If you upgraded using the push-button upgrades feature and encounter any difficulties with the current version of the operator within seven days of the upgrade, you can downgrade to the previously running version using the snapshot created during the upgrade process. Click the menu next to the cluster you would like to roll back and select *Roll back*.

Manual Upgrades

Determine whether an AgentConfiguration exists with the existing Operator (if your namespace is not the default *netapp-monitoring*, substitute the appropriate namespace):

```
kubectl -n netapp-monitoring get agentconfiguration netapp-ci-monitoring-
configuration
```

If an AgentConfiguration exists:

- Install the latest Operator over the existing Operator.
  - Ensure you are pulling the latest container images if you are using a custom repository.

If the AgentConfiguration does not exist:

- Make note of your cluster name as recognized by Data Infrastructure Insights (if your namespace is not the default netapp-monitoring, substitute the appropriate namespace):

  ```
  kubectl -n netapp-monitoring get agent -o
  jsonpath='{.items[0].spec.cluster-name}'
  ```

- Create a backup of the existing Operator (if your namespace is not the default netapp-monitoring,

substitute the appropriate namespace):

```
kubectl -n netapp-monitoring get agent -o yaml > agent_backup.yaml
```

- Uninstall the existing Operator.
- Install the latest Operator.
  - Use the same cluster name.
  - After downloading the latest Operator YAML files, port any customizations found in agent_backup.yaml to the downloaded operator-config.yaml before deploying.
  - Ensure you are pulling the latest container images if you are using a custom repository.

Stopping and Starting the Kubernetes Monitoring Operator

To stop the Kubernetes Monitoring Operator:

```
kubectl -n netapp-monitoring scale deploy monitoring-operator --replicas=0
```

To start the Kubernetes Monitoring Operator:

```
kubectl -n netapp-monitoring scale deploy monitoring-operator --replicas=1
```

Uninstalling

To remove the Kubernetes Monitoring Operator

Note that the default namespace for the Kubernetes Monitoring Operator is "netapp-monitoring". If you have set your own namespace, substitute that namespace in these and all subsequent commands and files.

Newer versions of the monitoring operator can be uninstalled with the following commands:

```
kubectl -n <NAMESPACE> delete agent -l installed-by=nkmo-<NAMESPACE>
kubectl -n <NAMESPACE> delete
clusterrole,clusterrolebinding,crd,svc,deploy,role,rolebinding,secret,sa
-l installed-by=nkmo-<NAMESPACE>
```

If the monitoring operator was deployed in its own dedicated namespace, delete the namespace:

```
kubectl delete ns <NAMESPACE>
```

Note: If the first command returns “No resources found”, use the following instructions to uninstall older versions of the monitoring operator.

Execute each of the following commands in order. Depending on your current installation, some of these

commands may return 'object not found' messages. These messages may be safely ignored.

```
kubectl -n <NAMESPACE> delete agent agent-monitoring-netapp
kubectl delete crd agents.monitoring.netapp.com
kubectl -n <NAMESPACE> delete role agent-leader-election-role
kubectl delete clusterrole agent-manager-role agent-proxy-role agent-
metrics-reader <NAMESPACE>-agent-manager-role <NAMESPACE>-agent-proxy-role
<NAMESPACE>-cluster-role-privileged
kubectl delete clusterrolebinding agent-manager-rolebinding agent-proxy-
rolebinding agent-cluster-admin-rolebinding <NAMESPACE>-agent-manager-
rolebinding <NAMESPACE>-agent-proxy-rolebinding <NAMESPACE>-cluster-role-
binding-privileged
kubectl delete <NAMESPACE>-psp-nkmo
kubectl delete ns <NAMESPACE>
```

If a Security Context Constraint was previously-created:

```
kubectl delete scc telegraf-hostaccess
```

About Kube-state-metrics

The NetApp Kubernetes Monitoring Operator installs its own kube-state-metrics to avoid conflict with any other instances.

For information about Kube-State-Metrics, see this page.

Configuring/Customizing the Operator

These sections contain information on customizing your operator configuration, working with proxy, using a custom or private docker repository, or working with OpenShift.

Configuration Options

Most commonly modified settings can be configured in the *AgentConfiguration* custom resource. You can edit this resource before deploying the operator by editing the *operator-config.yaml* file. This file includes commented-out examples of settings. See the list of available settings for the most recent version of the operator.

You can also edit this resource after the operator has been deployed by using the following command:

```
kubectl -n netapp-monitoring edit AgentConfiguration
```

To determine if your deployed version of the operator supports AgentConfiguration, run the following command:

```
kubectl get crd agentconfigurations.monitoring.netapp.com
```

If you see an “Error from server (NotFound)” message, your operator must be upgraded before you can use the AgentConfiguration.

Configuring Proxy Support

There are two places where you may use a proxy on your tenant in order to install the Kubernetes Monitoring Operator. These may be the same or separate proxy systems:

- Proxy needed during execution of the installation code snippet (using "curl") to connect the system where the snippet is executed to your Data Infrastructure Insights environment
- Proxy needed by the target Kubernetes cluster to communicate with your Data Infrastructure Insights environment

If you use a proxy for either or both of these, in order to install the Kubernetes Operating Monitor you must first ensure that your proxy is configured to allow good communication to your Data Infrastructure Insights environment. If you have a proxy and can access Data Infrastructure Insights from the server/VM from which you wish to install the Operator, then your proxy is likely configured properly.

For the proxy used to install the Kubernetes Operating Monitor, before installing the Operator, set the *http_proxy/https_proxy* environment variables. For some proxy environments, you may also need to set the *no_proxy environment* variable.

To set the variable(s), perform the following steps on your system **before** installing the Kubernetes Monitoring Operator:

1. Set the *https_proxy* and/or *http_proxy* environment variable(s) for the current user:
   a. If the proxy being setup does not have Authentication (username/password), run the following command:

```
export https_proxy=<proxy_server>:<proxy_port>
```

   b. If the proxy being setup does have Authentication (username/password), run this command:

```
export
http_proxy=<proxy_username>:<proxy_password>@<proxy_server>:<proxy_po
rt>
```

For the proxy used for your Kubernetes cluster to communicate with your Data Infrastructure Insights environment, install the Kubernetes Monitoring Operator after reading all of these instructions.

Configure the proxy section of AgentConfiguration in operator-config.yaml before deploying the Kubernetes Monitoring Operator.

```
agent:
  ...
  proxy:
    server: <server for proxy>
    port: <port for proxy>
    username: <username for proxy>
    password: <password for proxy>

    # In the noproxy section, enter a comma-separated list of
    # IP addresses and/or resolvable hostnames that should bypass
    # the proxy
    noproxy: <comma separated list>

    isTelegrafProxyEnabled: true
    isFluentbitProxyEnabled: <true or false> # true if Events Log enabled
    isCollectorsProxyEnabled: <true or false> # true if Network
Performance and Map enabled
    isAuProxyEnabled: <true or false> # true if AU enabled
  ...
...
```

Using a custom or private docker repository

By default, the Kubernetes Monitoring Operator will pull container images from the Data Infrastructure Insights repository. If you have a Kubernetes cluster used as the target for monitoring, and that cluster is configured to only pull container images from a custom or private Docker repository or container registry, you must configure access to the containers needed by the Kubernetes Monitoring Operator.

Run the “Image Pull Snippet” from the NetApp Monitoring Operator install tile. This command will log into the Data Infrastructure Insights repository, pull all image dependencies for the operator, and log out of the Data Infrastructure Insights repository. When prompted, enter the provided repository temporary password. This command downloads all images used by the operator, including for optional features. See below for which features these images are used for.

Core Operator Functionality and Kubernetes Monitoring

- netapp-monitoring
- ci-kube-rbac-proxy
- ci-ksm
- ci-telegraf
- distroless-root-user

Events Log

- ci-fluent-bit
- ci-kubernetes-event-exporter

Network Performance and Map

- ci-net-observer

Push the operator docker image to your private/local/enterprise docker repository according to your corporate policies. Ensure that the image tags and directory paths to these images in your repository are consistent with those in the Data Infrastructure Insights repository.

Edit the monitoring-operator deployment in operator-deployment.yaml, and modify all image references to use your private Docker repository.

```
image: <docker repo of the enterprise/corp docker repo>/ci-kube-rbac-
proxy:<ci-kube-rbac-proxy version>
image: <docker repo of the enterprise/corp docker repo>/netapp-
monitoring:<version>
```

Edit the AgentConfiguration in operator-config.yaml to reflect the new docker repo location. Create a new imagePullSecret for your private repository, for more details see *https://kubernetes.io/docs/tasks/configure-pod-container/pull-image-private-registry/*

```
agent:
  ...
  # An optional docker registry where you want docker images to be pulled
from as compared to CI's docker registry
  # Please see documentation link here:
xref:{relative_path}task_config_telegraf_agent_k8s.html#using-a-custom-or-
private-docker-repository
  dockerRepo: your.docker.repo/long/path/to/test
  # Optional: A docker image pull secret that maybe needed for your
private docker registry
  dockerImagePullSecret: docker-secret-name
```

OpenShift Instructions

If you are running on OpenShift 4.6 or higher, you must edit the AgentConfiguration in *operator-config.yaml* to enable the *runPrivileged* setting:

```
# Set runPrivileged to true SELinux is enabled on your kubernetes nodes
runPrivileged: true
```

Openshift may implement an added level of security that may block access to some Kubernetes components.

Tolerations and Taints

The *netapp-ci-telegraf-ds*, *netapp-ci-fluent-bit-ds*, and *netapp-ci-net-observer-l4-ds* DaemonSets must schedule a pod on every node in your cluster in order to correctly collect data on all nodes. The operator has been configured to tolerate some well known **taints**. If you have configured any custom taints on your nodes,

thus preventing pods from running on every node, you can create a **toleration** for those taints in the *AgentConfiguration*. If you have applied custom taints to all nodes in your cluster, you must also add the necessary tolerations to the operator deployment to allow the operator pod to be scheduled and executed.

Learn More about Kubernetes Taints and Tolerations.

Return to the **NetApp Kubernetes Monitoring Operator Installation** page

A Note About Secrets

To remove permission for the Kubernetes Monitoring Operator to view secrets cluster-wide, delete the following resources from the *operator-setup.yaml* file before installing:

```
 ClusterRole/netapp-ci<namespace>-agent-secret
 ClusterRoleBinding/netapp-ci<namespace>-agent-secret
```

If this is an upgrade, also delete the resources from your cluster:

```
 kubectl delete ClusterRole/netapp-ci-<namespace>-agent-secret-clusterrole
 kubectl delete ClusterRoleBinding/netapp-ci-<namespace>-agent-secret-
clusterrolebinding
```

If Change Analysis is enabled, modify the *AgentConfiguration* or *operator-config.yaml* to uncomment the change-management section and include *kindsToIgnoreFromWatch: '"secrets"'* under the change-management section. Note the presence and position of single and double quotes in this line.

```
change-management:
  ...
  # # A comma separated list of kinds to ignore from watching from the
default set of kinds watched by the collector
  # # Each kind will have to be prefixed by its apigroup
  # # Example: '"networking.k8s.io.networkpolicies,batch.jobs",
"authorization.k8s.io.subjectaccessreviews"'
  kindsToIgnoreFromWatch: '"secrets"'
  ...
```

Verifying Kubernetes Monitoring Operator Image Signatures

The image for the operator and all related images it deploys are signed by NetApp. You can manually verify the images before installation using the cosign tool, or configure a Kubernetes admission controller. For more details please see the Kubernetes documentation.

The public key used to verify the image signatures is available in the Monitoring Operator install tile under *Optional: Upload the operator images to your private repository > Image Signature Public Key*

To manually verify an image signature, perform the following steps:

1. Copy and run the Image Pull Snippet
2. Copy and enter the Repository Password when prompted
3. Store the Image Signature Public Key (dii-image-signing.pub in the example)
4. Verify the images using cosign. Refer to the following example of cosign usage

```
$ cosign verify --key dii-image-signing.pub --insecure-ignore-sct
--insecure-ignore-tlog <repository>/<image>:<tag>
Verification for <repository>/<image>:<tag> --
The following checks were performed on each of these signatures:
  - The cosign claims were validated
  - The signatures were verified against the specified public key
[{"critical":{"identity":{"docker-
reference":"<repository>/<image>"},"image":{"docker-manifest-
digest":"sha256:<hash>"},"type":"cosign container image
signature"},"optional":null}]
```

Troubleshooting

Some things to try if you encounter problems setting up the Kubernetes Monitoring Operator:

Problem:	Try this:
I do not see a hyperlink/connection between my Kubernetes Persistent Volume and the corresponding back-end storage device. My Kubernetes Persistent Volume is configured using the hostname of the storage server.	Follow the steps to uninstall the existing Telegraf agent, then re-install the latest Telegraf agent. You must be using Telegraf version 2.0 or later, and your Kubernetes cluster storage must be actively monitored by Data Infrastructure Insights.

Problem:	Try this:
I’m seeing messages in the logs resembling the following: E0901 15:21:39.962145 1 reflector.go:178] k8s.io/kube-state-metrics/internal/store/builder.go:352: Failed to list *v1.MutatingWebhookConfiguration: the server could not find the requested resource E0901 15:21:43.168161 1 reflector.go:178] k8s.io/kube-state-metrics/internal/store/builder.go:352: Failed to list *v1.Lease: the server could not find the requested resource (get leases.coordination.k8s.io) etc.	These messages may occur if you are running kube-state-metrics version 2.0.0 or above with Kubernetes versions below 1.20. To get the Kubernetes version: *kubectl version* To get the kube-state-metrics version: *kubectl get deploy/kube-state-metrics -o jsonpath='{..image}'* To prevent these messages from happening, users can modify their kube-state-metrics deployment to disable the following Leases: *mutatingwebhookconfigurations validatingwebhookconfigurations volumeattachments resources* More specifically, they can use the following CLI argument: resources=certificatesigningrequests,configmaps,cronjobs,daemonsets, deployments,endpoints,horizontalpodautoscalers,ingresses,jobs,limitranges, namespaces,networkpolicies,nodes,persistentvolumeclaims,persistentvolumes, poddisruptionbudgets,pods,replicasets,replicationcontrollers,resourcequotas, secrets,services,statefulsets,storageclasses The default resource list is: "certificatesigningrequests,configmaps,cronjobs,daemonsets,deployments, endpoints,horizontalpodautoscalers,ingresses,jobs,leases,limitranges, mutatingwebhookconfigurations,namespaces,networkpolicies,nodes, persistentvolumeclaims,persistentvolumes,poddisruptionbudgets,pods,replicasets, replicationcontrollers,resourcequotas,secrets,services,statefulsets,storageclasses, validatingwebhookconfigurations,volumeattachments"

Problem:	Try this:
I see error messages from Telegraf resembling the following, but Telegraf does start up and run: Oct 11 14:23:41 ip-172-31-39-47 systemd[1]: Started The plugin-driven server agent for reporting metrics into InfluxDB. Oct 11 14:23:41 ip-172-31-39-47 telegraf[1827]: time="2021-10-11T14:23:41Z" level=error msg="failed to create cache directory. /etc/telegraf/.cache/snowflake, err: mkdir /etc/telegraf/.ca che: permission denied. ignored\n" func="gosnowflake.(*defaultLogger).Errorf" file="log.go:120" Oct 11 14:23:41 ip-172-31-39-47 telegraf[1827]: time="2021-10-11T14:23:41Z" level=error msg="failed to open. Ignored. open /etc/telegraf/.cache/snowflake/ocsp_response_cache.json: no such file or directory\n" func="gosnowflake.(*defaultLogger).Errorf" file="log.go:120" Oct 11 14:23:41 ip-172-31-39-47 telegraf[1827]: 2021-10-11T14:23:41Z I! Starting Telegraf 1.19.3	This is a known issue. Refer to This GitHub article for more details. As long as Telegraf is up and running, users can ignore these error messages.
On Kubernetes, my Telegraf pod(s) are reporting the following error: "Error in processing mountstats info: failed to open mountstats file: /hostfs/proc/1/mountstats, error: open /hostfs/proc/1/mountstats: permission denied"	If SELinux is enabled and enforcing, it is likely preventing the Telegraf pod(s) from accessing the /proc/1/mountstats file on the Kubernetes node. To overcome this restriction, edit the agentconfiguration, and enable the runPrivileged setting. For more details, refer to the OpenShift Instructions.
On Kubernetes, my Telegraf ReplicaSet pod is reporting the following error: [inputs.prometheus] Error in plugin: could not load keypair /etc/kubernetes/pki/etcd/server.crt:/etc/kubernetes/pki/etcd/server.key: open /etc/kubernetes/pki/etcd/server.crt: no such file or directory	The Telegraf ReplicaSet pod is intended to run on a node designated as a master or for etcd. If the ReplicaSet pod is not running on one of these nodes, you will get these errors. Check to see if your master/etcd nodes have taints on them. If they do, add the necessary tolerations to the Telegraf ReplicaSet, telegraf-rs. For example, edit the ReplicaSet… kubectl edit rs telegraf-rs …and add the appropriate tolerations to the spec. Then, restart the ReplicaSet pod.

Problem:	Try this:
I have a PSP/PSA environment. Does this affect my monitoring operator?	If your Kubernetes cluster is running with Pod Security Policy (PSP) or Pod Security Admission (PSA) in place, you must upgrade to the latest Kubernetes Monitoring Operator. Follow these steps to upgrade to the current Operator with support for PSP/PSA: 1. Uninstall the previous monitoring operator: kubectl delete agent agent-monitoring-netapp -n netapp-monitoring kubectl delete ns netapp-monitoring kubectl delete crd agents.monitoring.netapp.com kubectl delete clusterrole agent-manager-role agent-proxy-role agent-metrics-reader kubectl delete clusterrolebinding agent-manager-rolebinding agent-proxy-rolebinding agent-cluster-admin-rolebinding 2. Install the latest version of the monitoring operator.
I ran into issues trying to deploy the Operator, and I have PSP/PSA in use.	1. Edit the agent using the following command: kubectl -n <name-space> edit agent 2. Mark 'security-policy-enabled' as 'false'. This will disable Pod Security Policies and Pod Security Admission and allow the Operator to deploy. Confirm by using the following commands: kubectl get psp (should show Pod Security Policy removed) kubectl get all -n <namespace> | grep -i psp (should show that nothing is found)
"ImagePullBackoff" errors seen	These errors may be seen if you have a custom or private docker repository and have not yet configured the Kubernetes Monitoring Operator to properly recognize it. Read more about configuring for custom/private repo.

Problem:	Try this:

I am having an issue with my monitoring-operator deployment, and the current documentation does not help me resolve it.	Capture or otherwise note the output from the following commands, and contact the Technical Support team.
	
	`kubectl -n netapp-monitoring get all`
	`kubectl -n netapp-monitoring describe all`
	`kubectl -n netapp-monitoring logs <monitoring-operator-pod> --all-containers=true`
	`kubectl -n netapp-monitoring logs <telegraf-pod> --all-containers=true`

net-observer (Workload Map) pods in Operator namespace are in CrashLoopBackOff	These pods correspond to Workload Map data collector for Network Observability. Try these: • Check the logs of one of the pods to confirm minimum kernel version. For example: --- {"ci-tenant-id":"your-tenant-id","collector-cluster":"your-k8s-cluster-name","environment":"prod","level":"error","msg":"failed in validation. Reason: kernel version 3.10.0 is less than minimum kernel version of 4.18.0","time":"2022-11-09T08:23:08Z"} --- • Net-observer pods requires the Linux kernel version to be at least 4.18.0. Check the kernel version using the command “uname -r” and ensure they are >= 4.18.0
Pods are running in Operator namespace (default: netapp-monitoring), but no data is shown in UI for workload map or Kubernetes metrics in Queries	Check the time setting on the nodes of the K8S cluster. For accurate audit and data reporting, it is strongly recommended to synchronize the time on the Agent machine using Network Time Protocol (NTP) or Simple Network Time Protocol (SNTP).
Some of the net-observer pods in Operator namespace are in Pending state	Net-observer is a DaemonSet and runs a pod in each Node of the k8s cluster. • Note the pod which is in Pending state, and check if it is experiencing a resource issue for CPU or memory. Ensure the required memory and CPU is available in the node.

Problem:	Try this:
I'm seeing the following in my logs immediately after installing the Kubernetes Monitoring Operator: [inputs.prometheus] Error in plugin: error making HTTP request to http://kube-state-metrics.<namespace>.svc.cluster.local:8080/metrics: Get http://kube-state-metrics.<namespace>.svc.cluster.local:8080/metrics: dial tcp: lookup kube-state-metrics.<namespace>.svc.cluster.local: no such host	This message is typically only seen when a new operator is installed and the *telegraf-rs* pod is up before the *ksm* pod is up. These messages should stop once all pods are running.
I do see not any metrics being collected for the Kubernetes CronJobs that exist in my cluster.	Verify your Kubernetes version (i.e. `kubectl version`). If it is v1.20.x or below, this is an expected limitation. The kube-state-metrics release deployed with the Kubernetes Monitoring Operator only supports v1.CronJob. With Kubernetes 1.20.x and below, the CronJob resource is at v1beta.CronJob. As a result, kube-state-metrics cannot find the CronJob resource.
After installing the operator, the telegraf-ds pods enter CrashLoopBackOff and the pod logs indicate "su: Authentication failure".	Edit the telegraf section in *AgentConfiguration*, and set *dockerMetricCollectionEnabled* to false. For more details, refer to the operator's configuration options. … spec: … telegraf: … - name: docker run-mode: - DaemonSet substitutions: - key: DOCKER_UNIX_SOCK_PLACEHOLDER value: unix:///run/docker.sock … …
I see repeating error messages resembling the following in my Telegraf logs: E! [agent] Error writing to outputs.http: Post "https://<tenant_url>/rest/v1/lake/ingest/influxdb": context deadline exceeded (Client.Timeout exceeded while awaiting headers)	Edit the telegraf section in *AgentConfiguration*, and increase *outputTimeout* to 10s. For more details, refer to the operator's configuration options.
I'm missing *involvedobject* data for some Event Logs.	Be sure you have followed the steps in the Permissions section above.

Problem:	Try this:
Why am I seeing two monitoring operator pods running, one named netapp-ci-monitoring-operator-<pod> and the other named monitoring-operator-<pod>?	As of October 12, 2023, Data Infrastructure Insights has refactored the operator to better serve our users; for those changes to be fully adopted, you must remove the old operator and install the new one.

My kubernetes events unexpectedly stopped reporting to Data Infrastructure Insights.	Retrieve the name of the event-exporter pod:
	`` `kubectl -n netapp-monitoring get pods |grep event-exporter |awk '{print $1}' |sed 's/event-exporter./event-exporter/'` ``
	It should be either "netapp-ci-event-exporter" or "event-exporter". Next, edit the monitoring agent `kubectl -n netapp-monitoring edit agent`, and set the value for LOG_FILE to reflect the appropriate event-exporter pod name found in the previous step. More specifically, LOG_FILE should be set to either "/var/log/containers/netapp-ci-event-exporter.log" or "/var/log/containers/event-exporter*.log"
	`fluent-bit:`
	`...`
	`- name: event-exporter-ci`
	`substitutions:`
	`- key: LOG_FILE`
	`values:`
	`- /var/log/containers/netapp-ci-event-exporter*.log`
	`...`
	Alternatively, one can also uninstall and reinstall the agent.

I'm seeing pod(s) deployed by the Kubernetes Monitoring Operator crash because of insufficient resources.	Refer to the Kubernetes Monitoring Operator configuration options to increase the CPU and/or memory limits as needed.
A missing image or invalid configuration caused the netapp-ci-kube-state-metrics pods to fail to startup or become ready. Now the StatefulSet is stuck and configuration changes are not being applied to the netapp-ci-kube-state-metrics pods.	The StatefulSet is in a broken state. After fixing any configuration problems bounce the netapp-ci-kube-state-metrics pods.

Problem:	Try this:
netapp-ci-kube-state-metrics pods fail to start after running a Kubernetes Operator upgrade, throwing ErrImagePull (failing to pull the image).	Try resetting the pods manually.
"Event discarded as being older then maxEventAgeSeconds" messages are being observed for my Kubernetes cluster under Log Analysis.	Modify the Operator *agentconfiguration* and increase the *event-exporter-maxEventAgeSeconds* (i.e. to 60s), *event-exporter-kubeQPS* (i.e. to 100), and *event-exporter-kubeBurst* (i.e. to 500). For more details on these configuration options, see the configuration options page.
Telegraf warns of, or crashes because of, insufficient lockable memory.	Try increasing the limit of lockable memory for Telegraf in the underlying operating system/node. If increasing the limit is not an option, modify the NKMO agentconfiguration and set *unprotected* to *true*. This will instruct Telegraf to no attempt to reserve locked memory pages. While this can pose a security risk as decrypted secrets might be swapped out to disk, it allows for execution in environments where reserving locked memory is not possible. For more details on the *unprotected* configuration options, refer to the configuration options page.
I see warning messages from Telegraf resembling the following: *W! [inputs.diskio] Unable to gather disk name for "vdc": error reading /dev/vdc: no such file or directory*	For the Kubernetes Monitoring Operator, these warning message are benign and can be safely ignored. Alternatively, edit the telegraf section in AgentConfiguration, and set *runDsPrivileged* to true. For more details, refer to the operator's configuration options.

Problem:	Try this:

My fluent-bit pod is failing with the following errors:
 	
[2024/10/16 14:16:23] [error] [/src/fluent-bit/plugins/in_tail/tail_fs_inotify.c:360 errno=24] Too many open files
[2024/10/16 14:16:23] [error] failed initialize input tail.0
[2024/10/16 14:16:23] [error] [engine] input initialization failed
	Try to change your *fsnotify* settings in your cluster:
	
	`sudo sysctl fs.inotify.max_user_instances (take note of setting)`
	
	`sudo sysctl fs.inotify.max_user_instances=<something larger than current setting>`
	
	`sudo sysctl fs.inotify.max_user_watches (take note of setting)`
	
	`sudo sysctl fs.inotify.max_user_watches=<something larger than current setting>`
	
	Restart Fluent-bit.
	
	Note: to make these settings persistent across node restarts, you need to put the following lines in */etc/sysctl.conf*
	
	`fs.inotify.max_user_instances=<something larger than current setting>`
	`fs.inotify.max_user_watches=<something larger than current setting>`

The telegraf DS pods are reporting errors pertaining to the kubernetes input plugin failing to make HTTP requests due to the inability to validate the TLS certificate. For example: E! [inputs.kubernetes] Error in plugin: error making HTTP request to "https://<kubelet_IP>:10250/stats/summary": Get "https://<kubelet_IP>:10250/stats/summary": tls: failed to verify certificate: x509: cannot validate certificate for <kubelet_IP> because it doesn't contain any IP SANs	This will occur if the kubelet is using self-signed certificates, and/or the specified certificate does not include the <kubelet_IP> in the certificates *Subject Alternative Name* list. To resolve this, the user can modify the agent configuration, and set *telegraf:insecureK8sSkipVerify* to *true*. This will configure the telegraf input plugin to skip verification. Alternatively, the user can configure the kubelet for serverTLSBootstrap, which will trigger a certificate request from the 'certificates.k8s.io' API.

Additional information may be found from the Support page or in the Data Collector Support Matrix.

Memcached Data Collector

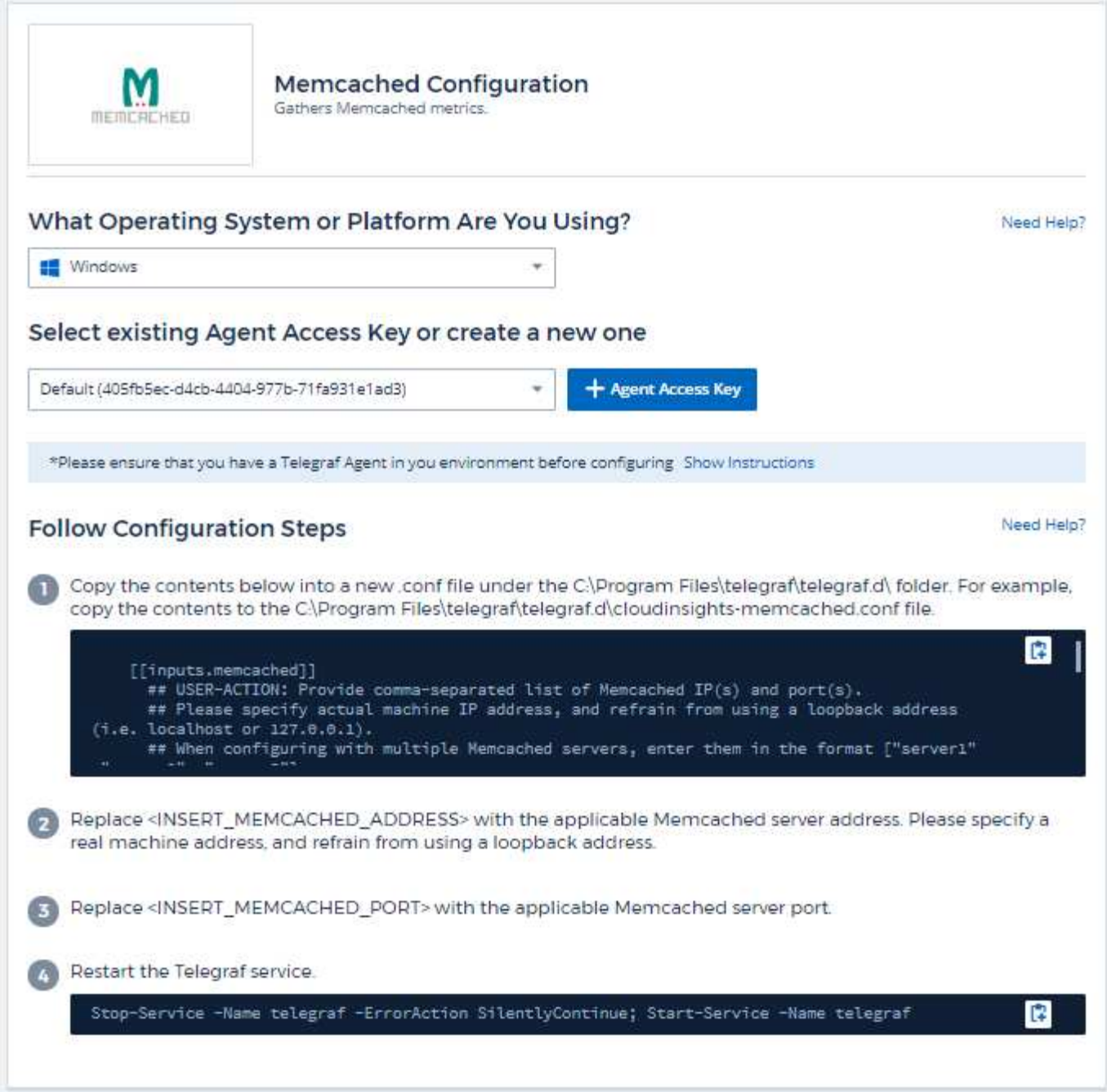
Data Infrastructure Insights uses this data collector to gather metrics from Memcached.

Installation

1. From **Observability > Collectors**, click **+Data Collector**. Choose Memcached.

   Select the Operating System or Platform on which the Telegraf agent is installed.

2. If you haven't already installed an Agent for collection, or you wish to install an Agent for a different Operating System or Platform, click *Show Instructions* to expand the Agent installation instructions.
3. Select the Agent Access Key for use with this data collector. You can add a new Agent Access Key by clicking the **+ Agent Access Key** button. Best practice: Use a different Agent Access Key only when you want to group data collectors, for example, by OS/Platform.
4. Follow the configuration steps to configure the data collector. The instructions vary depending on the type of Operating System or Platform you are using to collect data.



Setup

Information may be found in the Memcached wiki.

Objects and Counters

The following objects and their counters are collected:

Object:	Identifiers:	Attributes:	Datapoints:
Memcached	Namespace Server	Node IP Node Name	Accepting Connections Handled Authentication Requests Failed Authentications Bytes Used Bytes Read (per sec) Bytes Written (per sec) CAS Badval CAS Hits CAS Misses Flush Reqs (per sec) Get Reqs (per sec) Set Reqs (per sec) Touch Reqs (per sec) Connection Yields (per sec) Connection Structures Open Connections Current Stored Items Decr Requests Hits (per sec) Decr Requests Misses (per sec) Delete Requests Hits (per sec) Delete Requests Misses (per sec) Items Evicted Valid Evictions Expired Items Get Hits (per sec) Get Misses (per sec) Used Hash Bytes Hash Is Expanding Hash Power Level Incr Requests Hits (per sec) Incr Requests Misses (per sec) Server Max Bytes Listen Disabled Num Reclaimed Worker Threads Count Total Opened Connections Total Items Stored Touch Hits Touch Misses Server Uptime

Troubleshooting

Additional information may be found from the Support page.

MongoDB Data Collector

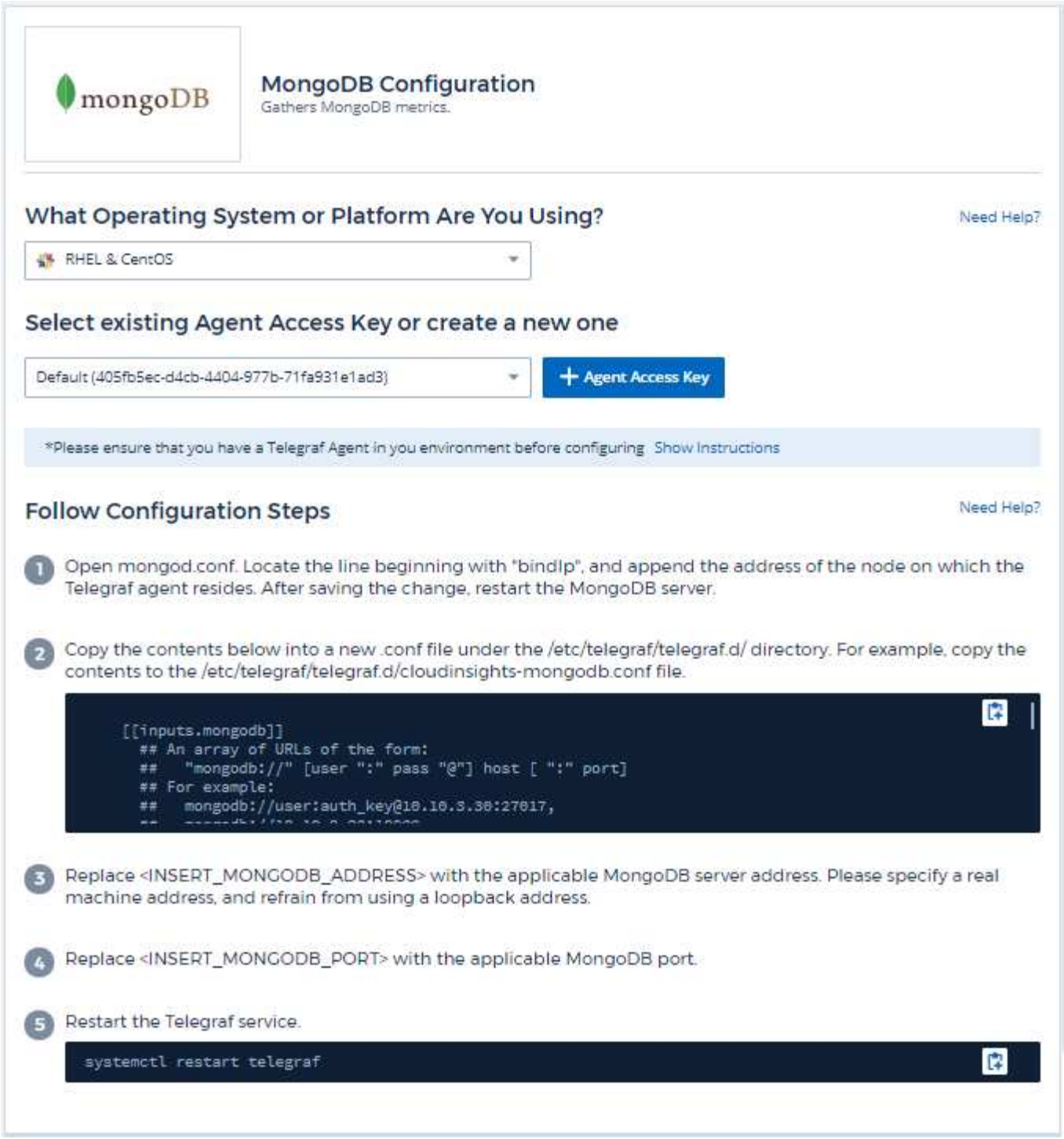
Data Infrastructure Insights uses this data collector to gather metrics from MongoDB.

Installation

1. From **Observability > Collectors**, click **+Data Collector**. Choose MongoDB.

   Select the Operating System or Platform on which the Telegraf agent is installed.

2. If you haven't already installed an Agent for collection, or you wish to install an Agent for a different Operating System or Platform, click *Show Instructions* to expand the Agent installation instructions.
3. Select the Agent Access Key for use with this data collector. You can add a new Agent Access Key by clicking the **+ Agent Access Key** button. Best practice: Use a different Agent Access Key only when you want to group data collectors, for example, by OS/Platform.
4. Follow the configuration steps to configure the data collector. The instructions vary depending on the type of Operating System or Platform you are using to collect data.



Setup

Information may be found in the MongoDB documentation.

Objects and Counters

The following objects and their counters are collected:

Object:	Identifiers:	Attributes:	Datapoints:
MongoDB	Namespace Hostname		
MongoDB Database	Namespace Hostname Database name		

Troubleshooting

Information may be found from the Support page.

MySQL Data Collector

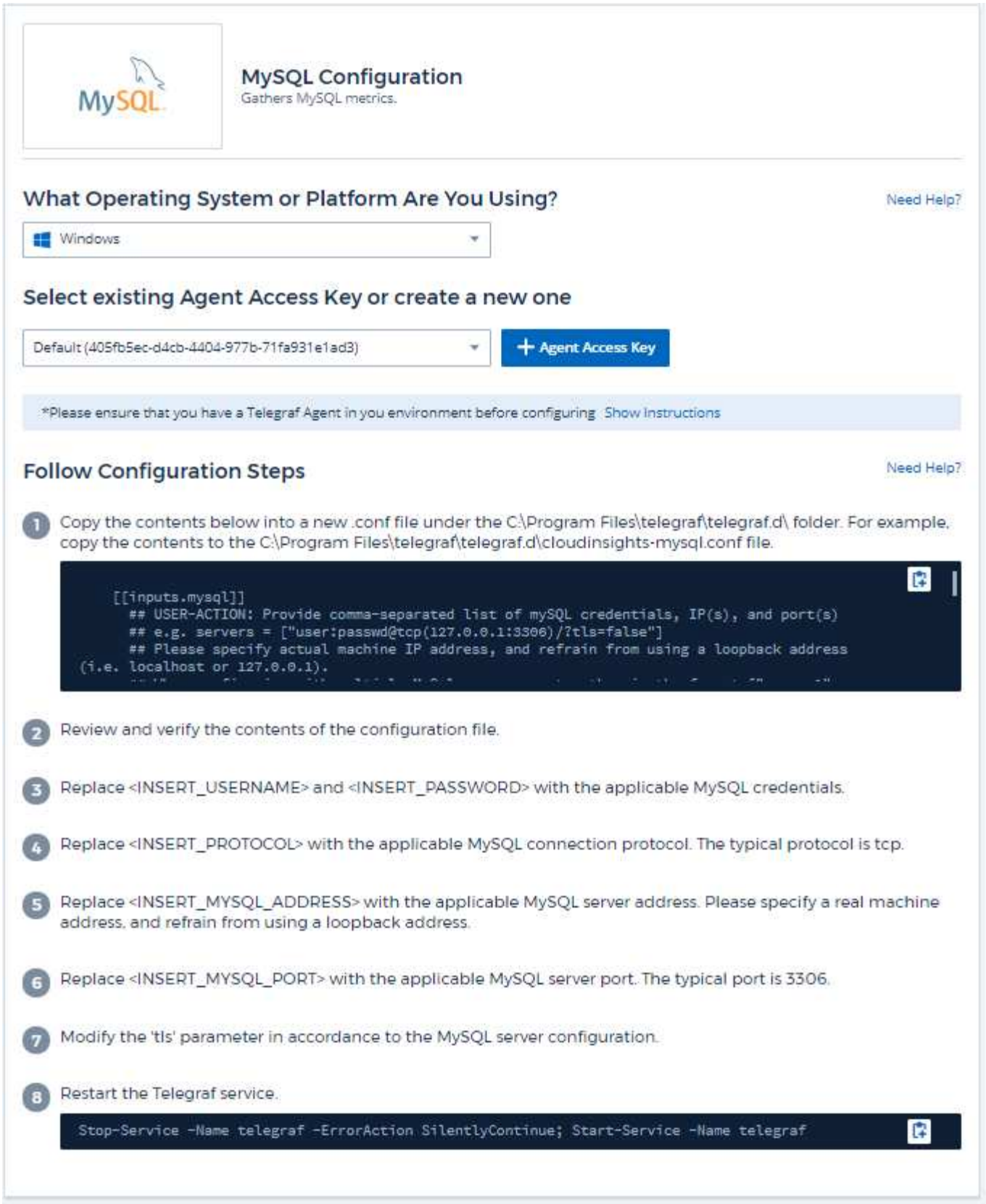
Data Infrastructure Insights uses this data collector to gather metrics from MySQL.

Installation

1. From **Observability > Collectors**, click **+Data Collector**. Choose MySQL.

   Select the Operating System or Platform on which the Telegraf agent is installed.

2. If you haven't already installed an Agent for collection, or you wish to install an Agent for a different Operating System or Platform, click *Show Instructions* to expand the Agent installation instructions.
3. Select the Agent Access Key for use with this data collector. You can add a new Agent Access Key by clicking the **+ Agent Access Key** button. Best practice: Use a different Agent Access Key only when you want to group data collectors, for example, by OS/Platform.
4. Follow the configuration steps to configure the data collector. The instructions vary depending on the type of Operating System or Platform you are using to collect data.



Setup

Information may be found in the MySQL documentation.

Objects and Counters

The following objects and their counters are collected:

Object:	Identifiers:	Attributes:	Datapoints:
MySQL	Namespace MySQL Server	Node IP Node Name	Aborted Clients (per sec) Aborted Connects (per sec) RX Bytes (per sec) TX Bytes (per sec) Commands Admin (per sec) Commands Alter Event Commands Alter Function Commands Alter Instance Commands Alter Procedure Commands Alter Server Commands Alter Table Commands Alter Tablespace Commands Alter User Commands Analyze Commands Assign To Keycache Commands Begin Commands Binlog Commands Call Procedure Commands Change DB Commands Change Master Commands Change Repl Filter Commands Check Commands Checksum Commands Commit Commands Create DB Commands Create Event Commands Create Function Commands Create Index Commands Create Procedure Commands Create Server Commands Create Table Commands Create Trigger Commands Create UDF Commands Create User Commands Create View Commands Dealloc SQL Connection Errors Accept Created Tmp Disk Tables Delayed Errors Flush Commands Handler Commit Innodb Buffer Pool Bytes Data Key Blocks Not Flushed

Troubleshooting

Additional information may be found from the Support page.

Netstat Data Collector

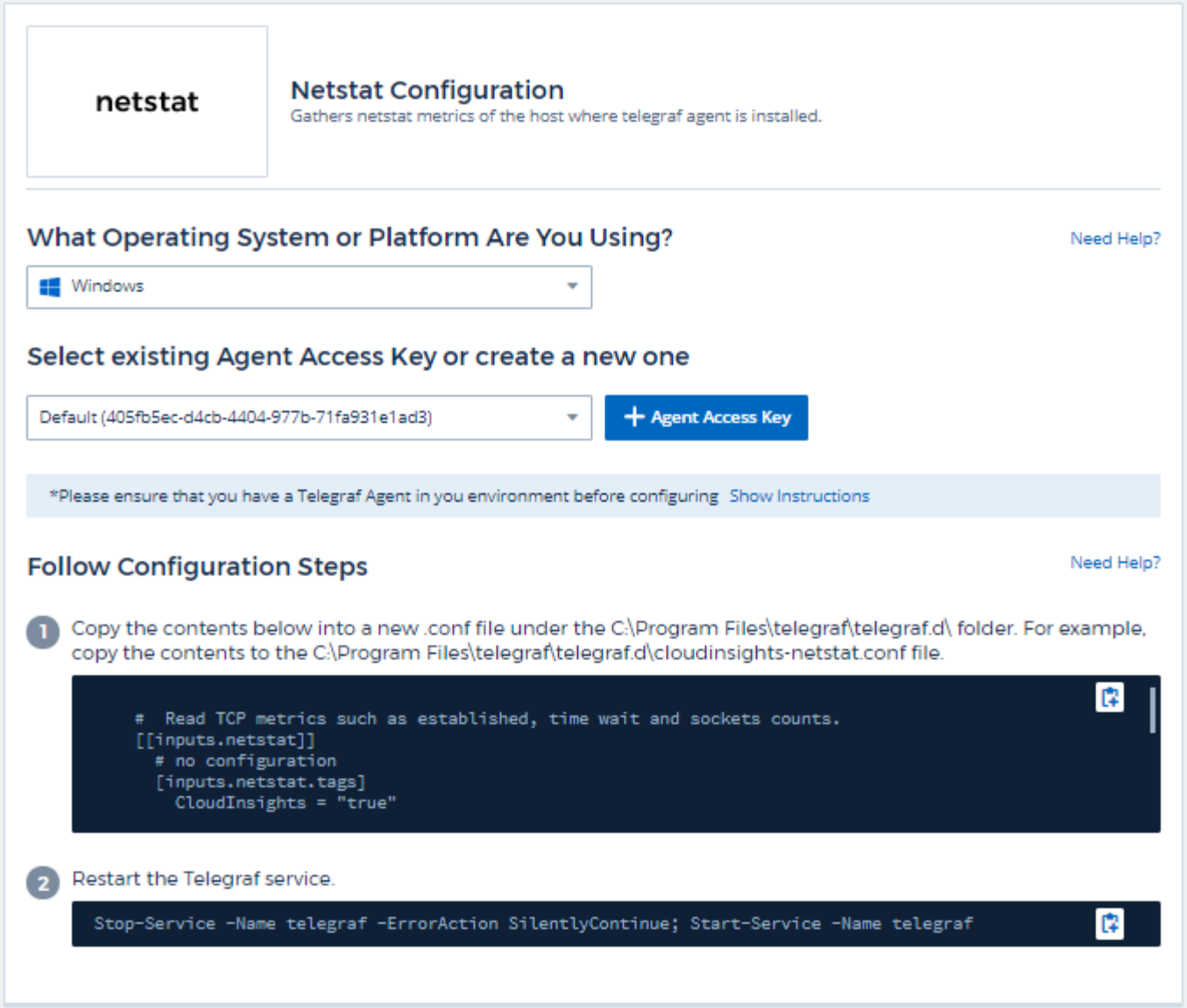
Data Infrastructure Insights uses this data collector to gather Netstat metrics.

Installation

1. From **Observability > Collectors**, click **+Data Collector**. Choose Netstat.

   Select the Operating System or Platform on which the Telegraf agent is installed.

2. If you haven't already installed an Agent for collection, or you wish to install an Agent for a different Operating System or Platform, click *Show Instructions* to expand the Agent installation instructions.
3. Select the Agent Access Key for use with this data collector. You can add a new Agent Access Key by clicking the **+ Agent Access Key** button. Best practice: Use a different Agent Access Key only when you want to group data collectors, for example, by OS/Platform.
4. Follow the configuration steps to configure the data collector. The instructions vary depending on the type of Operating System or Platform you are using to collect data.



Setup

Objects and Counters

The following objects and their counters are collected:

Object:	Identifiers:	Attributes:	Datapoints:
Netstat	Node UUID	Node IP Node Name	

Troubleshooting

Additional information may be found from the Support page.

Nginx Data Collector

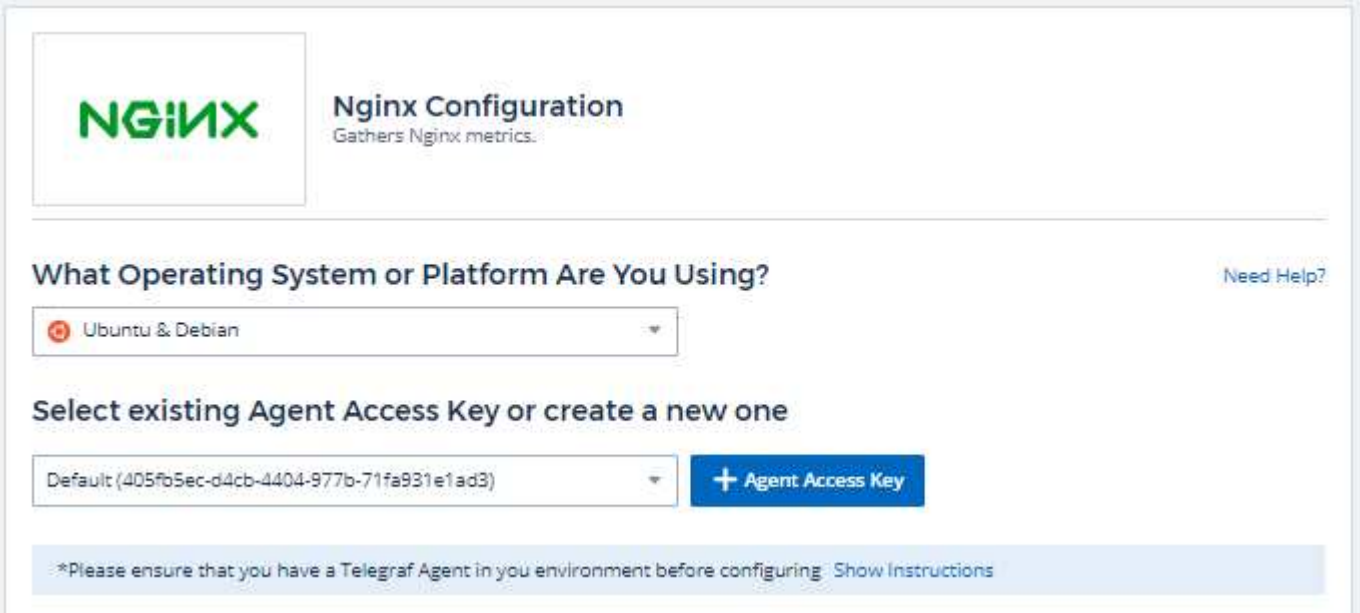
Data Infrastructure Insights uses this data collector to gather metrics from Nginx.

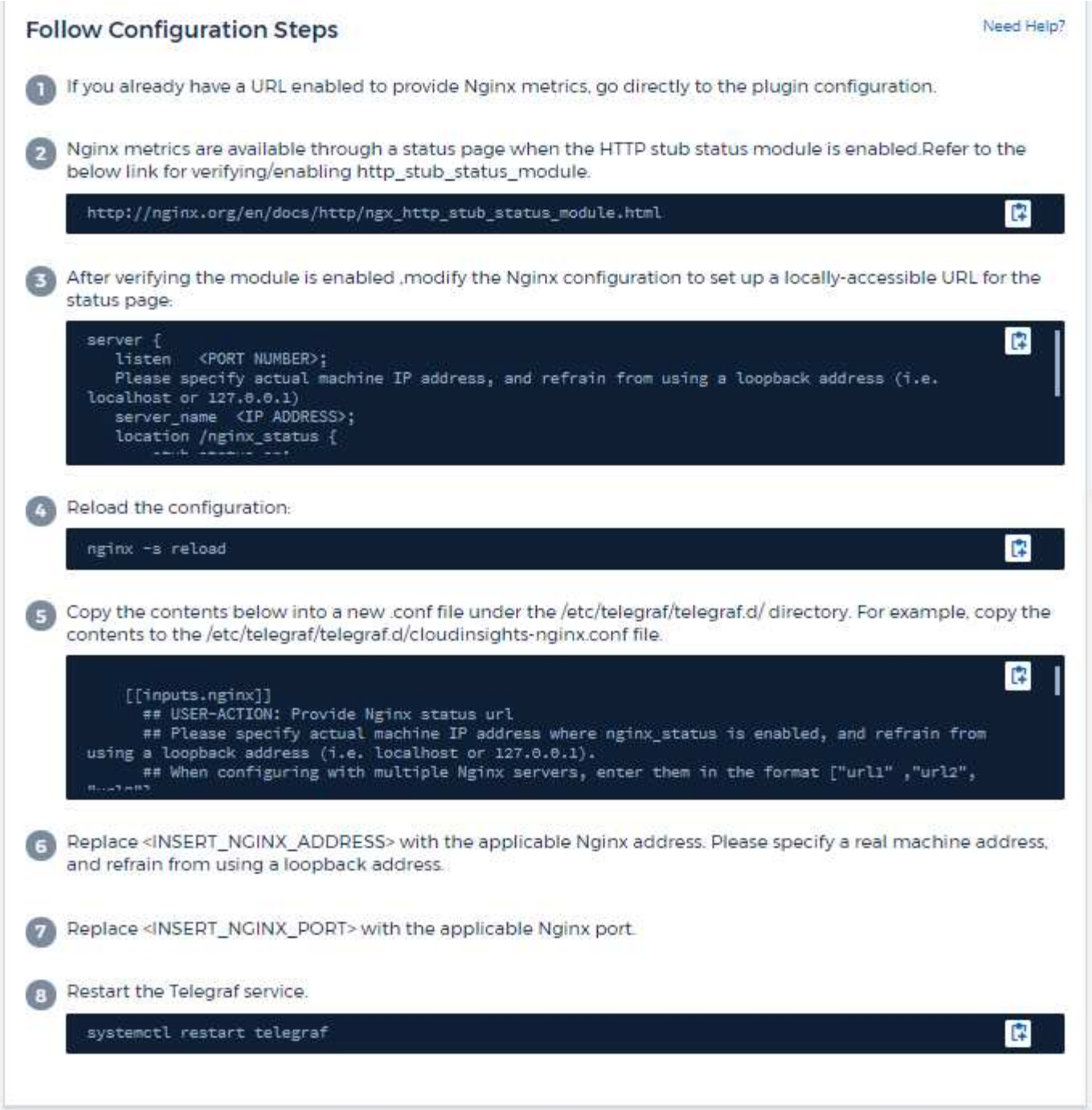
Installation

1. From **Observability > Collectors**, click **+Data Collector**. Choose Nginx.

   Select the Operating System or Platform on which the Telegraf agent is installed.

2. If you haven't already installed an Agent for collection, or you wish to install an Agent for a different Operating System or Platform, click *Show Instructions* to expand the Agent installation instructions.
3. Select the Agent Access Key for use with this data collector. You can add a new Agent Access Key by clicking the **+ Agent Access Key** button. Best practice: Use a different Agent Access Key only when you want to group data collectors, for example, by OS/Platform.
4. Follow the configuration steps to configure the data collector. The instructions vary depending on the type of Operating System or Platform you are using to collect data.





Setup

Nginx metric collection requires that Nginx http_stub_status_module be enabled.

Additional information may be found in the Nginx documentation.

Objects and Counters

The following objects and their counters are collected:

Object:	Identifiers:	Attributes:	Datapoints:
Nginx	Namespace Server	Node IP Node Name Port	Accepts Active Handled Reading Requests Waiting Writing

Troubleshooting

Additional information may be found from the Support page.

PostgreSQL Data Collector

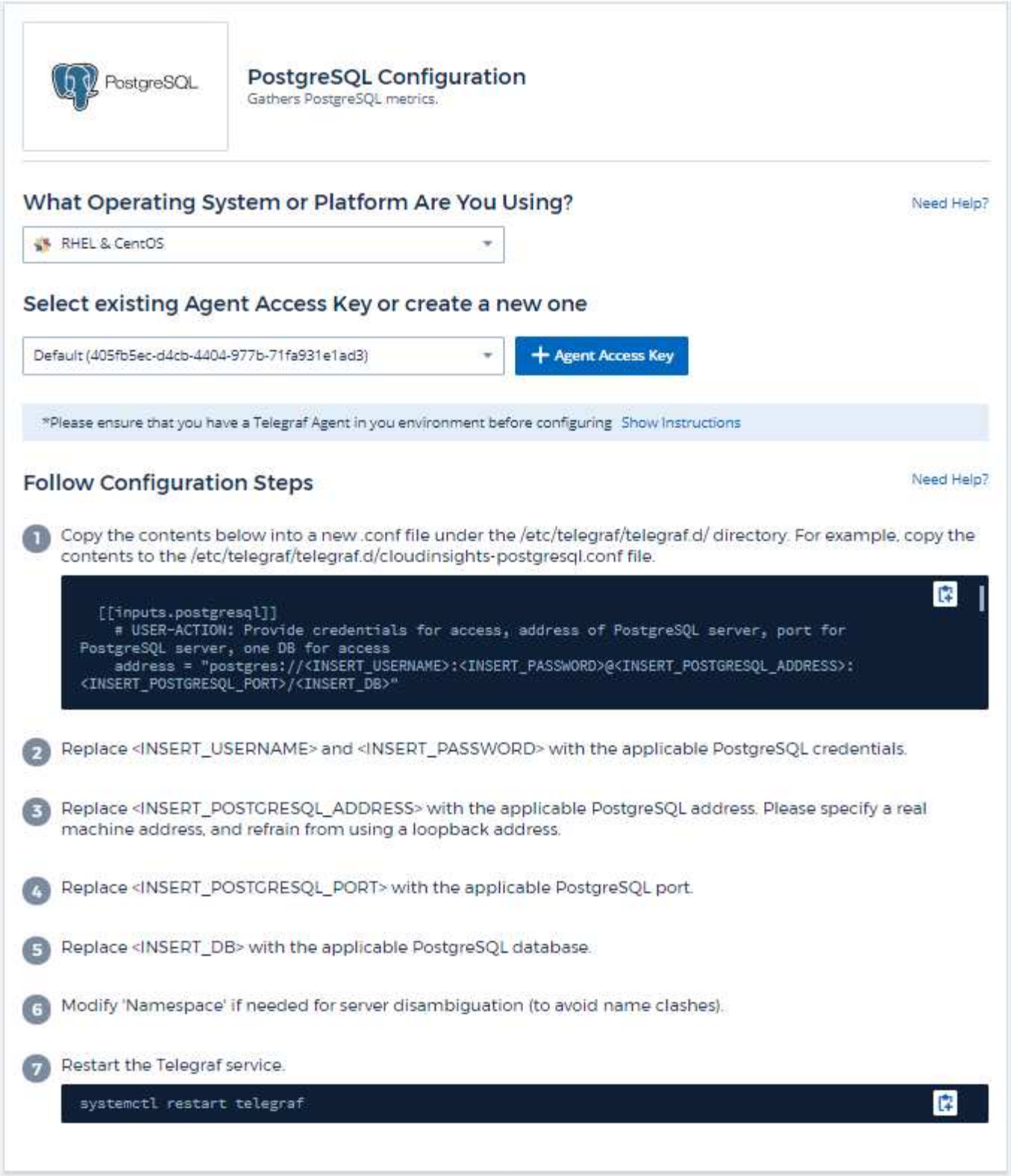
Data Infrastructure Insights uses this data collector to gather metrics from PostgreSQL.

Installation

1. From **Observability > Collectors**, click **+Data Collector**. Choose PostgreSQL.

   Select the Operating System or Platform on which the Telegraf agent is installed.

2. If you haven't already installed an Agent for collection, or you wish to install an Agent for a different Operating System or Platform, click *Show Instructions* to expand the Agent installation instructions.
3. Select the Agent Access Key for use with this data collector. You can add a new Agent Access Key by clicking the **+ Agent Access Key** button. Best practice: Use a different Agent Access Key only when you want to group data collectors, for example, by OS/Platform.
4. Follow the configuration steps to configure the data collector. The instructions vary depending on the type of Operating System or Platform you are using to collect data.



Setup

Information may be found in the PostgreSQL documentation.

Objects and Counters

The following objects and their counters are collected:

Object:	Identifiers:	Attributes:	Datapoints:
PostgreSQL Server	Namespace Database Server	Node Name Node IP	Buffers Allocated Buffers Backend Buffers Backend File Sync Buffers Checkpoint Buffers Clean Checkpoints Sync Time Checkpoints Write Time Checkpoints Requests Checkpoints Timed Max Written Clean
PostgreSQL Database	Namespace Database Server	Database OID Node Name Node IP	Blocks Read Time Blocks Write Time Blocks Hits Blocks Reads Conflicts Deadlocks Client Number Temp Files Bytes Temp Files Number Rows Deleted Rows Fetched Rows Inserted Rows Returned Rows Updated Transactions Committed Transactions Rollbacked

Troubleshooting

Additional information may be found from the Support page.

Puppet Agent Data Collector

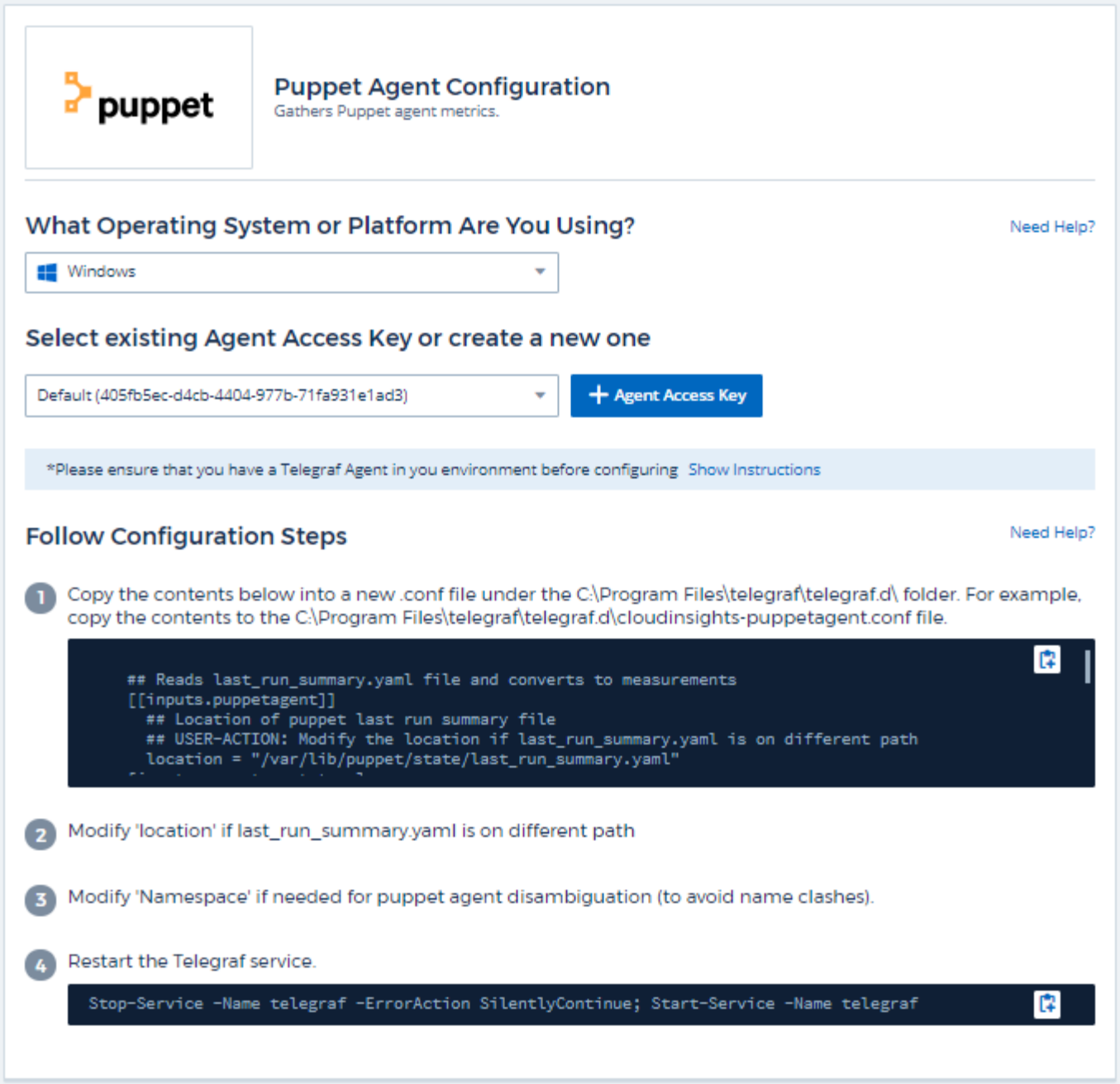
Data Infrastructure Insights uses this data collector to gather metrics from Puppet Agent.

Installation

1. From **Observability > Collectors**, click **+Data Collector**. Choose Puppet.

   Select the Operating System or Platform on which the Telegraf agent is installed.

2. If you haven't already installed an Agent for collection, or you wish to install an Agent for a different Operating System or Platform, click *Show Instructions* to expand the Agent installation instructions.
3. Select the Agent Access Key for use with this data collector. You can add a new Agent Access Key by clicking the **+ Agent Access Key** button. Best practice: Use a different Agent Access Key only when you want to group data collectors, for example, by OS/Platform.
4. Follow the configuration steps to configure the data collector. The instructions vary depending on the type of Operating System or Platform you are using to collect data.



Setup

Information may be found in the Puppet documentation

Objects and Counters

The following objects and their counters are collected:

Object:	Identifiers:	Attributes:	Datapoints:

Puppet Agent	Namespace Node UUID	Node Name Location Node IP Version Configstring Version Puppet	Changes Total Events Failure Events Success Events Total Resources Changed Resources Failed Resources Failed To Restart Resources Outofsync Resources Restarted Resources Scheduled Resources Skipped Resources Total Time Anchor Time Configretrieval Time Cron Time Exec Time File Time Filebucket Time Lastrun Time Package Time Schedule Time Service Time Sshauthorizedkey Time Total Time User

Troubleshooting

Additional information may be found from the Support page.

Redis Data Collector

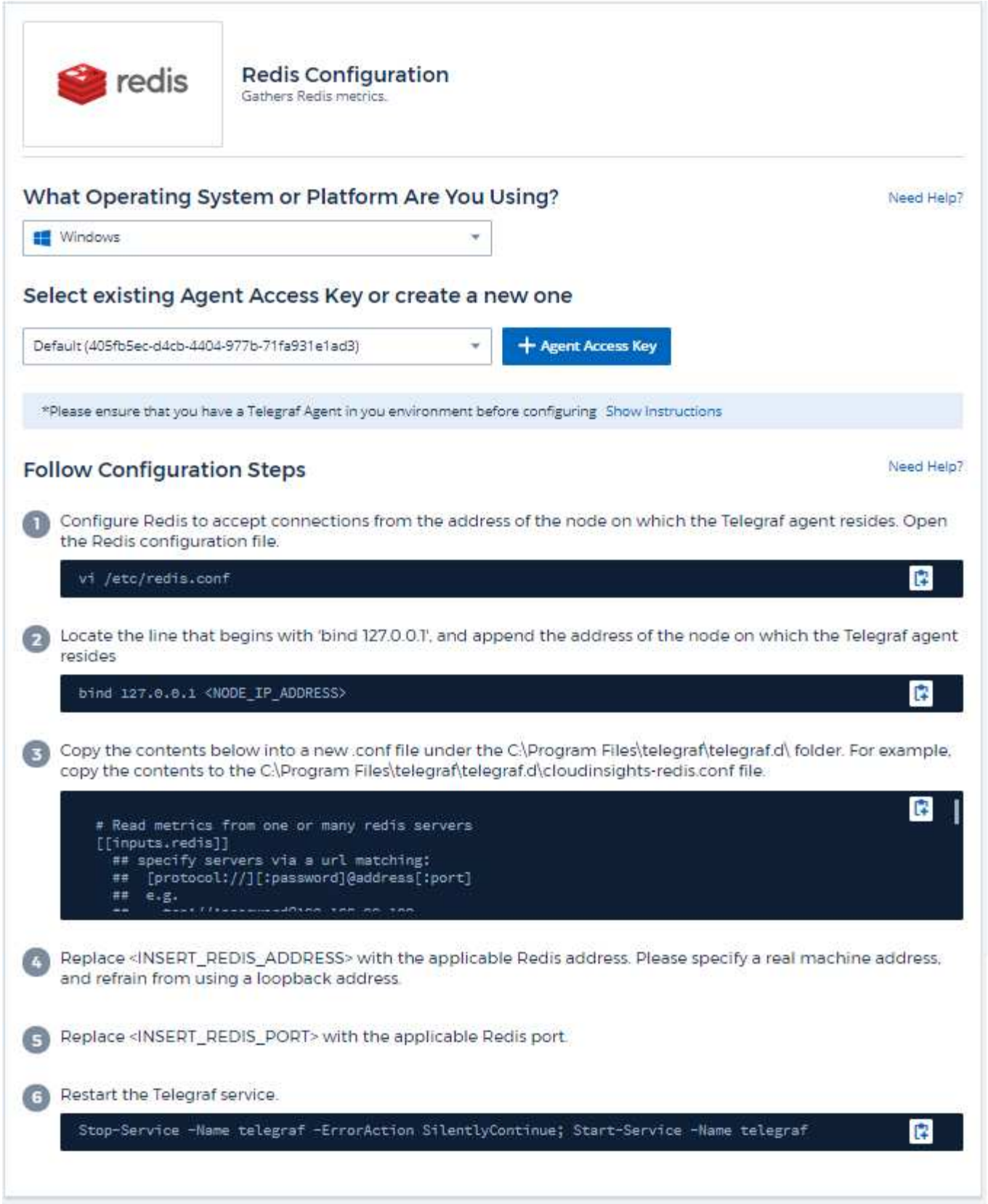
Data Infrastructure Insights uses this data collector to gather metrics from Redis. Redis is an open source, in-memory data structure store used as a database, cache, and message broker, supporting the following data structures: strings, hashes, lists, sets, and more.

Installation

1. From **Observability > Collectors**, click **+Data Collector**. Choose Redis.

   Select the Operating System or Platform on which the Telegraf agent is installed.

2. If you haven't already installed an Agent for collection, or you wish to install an Agent for a different Operating System or Platform, click *Show Instructions* to expand the Agent installation instructions.
3. Select the Agent Access Key for use with this data collector. You can add a new Agent Access Key by clicking the **+ Agent Access Key** button. Best practice: Use a different Agent Access Key only when you want to group data collectors, for example, by OS/Platform.
4. Follow the configuration steps to configure the data collector. The instructions vary depending on the type of Operating System or Platform you are using to collect data.



Setup

Information may be found in the Redis documentation.

Objects and Counters

The following objects and their counters are collected:

Object:	Identifiers:	Attributes:	Datapoints:
Redis	Namespace Server		

Troubleshooting

Additional information may be found from the Support page.

Object Icon Reference

Object icons used in Data Infrastructure Insights.

Infrastructure Icons:

Kubernetes Icons:

Kubernetes Network Performance Monitoring and Map Icons:

Copyright information

Trademark information