



CROWDSTRIKE

CrowdStrike
Falcon Data Replicator (FDR): SQS
Add-on for Splunk
Installation and Configuration Guide v2.0+

Introduction	3
Requirements	4
Before Getting Started	5
Getting Started	6
FDR Communication Flow	6
FDR Data Volume Considerations	6
The FDR Event Classifications	7
FDR Folder Structure	7
FDR Event Classifications	7
High Level Data Flow	8
Validating that FDR is Enabled	9
Generating/Collecting FDR Credentials	10
Generating New FDR Credentials	10
Collecting FDR Credentials	10
Proxy Considerations	11
Splunk Architecture	11
Configuring the TA	13
TA Layout	13
Inputs Section	13
Configuration Section	14
Search Section	14
Configuring the TA to collect data	15
Configure Proxy Settings (optional)	15
Configure an Account	16
Creating an Input	17
Configure an Input	19
Configuring CrowdStrike FDR Data Inputs	20
Configuring CrowdStrike FDRv2 Based Inputs	21
Search Macros	23
Recommendations	24
Custom Indexes	24
AID Master Data	24
Troubleshooting	25
Configuring the TA to collect log data	25

Change Logging Level	25
Obtaining and Contacting Support	26
Additional Resources	27

Introduction

This guide covers the deployment, configuration and usage of the CrowdStrike Falcon Data Replicator: SQS Technical Add-on (TA) for Splunk.

The CrowdStrike Falcon Data Replicator Technical Add-on for Splunk allows CrowdStrike customers to retrieve FDR data from the CrowdStrike hosted S3 buckets via the CrowdStrike provide SQS Queue.

To get more information about this CrowdStrike Falcon Data Replicator (FDR), please refer to the FDR documentation which can be found in the CrowdStrike Falcon UI:

[CrowdStrike Falcon Data Replicator Guide](#)

For information about the event types contained in FDR, please refer to the Events Data Dictionary documentation which can be found in the CrowdStrike Falcon UI:

[CrowdStrike Events Data Dictionary](#)

Multitenancy - This TA is able to have multiple independent inputs enabled at the same time, each collecting data from different Falcon Instances and storing it in independent indexes.

Requirements

The following are the requirements to leverage this technical add-on:

1. An active subscription to the CrowdStrike Falcon Data Replicator
2. A Splunk Heavy forwarder or Input Data Manager (IDM)
3. A Splunk account with proper access to deploy and configure technical add-ons
4. An active FDR credential and SQS URL or proper access to the CrowdStrike Falcon instance to create one
5. The CrowdStrike Cloud environment that the Falcon instance resides in

NOTE: This technical add-on is not designed to work with any S3 or SQS configuration except the one provided and maintained directly by CrowdStrike

Before Getting Started

Before deploying this technical add on please review the following:

1. CrowdStrike's Falcon Data Replicator is a data 'dump', as opposed to an API, to an AWS hosted S3 bucket that is associated with an SQS queue that can be monitored to notify customers when a new data package is available.
2. As it is not an API there is no capability to filter the requested FDR data packages. Data packages must be downloaded and decompressed in order to be examined and processed. This means that any filtering of the data packages will have to be done after downloading and decompression have taken place and would require that the evaluation be done on an event-by-event basis.
3. In order to prevent data from being ingested multiple times the TA input will put a hold of the file for approximately 5 minutes. During this hold period, no other clients will be allowed to attempt to download the data. In the event that the client is unable to properly process the data and fails to delete the message from the queue, the message will be accessible once the time has expired.
4. The SQS queue does not provide data in chronological order in regards to the timestamps of the events contained within the package. In addition, the data packages are constructed by data being received at that moment in time so it is very possible for packages to have events with very different timestamps.

FDR SQS MESSAGES ARE DELETED AFTER THEY'RE PROCESSED. THEREFORE, IT'S RECOMMENDED THAT THE CLIENT OR CLIENTS COLLECTING THE DATA PLACE IT IN A CENTRAL LOCATION SUCH AS A SINGLE SPLUNK ENVIRONMENT.

Getting Started

FDR Communication Flow

The CrowdStrike FDR TA for Splunk leverages the SQS message queue provided by CrowdStrike to identify that data is available to be retrieved in the CrowdStrike provided S3 bucket. The TA communication process is as follows:

1. The TA will query the CrowdStrike SQS queue for a maximum of 10 messages (this is the maximum allowed per AWS published documentation) matching the input type to the type of data in the message
2. The TA will use the information contained in the message to retrieve the compressed data file from the CrowdStrike S3 bucket
3. The TA will then decompress the file and prepare it to send to Splunk
4. The TA will post the file data to Splunk to be indexed

FDR Data Volume Considerations

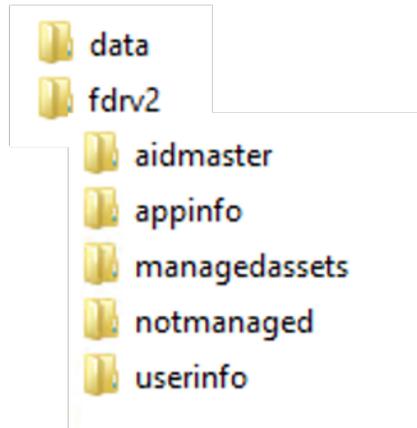
CrowdStrike FDR produces an extremely large amount of data that can be problematic to ingest into Splunk for some customers. Some specific issues that should be taken into account are:

1. The amount of data that will be ingested
2. The available resources to collect and process the data

The FDR Event Classifications

The CrowdStrike FDR TA can collect the both primary and secondary events from FDR and follows the FDR folder structure for data types.

FDR Folder Structure

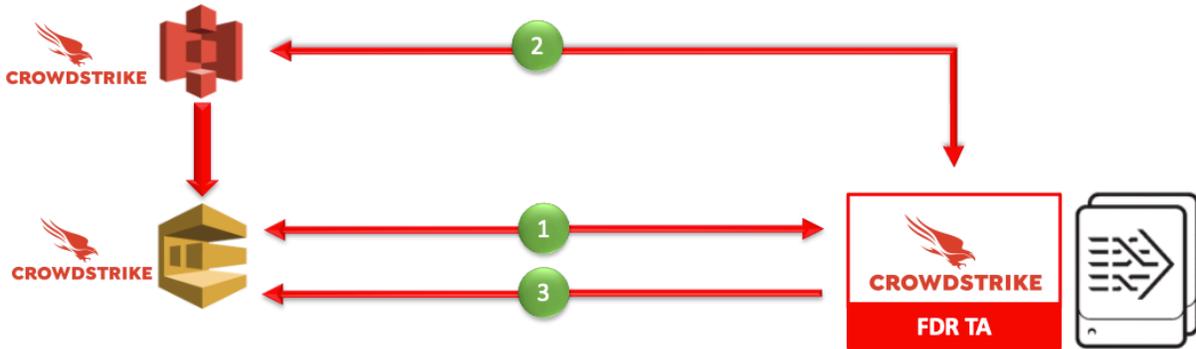


FDR Event Classifications

- **Primary events:** Describe specific data and individual actions taking place on CrowdStrike protected hosts. The following folders contain primary events:
 - **data:** contains the raw sensor telemetry from Falcon sensor and processed event data
 - **The following events require the Falcon Insight module**
 - **fdrv2 - aidmaster:** contains basic host information collected by the Falcon sensor
 - **fdrv2 - managedassets:** contains basic network configuration information collected by the Falcon sensor
 - **fdrv2 - notmanagedassets:** contains basics network configuration information collected by the Falcon sensor about devices in the network not running a Falcon sensor
- **Secondary events:** Events containing higher-level information Falcon Sensors have collected about the environment.
 - **The following events requires the Falcon Discover module**
 - **fdrv2 - appinfo:** contains application information collected by hosts running Falcon sensors
 - **fdrv2 - userinfo:** contains user information collected by hosts running the Falcon sensor

High Level Data Flow

The CrowdStrike FDR: SQS TA leverages a CrowdStrike provided AWS SQS queue to determine when a new FDR data package is available for download from the CrowdStrike provided AWS S3 bucket:



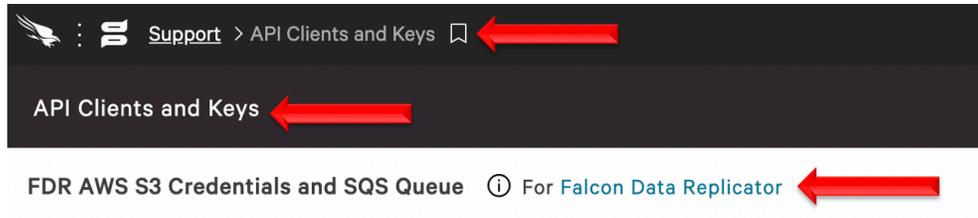
1. The TA accesses the SQS queue
 - a. collects 10 SQS messages with the data package IDs*
 - b. 'hides' the package IDs for the VisibilityTimeout setting*
2. The TA collects the FDR data from the S3 bucket
3. The TA removes the message from the SQS queue

*Notes:

- 10 SQS messages is the maximum number allowed by AWS at the time the TA was created
- The VisibilityTimeout function is used to prevent another client from attempting to collect the same message that's currently being processed. In the event that the message can't be processed by the original client within the VisibilityTimeout timeframe it will be released and returned to being visible in the queue. The current setting for the FDR: SQS TA is 300 seconds.

Validating that FDR is Enabled

The CrowdStrike FDR:SQS TA requires that FDR be enabled on the CrowdStrike instance.



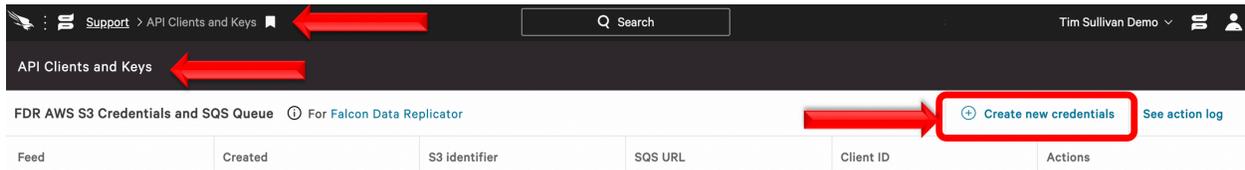
1. Access the CrowdStrike Falcon user interface (UI) with an account that is able to view/create the API clients and keys page
2. Navigate to 'Support'>'API Client and Keys' page
3. Validate that 'FDR AWS S3 Credentials and SQS Queue' is present

CrowdStrike FDR can only be enabled by CrowdStrike Support
If FDR is not enabled, please submit a support ticket through the support portal:
<https://supportportal.crowdstrike.com/>

Generating/Collecting FDR Credentials

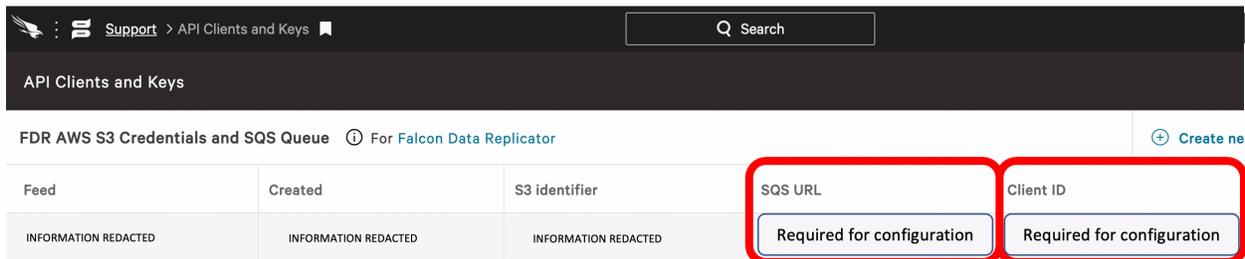
The FDR:SQS TA requires the FDR credentials that are located in the CrowdStrike Falcon UI in order to access the data. These can be existing FDR credentials or can be newly generated credentials.

Generating New FDR Credentials



1. Access the CrowdStrike Falcon user interface (UI) with an account that is able to view/create the API clients and keys page
2. Navigate to 'Support'>'API Client and Keys' page
3. On the same line as 'FDR AWS S3 Credentials and SQS Queue' select 'Create new credentials'

Collecting FDR Credentials



1. Access the CrowdStrike Falcon user interface (UI) with an account that is able to view/create the API clients and keys page
2. Navigate to 'Support'>'API Client and Keys' page
3. Collect the SQS URL and Client ID
4. The Secret is only available via the UI when the credential is created. If the Secret value is no longer available and a new credential cannot be created, an existing one will need to be deleted and a new one recreated. *

***NOTE:** The FDR:SQS TA uses the SQS URL to determine that there's data available for collection and then delete the message from the queue. Therefore, existing FDR credentials that are already in use **SHOULD NOT** be used unless it's replacing an existing client.

Proxy Considerations

The CrowdStrike FDR:SQS Add-On communicates with the AWS infrastructure and any proxy systems in the environment should be configured to allow this communication.

Splunk Architecture

Splunk Search Head(s) and Splunk Cloud: The TA should be installed to provide field mapping and search macro support. These are often required to support CrowdStrike Apps. The TA should be deployed without any accounts or inputs configured and any search macros should be properly configured for use.

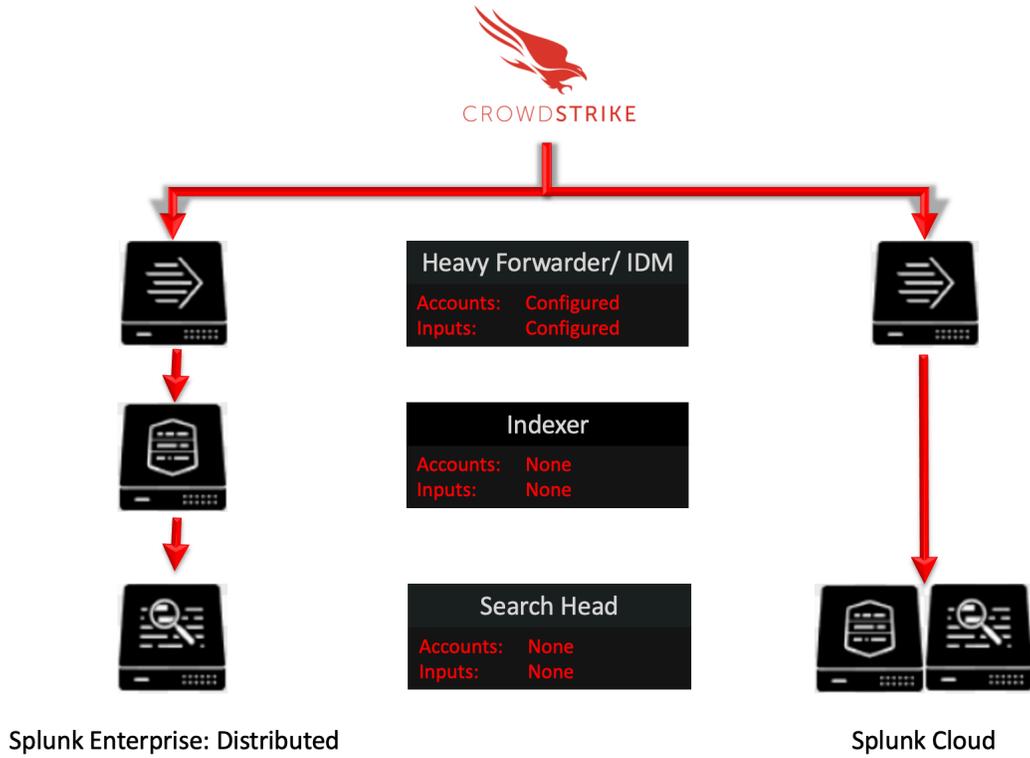
Splunk Indexer(s): The TA can be installed to provide field mapping and search macro support. The TA should be deployed without any accounts or inputs configured and any search macros should be properly configured for use. If a custom index is going to be used, then it should be created here.

Splunk Heavy Forwarder(s) & Information Data Managers (IDMs): The TA is required to be installed here as this is where the data will be collected. The appropriate accounts and inputs should be properly configured for data collection. Ensure that if a custom index is being used, which is highly recommended, that the index has been created on the indexer tier. If the Heavy Forwarder is storing events (not required but is an optional Splunk configuration) prior to forwarding them to the Indexer and a custom index is being used, ensure that the index has been created on both the Heavy Forwarder as well as the Indexer(s).

Note:

Due to python requirements the TA can only be configured for data collection on Heavy Forwarders and IDMs.

The following diagram shows the flow of data from the CrowdStrike FDR and the FDR:SQS TA configuration within a distributed Splunk Enterprise and Splunk Cloud environment:



Configuring the TA

TA Layout

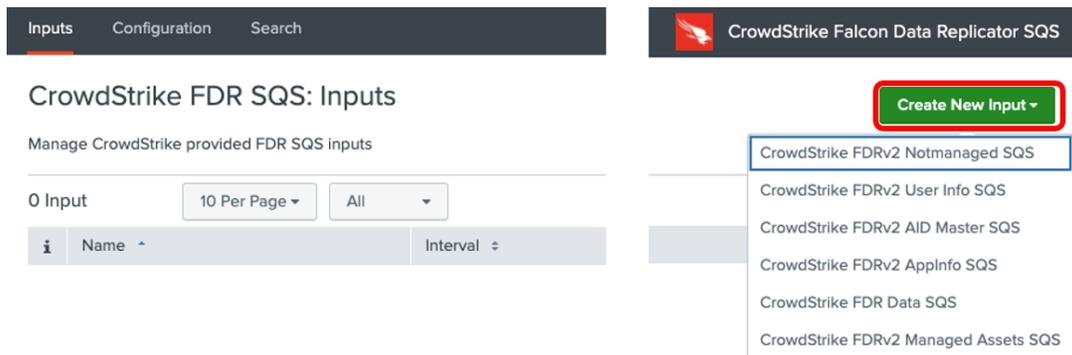
The TA contains 3 sections.



- The Inputs section
- The Configuration section
- The Search section

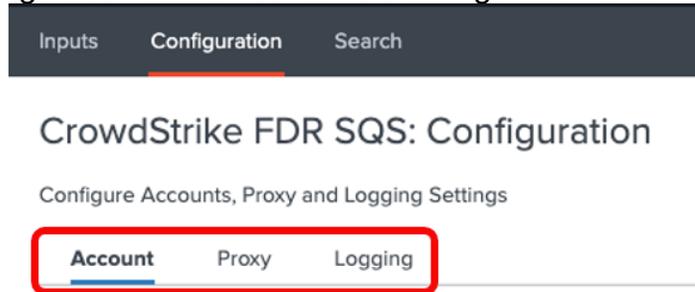
Inputs Section

The Inputs section is where inputs are configured, modified and listed. Prior to configuring any inputs an account needs to be created under the Configuration section (see below). In the far-right corner of the Inputs section contains a pull-down menu to create a new input configuration.



Configuration Section

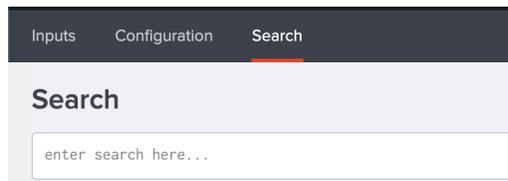
The Configuration section contains 3 configuration tabs:



- **FDR Account:** This is where the FDR credentials are entered.
- **Proxy:** This is where proxy server configurations are entered.
- **Logging:** This is where the logging level is configured.

Search Section

The Search section opens a standard Splunk search page but within the context of the TA.

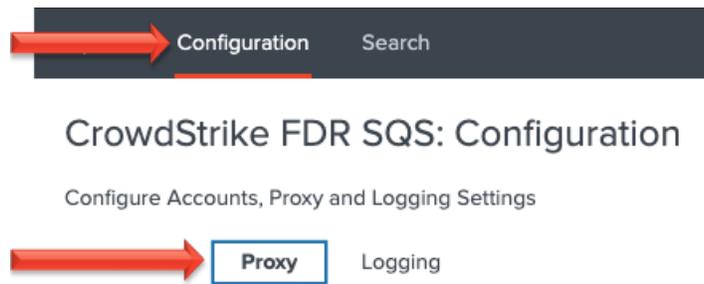


Configuring the TA to collect data

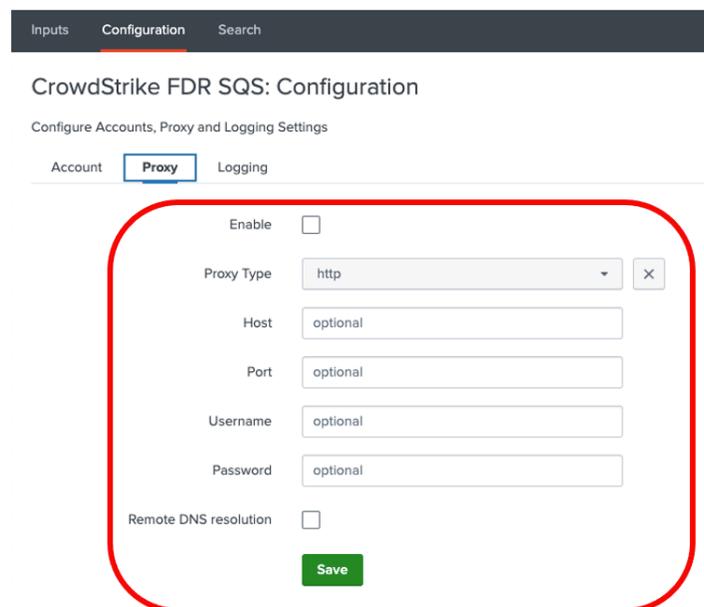
NOTE This action should only be performed on:
Splunk Heavy Forwarders and Splunk IDMs

Configure Proxy Settings (optional)

1. Proxy settings are configured under the Configuration section, Proxy tab. Proxies can cause authentication issue if not configured correctly, ensure that the proxy does not interfere with communication between the TA and the AWS SQS URL and S3 bucket.



2. Configure the following fields as appropriate:

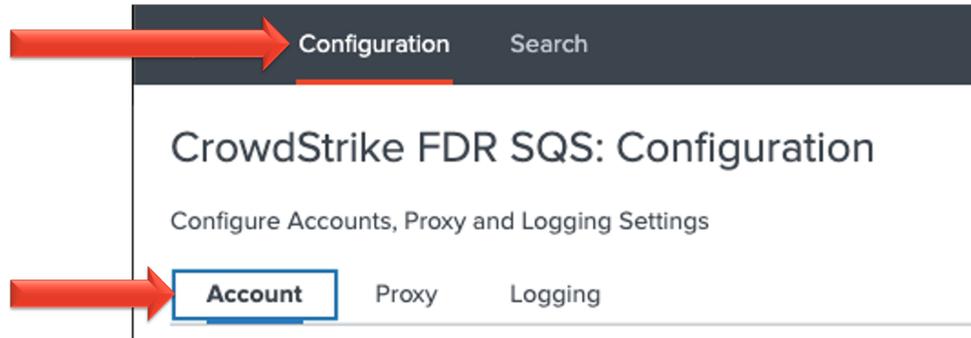


- **Enable:** This checkbox is used to enable/disable the proxy settings
- **Proxy Type:** This dropdown is used to select the proxy type
- **Host:** The hostname/IP address for the proxy server
- **Port:** The communication port for the proxy server
- **Username:** The authentication username for the proxy (optional)

- **Password:** The authentication password for the proxy (optional)
- **Save:** This button is used to save the configuration

Configure an Account

1. An account is configured using an FDR credential from the CrowdStrike Falcon UI.
2. An account is created under the Configuration section, FDR Account tab:



3. On the right side of the screen click the "Add" button:



4. Configure the following fields:

Add Account ✕

Account name
Enter a unique name for this account.

ClientID
Enter the FDR ClientID here
Enter the FDR ClientID provided in the Falcon UI.

Secret
Enter the FDR Secret provided in the Falcon UI.

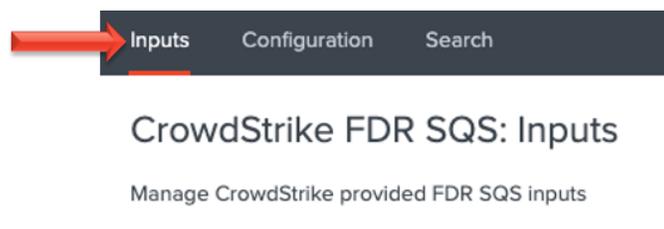
Cancel Add

- **Account Name:** A name unique for the Splunk instance
- **ClientID:** The ClientID of the FDR credential created in the CrowdStrike Falcon UI.
- **Secret:** The Secret of the FDR credential created in CrowdStrike Falcon UI.

5. Click the 'Add' button in the bottom right corner to save the account.

Creating an Input

1. An input will require a valid FDR account be created already.
2. An input is created under the Inputs section:

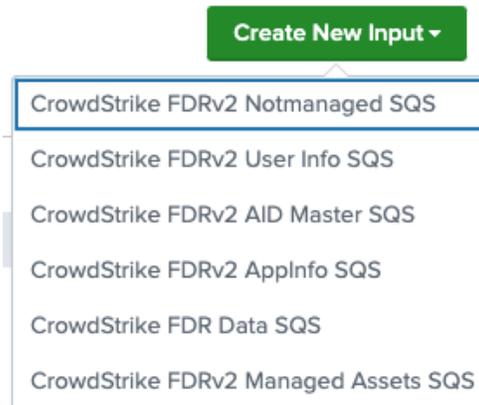


3. In the top right corner select the 'Create New Input' dropdown to display the available input types.

 CrowdStrike Falcon Data Replicator SQS

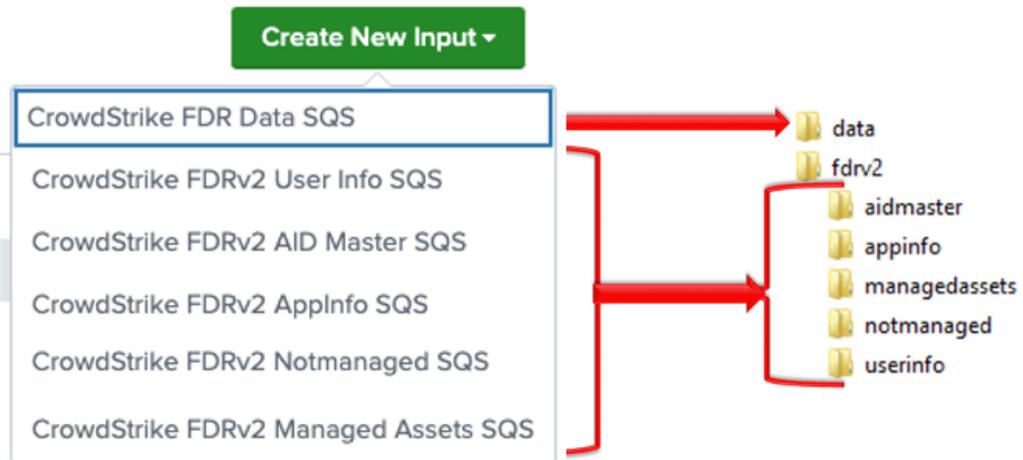


4. Select the input type to configure.



Configure an Input

The CrowdStrike FDR TA provides the ability to configure multiple input types. These input types align with the current folder structure in the FDR S3 bucket.



- **CrowdStrike FDR Data SQS:** Collects information in the 'Data' folder
- **CrowdStrike FDRv2 Notmanaged SQS:** Collects information in the 'FDRv2', 'Notmanaged' folder
- **CrowdStrike FDRv2 AIDMaster SQS:** Collects information in the 'FDRv2', 'AIDMaster' folder
- **CrowdStrike FDRv2 UserInfo SQS:** Collects information in the 'FDRv2', 'UserInfo' folder
- **CrowdStrike FDRv2 Managed Assets SQS:** Collects information in the 'FDRv2', 'Managed' folder
- **CrowdStrike FDRv2 AppInfo SQS:** Collects information in the 'FDRv2', 'UserInfo' folder

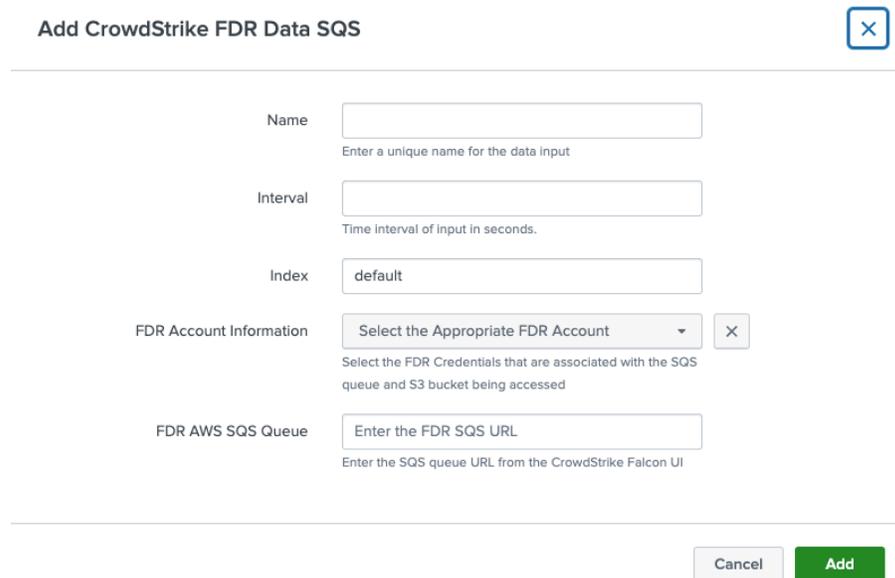
Configuring CrowdStrike FDR Data Inputs

The FDR Data Input contains the sensor telemetry and event data. Since the CrowdStrike FDR TA leverages the AWS SQS Queue for message tracking, it is possible to create multiple inputs for a single FDR S3 bucket.

1. Under 'Create New Input', select the 'CrowdStrike FDR Data' input type



2. Configure the appropriate fields:

A screenshot of a configuration form titled 'Add CrowdStrike FDR Data SQS'. The form contains several fields: 'Name' (text input), 'Interval' (text input), 'Index' (text input with 'default' selected), 'FDR Account Information' (dropdown menu with 'Select the Appropriate FDR Account' and a close button), and 'FDR AWS SQS Queue' (text input). Below the form are 'Cancel' and 'Add' buttons. A close button (X) is in the top right corner of the form area.

- **Name:** (required) A name unique to the Splunk Environment
 - **Interval:** (required) How often the specific input will run, expressed in seconds
 - **Index:** (required) The Splunk Index that the data will be stored in
 - **FDR Account Information:** (required) The appropriate FDR credential set configured in the 'FDR Account' tab under 'Configuration'
 - **FDR AWS SQS Queue:** (required) The SQS Queue URL listed in the CrowdStrike Falcon UI for the particular FDR S3 bucket
3. Click the 'Add' button in the bottom right corner to save and active the input.

Configuring CrowdStrike FDRv2 Based Inputs

All FDRv2 based Inputs share the same configuration settings. These currently include:



When configuring the 'interval' setting for FDRv2 inputs it is recommended to keep in mind the interval at which this data is posted. Configuring the interval accordingly can prevent the TA from making unnecessary queries and utilizing system resources that are not needed. Please consult with CrowdStrike documentation or CrowdStrike support for more information.

1. Under 'Create New Input', select the 'CrowdStrike FDR Data' input type
2. Configure the appropriate fields:

The screenshot shows a configuration form for a new input. It contains the following fields and labels:

- Name:** A text input field with the placeholder text "Enter a unique name for the data input".
- Interval:** A text input field with the placeholder text "Time interval of input in seconds".
- Index:** A text input field with the value "default".
- FDR Account Information:** A dropdown menu with the text "Select the Appropriate FDR Account" and a close button (X). Below it is the text "Select the FDR Credentials that are associated with the SQS queue and S3 bucket being accessed".
- FDR AWS SQS Queue:** A text input field with the placeholder text "Enter the FDR SQS URL" and "Enter the SQS Queue URL from the CrowdStrike Falcon UI".

At the bottom right of the form, there are two buttons: "Cancel" and "Add".

- **Name:** (required) A name unique to the Splunk Environment
 - **Interval:** (required) How often the specific input will run, expressed in seconds
 - **Index:** (required) The Splunk Index that the data will be stored in
 - **FDR Account Information:** (required) The appropriate FDR credential set configured in the 'FDR Account' tab under 'Configuration'
 - **FDR AWS SQS Queue:** (required) The SQS Queue URL listed in the CrowdStrike Falcon UI for the particular FDR S3 bucket
3. Click the 'Add' button in the bottom right corner to save and active the input.

Search Macros

The FDR TA contains 6 configurable search macros:

Name ↕	Definition ↕	Arguments ↕	Owner ↕	App ↕
cs_fdr_data_sqs_get_index	index=*		No owner	TA-crowdstrike-falcon-data-replicator-sqs
cs_fdrv2_aidmaster_sqs_get_index	index=*		No owner	TA-crowdstrike-falcon-data-replicator-sqs
cs_fdrv2_appinfo_sqs_get_index	index=*		No owner	TA-crowdstrike-falcon-data-replicator-sqs
cs_fdrv2_managed_sqs_get_index	index=*		No owner	TA-crowdstrike-falcon-data-replicator-sqs
cs_fdrv2_notmanaged_sqs_get_index	index=*		No owner	TA-crowdstrike-falcon-data-replicator-sqs
cs_fdrv2_userinfo_sqs_get_index	index=*		No owner	TA-crowdstrike-falcon-data-replicator-sqs

There are 6 that are configured to indicate the index(es) for certain input types and are configured by default to point to all indexes. These Search Macros should be updated to point to the correct index(es) prior to being leveraged.

Ensure the following:

1. Search Macros must be enclosed by 'back ticks', not single quotes. This key is located above the 'Tab' key, to the left of the number 1 on most US style keyboards.
2. Ensure that the account leveraging the macro has the correct permissions to use the macro or adjust the permission of the macro accordingly.
3. Ensure that the account leveraging the macro has the correct permissions to access the FDR data.
4. Ensure that the index(es) have been designated correctly.

Recommendations

The following are general recommendations. They may not be optimal in all situations and should be evaluated on an environment-by-environment basis.

Custom Indexes

The use of a dedicated custom index is strongly recommended for the CrowdStrike FDR data. The FDR TA was specifically designed to facilitate the indexing of different data types and event different event types to specific indexes.

Some examples of benefits that leveraging custom indexes can provides:

- Allows multiple teams to reference the data without exposing other data sets that may be more sensitive.
- Allows data collection types to be assigned to different Heavy Forwarders/IDM for access and resource allocation considerations.
- Improves searching response times and reduces resources needed.

AID Master Data

The AID Master data was designed to provide the ability to relate a hostname with the associated AID (Agent ID) while also providing some basic host information. While this information is useful and may satisfy some use cases, it's recommended that customers leverage the CrowdStrike Falcon Device TA to collect a much more comprehensive data set. This can be in place of or in addition to the data collected in the AID Master Data input.

Troubleshooting

CrowdStrike only provides support for:

- TA code-based functionality errors
- CrowdStrike provided FDR architecture SQS/S3 based access errors

Examples of issues that are outside the scope of CrowdStrike support:

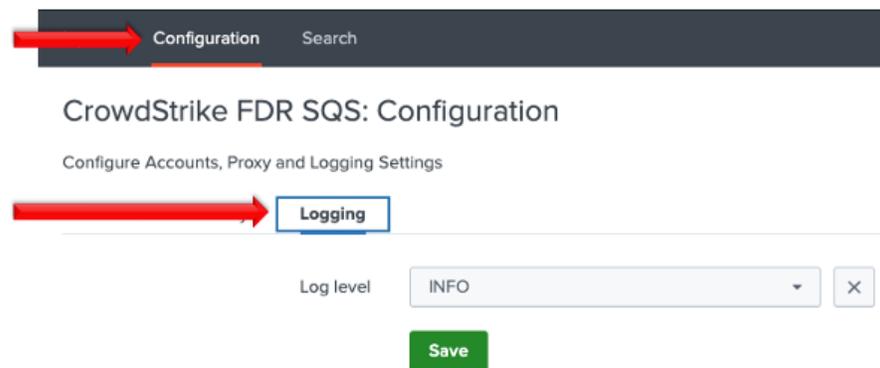
- Proxy based issues
- Firewall based issues
- Network connectivity issues
- Authentication issues (based on misconfigured credentials)
- Splunk CIM field mapping

Configuring the TA to collect log data

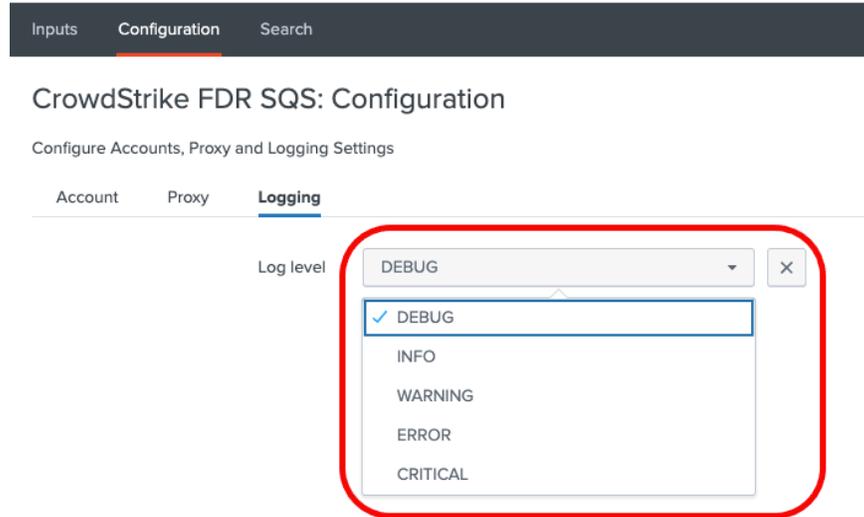
The TA logging level is set to 'info' by default and will only log a minimal amount of information. To properly troubleshoot issues with the TA the logging level should be set to 'debug'.

Change Logging Level

1. Navigate to the Configuration section, Logging tab:



2. Select the logging level from the drop-down menu:



3. Click 'Save' to save the logging level.

Obtaining and Contacting Support

1. Ensure that the FDR credential has been entered correctly
2. Ensure that the FDR data type is available in the FDR configuration
3. Set the TA log level to 'DEBUG'
4. Repeat and record the action(s) that are associated with the issue you are reporting
5. Download the all log files containing 'ta_crowdstrike_falcon_data_replicator_sqs' under the \$Splunk/var/log/splunk/ directory
6. Record the following information about the Splunk system:
 - Splunk environment type
 - Splunk version
 - TA version
7. Identify the types of networks devices that the connection will traverse and ensure that they have been properly configured
8. Navigate to <https://supportportal.crowdstrike.com/>
9. Provide (at a minimum) the information from steps 4-7

Additional Resources

(Access to the CrowdStrike Falcon UI Required)

[Falcon Data Replicator Feature Guide](#)

[Events Data Dictionary](#)