CROWDSTRIKE

Beta

CrowdStrike Falcon Data Replicator (FDR) Add-on for Splunk

Installation and Configuration Guide v1.5

(Splunkbase Posted BETA Release)

Introduction

This guide covers the deployment, configuration and usage of the CrowdStrike Falcon Data Replicator 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 and index it into Splunk.

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

Getting Started

FDR S3 Communication

The CrowdStrike FDR TA for Splunk leverages a different data access methodology than FDR clients have in the past. Historically the way that FDR data was handled was to configure a client to listen to an AWS SQS queue and when a new data package was available the client would get the information from the SQS queue, download the data package and remove the message from the SQS queue. While this method was efficient, it was also limiting in that only one client could be configured per SQS queue.

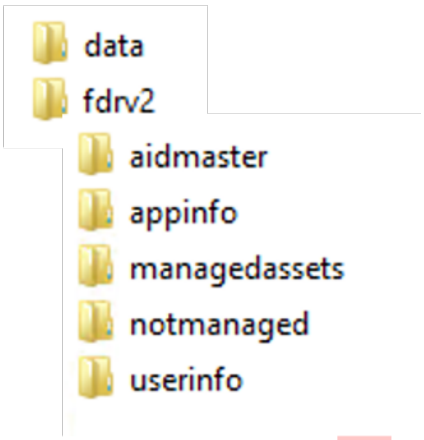
The FDR TA for Splunk does require the SQS queue URL in the input configuration, however this is only to get specific information to connect to the FDR S3 bucket. The FDR TA for Splunk does not communicate with the AWS SQS infrastructure but instead communicated directly with the S3 bucket. This provides significant benefits over the legacy client configuration:

1. It allows multiple clients to collect data from the FDR S3 bucket without needing to rely on the SQS queue for tracking.
2. It allows specific/different FDR data collections to be collected a different time intervals, reducing the impact on the collecting system(s).
3. It allows different Heavy Forwarders/ Inputs Data Manager (IDM) to collect different data types, enabling a distributed collection architecture.
4. It provides the ability to download and index specific data for use by different groups/teams while ensuring proper access controls to sensitive 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

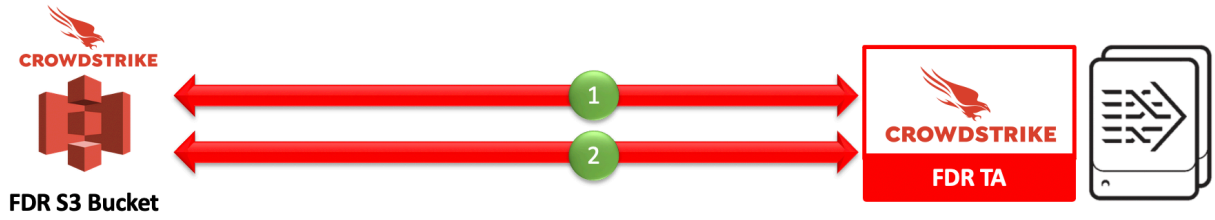
    The following events require the Falcon Insight module
  - **fdrv2 - aidmaster**: contains basic host information collected by the Falcon sensor
  - **fdrv2 - managedassets**: contains basic network configuration information collected by the Falcon sensor
  - **fdrv2 - notmanagedassets**: contains basics 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

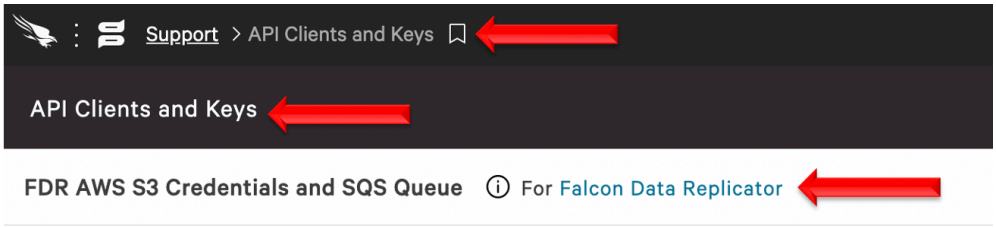
The CrowdStrike FDR TA performs the same API calls at each time interval that's configured within the specific TA input:



1. The TA accesses the CrowdStrike FDR S3 bucket and gets a list of files matching the desired event category and timeframe
2. The TA downloads the list of files identified, decompresses them, filters if necessary and posts the data to Splunk

Validating that FDR is Enabled

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



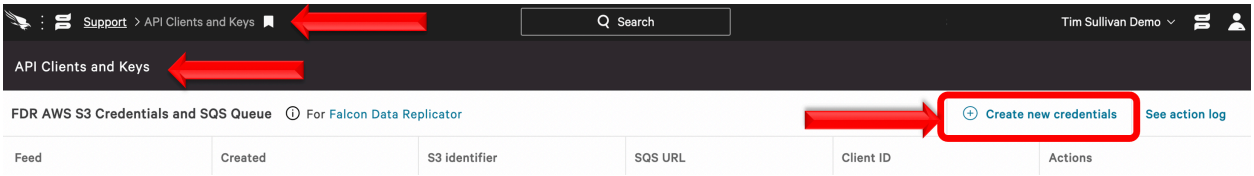
1. Access the CrowdStrike Falcon user interface (UI) with an account that is able to view/create the API clients and keys page
2. Navigate to 'Support'>'API Client and Keys' page
3. Validate that 'FDR AWS S3 Credentials and SQS Queue' is present

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

Generating/Collecting FDR Credentials

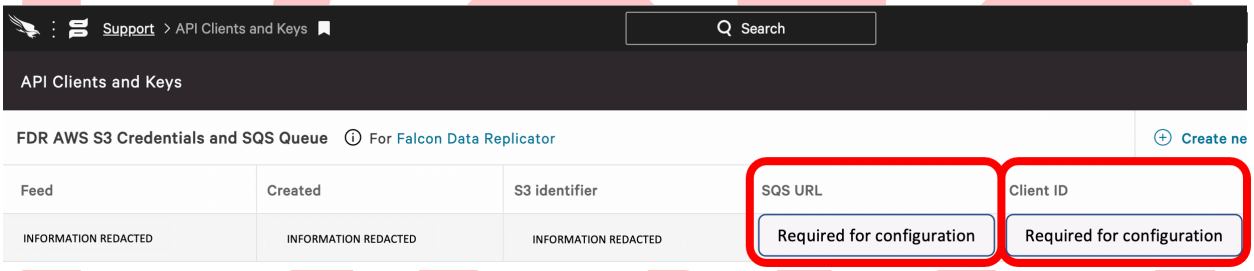
The FDR TA requires the FDR credentials that are located in the CrowdStrike Falcon UI in order to access the data. These can be existing FDR credentials or can be newly generated credentials.

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 not available via the UI, with the exception of when the credential is created, it is still required for configuration. 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 TA uses the SQS URL information for access but does **NOT** use nor access the SQS Queue itself. Therefore, existing FDR credentials that are already in use can used without the SQS Queue will be impacted.

Proxy Considerations

The CrowdStrike FDR Add-On communicated with the AWS S3 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.

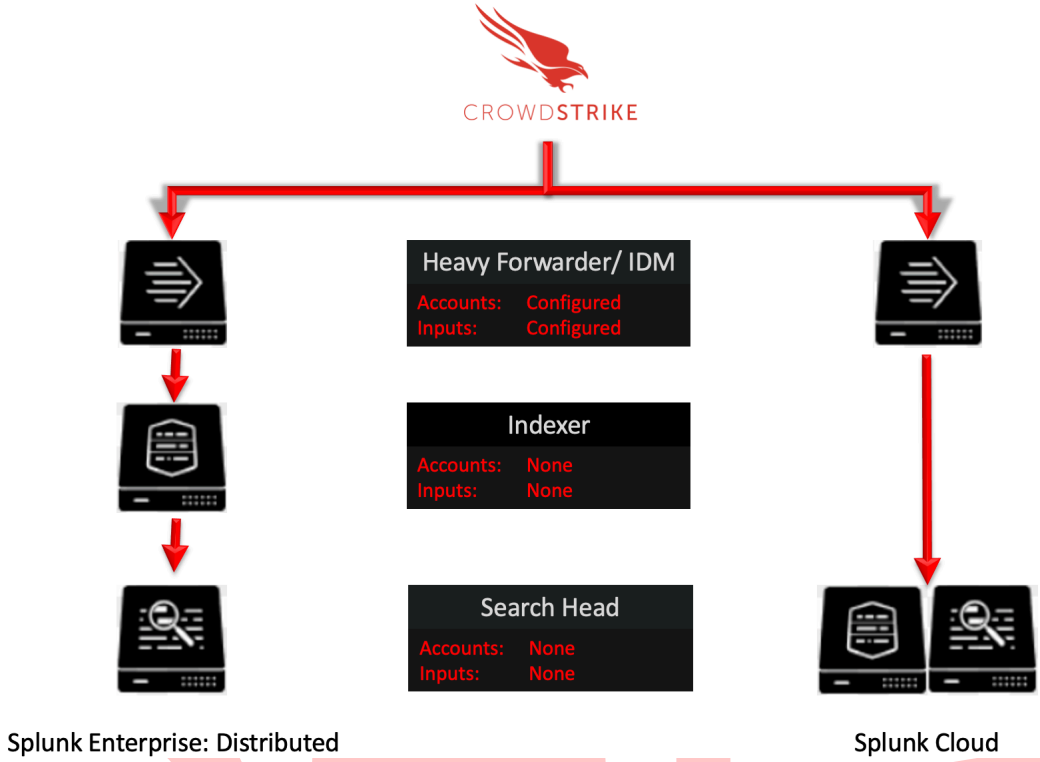
Splunk Indexer(s): The TA can be installed to provide field mapping and search macro support. The TA should be deployed without any accounts or inputs configured and any search macros should be properly configured for use. If a custom index is going to be used, then it should be created here.

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.

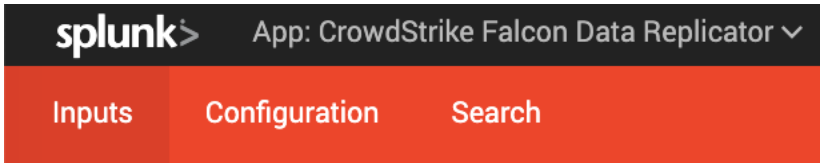
The following diagram shows the flow of data from the CrowdStrike FDR S3 bucket and the FDR TA configuration within a distributed Splunk Enterprise and Splunk Cloud environment:



Configuring the TA

TA Layout

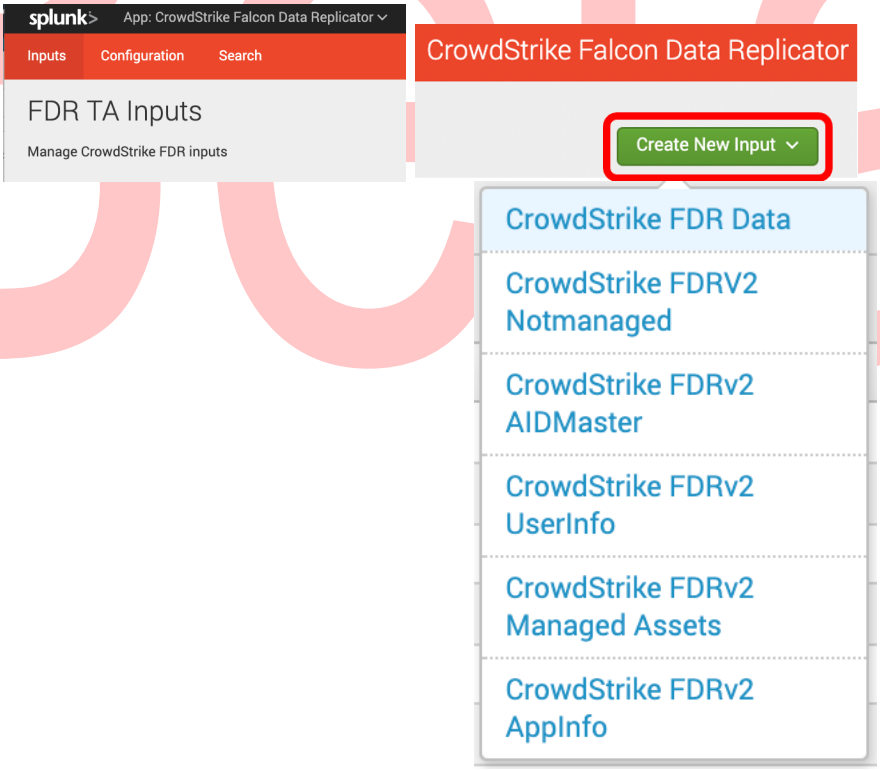
The TA contains 3 sections.



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

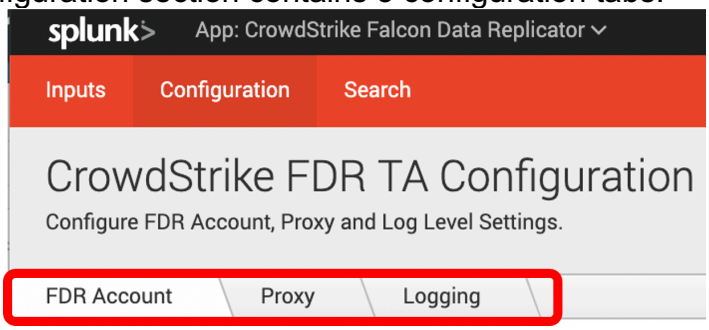
Inputs Section

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

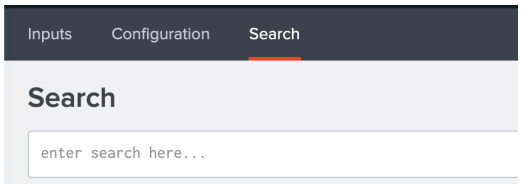
The Configuration section contains 3 configuration tabs:



- **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.

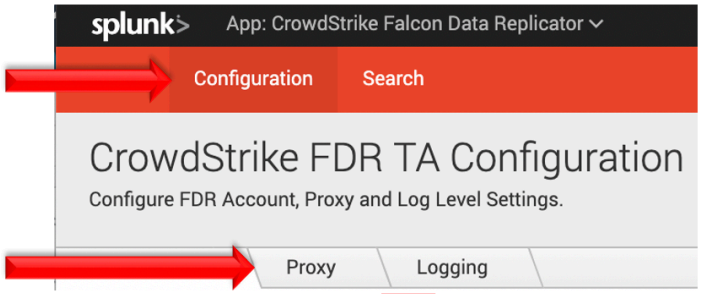


Configuring the TA to collect data

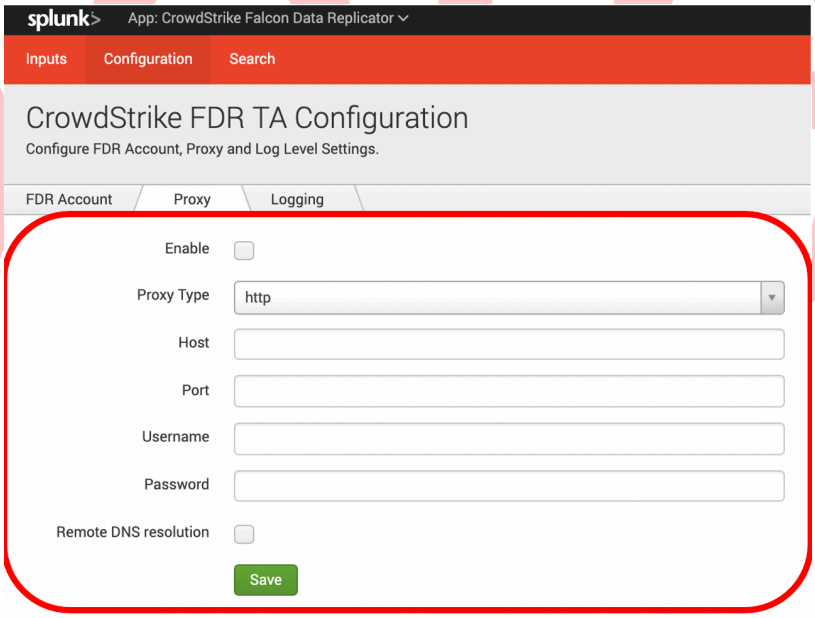
NOTE This action should only be performed on a Splunk Heavy Forwarder or Splunk IDM

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 S3.



2. Configure the following fields as appropriate:

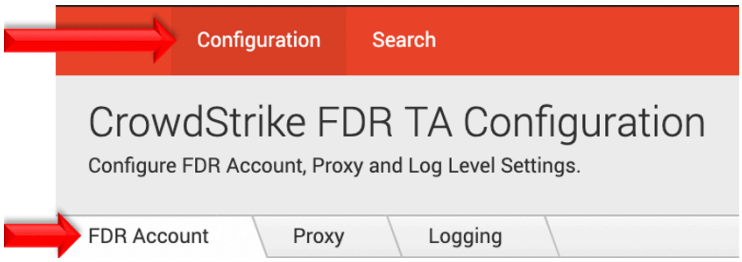


- **Enable**: This checkbox is used to enable/disable the proxy settings
- **Proxy Type**: This dropdown is used to select the proxy type
- **Host**: The hostname/IP address for the proxy server
- **Port**: The communication port for the proxy server
- **Username**: The authentication username for the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

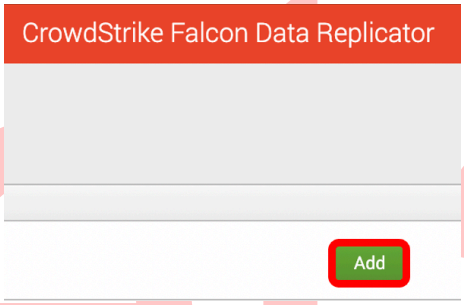
- **Password**: The authentication password for the proxy (optional)
- **Save**: This button is used to safe the configuration

Configure an Account

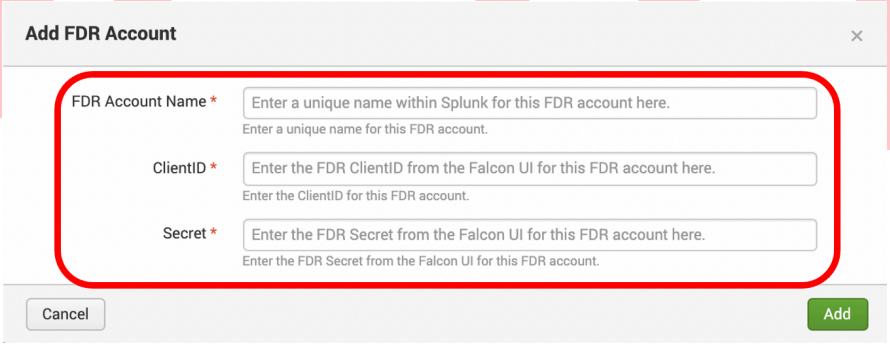
1. An account is configured using an FDR credential from the CrowdStrike Falcon UI.
2. An account is created under the Configuration section, FDR Account tab:



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



4. Configure the following fields:

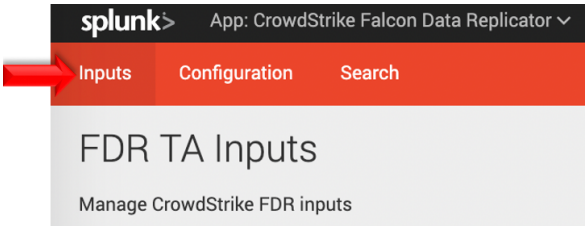


- **FDR Account Name**: A name unique for the Splunk instance
- **ClientID**: The ClientID of the FDR credential created in the CrowdStrike Falcon UI.
- **Secret**: The Secret of the FDR credential created in CrowdStrike Falcon UI.

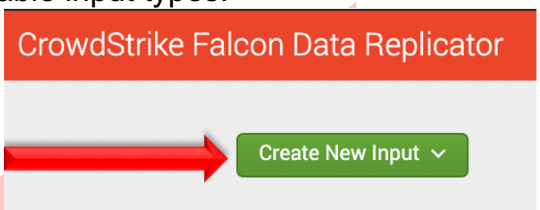
5. Click the ‘Add’ button in the bottom right corner to save the account.

Creating an Input

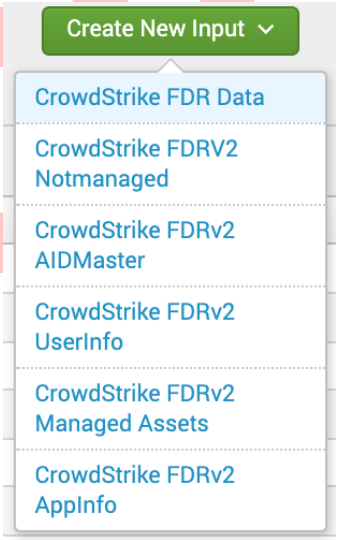
1. An input will require a valid FDR account be created already.
2. An input is created under the Inputs section:



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

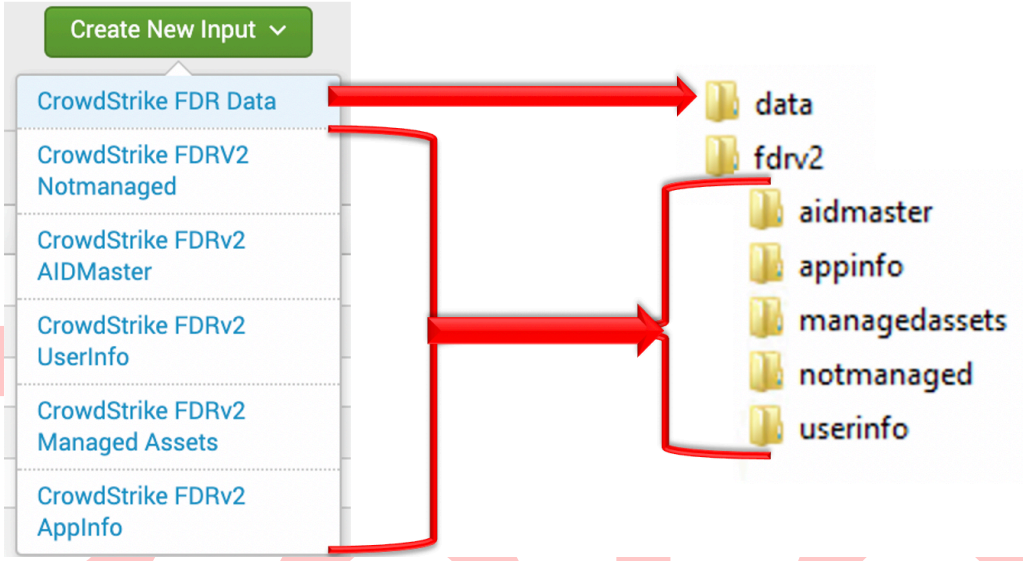


4. Select the input type to configure.



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.



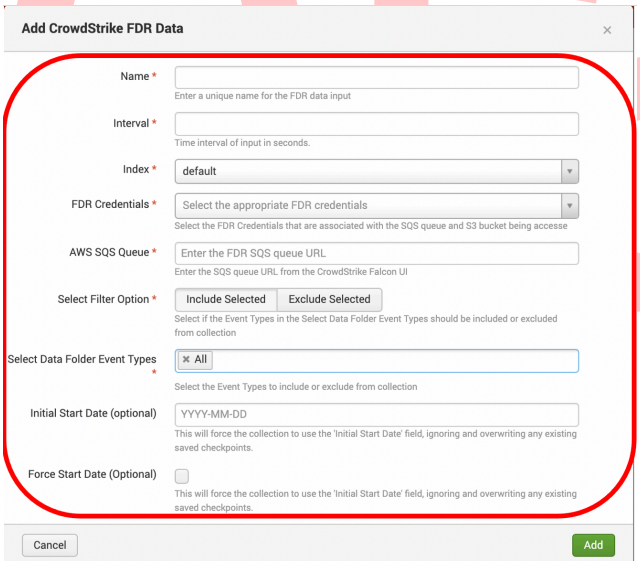
- **CrowdStrike FDR Data**: Collects information in the 'Data' folder within FDR
  - This input provides the ability to filter by event_simpleName and event_types field values
- **CrowdStrike FDRv2 Notmanaged**: Collects information in the 'FDRv2', 'Notmanaged' folder within FDR
- **CrowdStrike FDRv2 AIDMaster**: Collects information in the 'FDRv2', 'AIDMaster' folder within FDR
- **CrowdStrike FDRv2 UserInfo**: Collects information in the 'FDRv2', 'UserInfo' folder within FDR
- **CrowdStrike FDRv2 Managed Assets**: Collects information in the 'FDRv2', 'Managed' folder within FDR
- **CrowdStrike FDRv2 AppInfo**: Collects information in the 'FDRv2', 'UserInfo' folder within FDR

Configuring CrowdStrike FDR Data Inputs

The FDR Data Input contains the sensor telemetry and event data. This input has the ability to provide filtering functionality based on the FDR 'event_simpleName' filed value. Filtering can either be either inclusive or exclusive depending on the requirements. Since the CrowdStrike FDR TA does leverage the AWS SQS Queue for message tracking, it is possible to create multiple inputs for a single FDR S3 bucket. However, because of the way that the data is currently maintained in FDR, all the data will need to examined to determine if it matches the filtering criteria.

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

2. Configure the appropriate fields:



- **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 Credentials**: (required) The appropriate FDR credential set configured in the 'FDR Account' tab under 'Configuration'

- **AWS SQS Queue**: (required) The SQS Queue URL listed in the CrowdStrike Falcon UI for the particular FDR S3 bucket
- **Select Filter Option**: (required) Indicates how any filtering options should act
    - **Include Selected**: Events with 'event_simpleName' field values matching those associated with selected Event Types will be processed and sent to Splunk.
    - **Exclude Selected**: Events with 'event_simpleName' field values matching those associated with selected Event Types **will not** be processed and sent to Splunk.
- **Select Data Folder Event Types**: (required) Indicates what groups of events are within scope of the selected filtering option (all is the default)
    - This selection has prepopulated options but does also allow for custom groups of 'event_simpleName' selections.
- **Initial Start Date**: (optional) A date in YYYY-MM-DD format that serves as a starting point from which to collect data. This is the date that the data was posted to the FDR S3 bucket not the date of the event itself.
- **Force Start Date**: (optional) Forces the TA to collect data regardless of checkpoints from previous collections.
    - This setting should be cleared after the it's been utilized to prevent the next collection from starting at the same date
    - Utilizing this setting will overwrite any existing time stamp for the input

3. Click the 'Add' button in the bottom right corner to save and active the input.

Configuring CrowdStrike FDRv2 Based Inputs

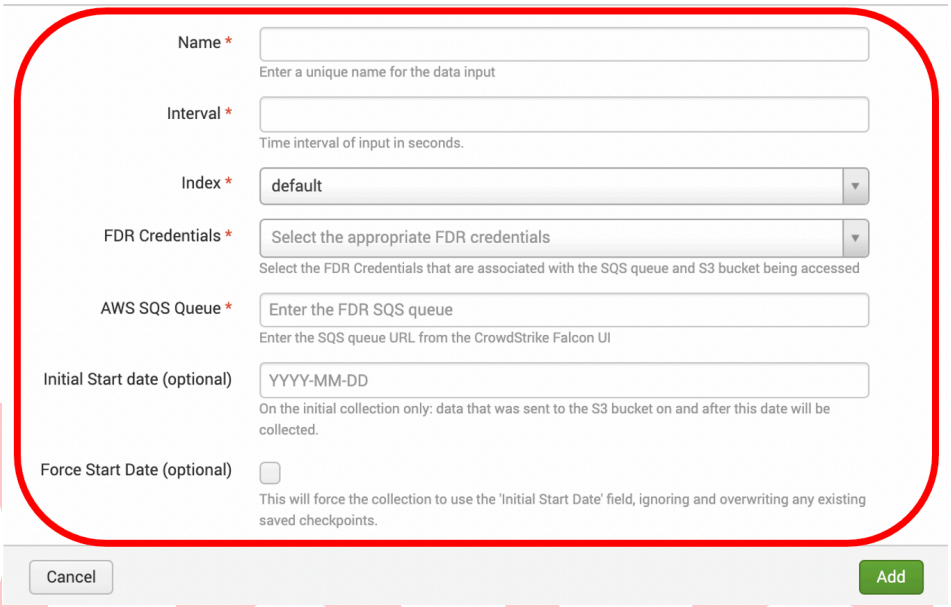
All FDRv2 based Inputs share the same configuration settings. These currently include:

- CrowdStrike FDRV2 Notmanaged
- CrowdStrike FDRv2 AIDMaster
- CrowdStrike FDRv2 UserInfo
- CrowdStrike FDRv2 Managed Assets
- CrowdStrike FDRv2 AppInfo

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:



- **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 Credentials**: (required) The appropriate FDR credential set configured in the ‘FDR Account’ tab under ‘Configuration’
- **AWS SQS Queue**: (required) The SQS Queue URL listed in the CrowdStrike Falcon UI for the particular FDR S3 bucket
- **Initial Start Date**: (optional) A date in YYYY-MM-DD format that serves as a starting point from which to collect data. This is the date that the data was posted to the FDR S3 bucket not the date of the event itself.
- **Force Start Date**: (optional) Forces the TA to collect data regardless of checkpoints from previous collections.
    - This setting should be cleared after the it’s been utilized to prevent the next collection from starting at the same date
    - Utilizing this setting will overwrite any existing time stamp for the input

3. Click the ‘Add’ button in the bottom right corner to save and active the input.

Data Input Filters

The FDR Data input provides for the ability to filter events based on the 'event_simpleName' or the 'event_type'* field value within the raw data. These fields can be found under the 'data' field within the processed FDR data – 'data.event_simpleName' and 'data.event_type' respectively.

*Note as of the writing of this document only Zero Trust Host Assessments currently leverage the 'event_type' field.

Data Input Filters – Standard Collections

CrowdStrike will provide groups of event_simpleName/eventtype field values that have certain associations to facilitate making filtering easier for some use cases. These collections will contain specific field values available at the time that they are released and may not contain all the field values desired. It is highly recommended that they be reviewed to ensure that the expected data will be collected/filtered.

Collections will be added and revised by CrowdStrike as needed.

Data Input Filters – Custom Collections

If there is a specific use case that requires a custom list of event_simpleName/event_type field values, the TA currently provides the ability to create 3 custom filter lists. These lists are located in the 'FDR_Event_Types.py' file within the 'bin' folder of the TA and are labeled 'custom01', 'custom02' and 'custom03'.

```
'custom01' : [],

'custom02' : [],

'custom03' : []
```

These lists must maintain their current labels. They can be populated with the field values desired but must be populated in the correct format (Python list format). The format requires that:

1. The field value be surrounded by a set of single quotes or a set of double quotes.
2. Each value should be separated by a comma.

The proper format can be observed by examining the other lists in the file.

WARNING:
Failure to maintain the proper syntax structure in this file can result in loss of filtering functionality and/or complete loss of TA functionality.

Search Macros

The FDR TA contains 7 configurable search macros:

Name	Definition	Arguments
cs_fdr_data_get_index	index=*	
cs_fdrv2_aidmaster_get_index	index=*	
cs_fdrv2_appinfo_get_index	index=*	
cs_fdrv2_managed_get_index	index=*	
cs_fdrv2_notmanaged_get_index	index=*	
cs_fdrv2_userinfo_get_index	index=*	
cs_fdr_data_input(1)	`cs_fdr_data_get_index` "ta_data.Input"=$input$	input

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.

In addition, there is a 7th Search Macro that is used to search for a specific FDR Data Input name. This Search Macro requires the `cs_fdr_data_get_index` macro to be properly configured prior to use. The Search Macro takes 1 input (which is identified by the '1' in parenthesis) which should be input when use.

For example, if an input was configured to only collect events for the 'Zero Trust Host Assessment' Event Types and the input name was "Zero_Trust_Assessments". The Search Macro would be input into Splunk search bar as `cs_fdr_data_input(Zero_Trust_Assessments)`.

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
- S3 Access based errors

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

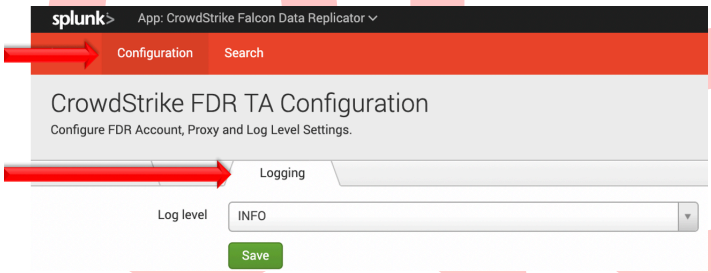
- Proxy based issues
- Firewall based issues
- Network connectivity issues
- Authentication issues (based on misconfigured credentials)
- Splunk CIM field mapping

Configuring the TA to collect log data

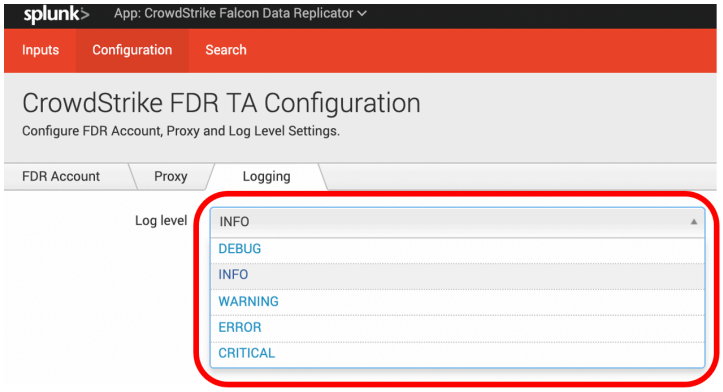
The TA logging level is set to 'info' by default and will only log a minimal amount of information. To properly troubleshoot issues with the TA the logging level should be set to 'debug'.

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.

Contacting Support

1. Ensure that the OAuth2 credential has been scoped correctly
2. Set the TA log level to 'DEBUG'
3. Repeat and record the action(s) that are associated with the issue you are reporting
4. Download the all log files containing 'ta_crowdstrike_falcon_data_replicator' under the $Splunk/var/log/splunk/ directory
5. Record the following information about the Splunk system:
    - Splunk environment type
    - Splunk version
    - TA version
6. Identify the types of networks devices that the connection will traverse and ensure that they have been properly configured
7. Collect API audit logs from the Falcon instance for the time frame when the issue is occurring
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

Appendix A: Current Pre-Defined Filter Lists

AWS:

AwsEc2Image
AwsEc2Instance
AwsEc2NetworkAcl
AwsEc2NetworkAclEntry
AwsEc2NetworkInterface
AwsEc2NetworkInterfacePrivateIP
AwsEc2SecurityGroup
AwsEc2SecurityGroupRule
AwsEc2Subnet
AwsEc2Volume
AwsEc2Vpc
AwsIamAccountAlias

AZURE:

AzureApplicationFirewallRule
AzureDisk
AzureFirewall
AzureFirewallRuleCollection
AzureIpConfiguration
AzureNatFirewallRule
AzureNetworkFirewallRule
AzureNetworkInterface
AzureNetworkSecurityGroup
AzureNetworkSecurityGroupRule
AzurePrivateEndpoint
AzurePublicIpAddress
AzureResourceGroup
AzureSubnet
AzureSubscription
AzureVirtualMachine
AzureVirtualMachineState
AzureVirtualNetwork
AzureVirtualNetworkPeering

USB Events:

DcUsbConfigurationDescriptor
DcUsbDeviceBlocked
DcUsbDeviceConnected
DcUsbDeviceDisconnected
DcUsbDevicePolicyViolation
DcUsbDeviceWhitelisted
DcUsbEndpointDescriptor
DcUsbHIDDescriptor
DcUsbInterfaceDescriptor

Mobile:

AccessoryConnected
AccessoryDisconnected
AndroidIntentSentIPC
AndroidManifestFragmentData
AndroidManifestXmlUploaded
AndroidModuleStateUpdate
APKMetadata
AppSideLoaded
BootLoaderStatus
ClipboardCopy
ClipboardPaste
DeactivateMobileSensorResponse
DebuggableFlagTurnedOn
DebuggedState
DeveloperOptionsStatus
DexFileWritten
DnsRequestBlocked
DNSRequestDetectInfo
DnsRequestResult
DuplicateInstallFromPlayStore
HarmfulAppData
InstallFromUnknownSourcesStatus
LockScreenStatus
MobileAppIdentifiers
MobileAppsManifest
MobileDetection
MobileOsIntegrityStatus
MobilePowerStats
NetworkConnectIP4DetectInfo
NetworkConnectIP6DetectInfo
ObjCRuntimeAltered
PathUnexpectedlyReadable
ProcessWitness
RemediationActionKillIP4Connection
RemediationActionKillIP6Connection
SafetyNetCheckFailed
SafetyNetCompatibilityStatus

SecureTrafficDecrypted
SELinuxStatus
StorageEncryptionStatus
SystemPartitionAltered
SystemPartitionStatus
TrampolineDetected
UncontainerizeAppResponse
UnexpectedDynamicLibraryLoaded
UnexpectedEnvironmentVariable
UnexpectedFileFound
VerifyAppsDisabled
WiFiConnect
WiFiDisconnect

CrowdStrike Sensor Communications:

AgentConnect
AgentOnline
ChannelDataDownloadComplete
ChannelVersionRequired
CurrentSystemTags
ECBDownloadComplete
LFODownloadConfirmation
LfoUploadDataComplete
LfoUploadDataFailed
LfoUploadDataUnneeded
NetworkUncontainmentCompleted
SensorHeartbeat
AgentConnect
AgentOnline
ChannelDataDownloadComplete
ChannelVersionRequired
CurrentSystemTags
ECBDownloadComplete
LFODownloadConfirmation
LfoUploadDataComplete
LfoUploadDataFailed
LfoUploadDataUnneeded
NetworkUncontainmentCompleted
SensorHeartbeat

Network

DnsRequest
DnsRequestBlocked
DNSRequestDetectInfo
DnsRequestResult
DnsServerSigRedExploitAttemptEtw
HttpRequestDetect
HttpVisibilityStatus
LocalIpAddressIP4
LocalIpAddressIP6
LocalIpAddressRemovedIP4
LocalIpAddressRemovedIP6
NeighborListIP4
NeighborListIP6
NetShareAdd
NetShareDelete
NetShareSecurityModify
NetworkCloseIP4
NetworkCloseIP6
NetworkConnectIP4
NetworkConnectIP4Blocked
NetworkConnectIP4DetectInfo
NetworkConnectIP6
NetworkConnectIP6Blocked
NetworkConnectIP6DetectInfo
NetworkContainmentCompleted
NetworkListenIP4
NetworkListenIP6
NetworkModuleLoadAttempt
NetworkReceiveAcceptIP4
NetworkReceiveAcceptIP6
NetworkUncontainmentCompleted
RawBindIP4
RawBindIP6
SuspiciousDnsRequest

About CrowdStrike

CrowdStrike® Inc. (Nasdaq: CRWD), a global cybersecurity leader, is redefining security for the cloud era with an endpoint protection platform built from the ground up to stop breaches. The CrowdStrike Falcon® platform's single lightweight-agent architecture leverages cloud-scale artificial intelligence (AI) and offers real-time protection and visibility across the enterprise, preventing attacks on endpoints on or off the network. Powered by the proprietary CrowdStrike Threat Graph®, CrowdStrike Falcon correlates over 3 trillion endpoint-related events per week in real time from across the globe, fueling one of the world's most advanced data platforms for security.

There's only one thing to remember about CrowdStrike: **We stop breaches**.