Micro Focus Security
ArcSight Micro Focus Security

Software Version: 8.3.0

SmartConnector for ArcSight Common Event Format Hadoop

Document Release Date: February 2022
Software Release Date: February 2022

Legal Notices

Micro Focus
The Lawn
22-30 Old Bath Road
Newbury, Berkshire RG14 1QN
UK

https://www.microfocus.com

Copyright Notice

Trademark Notices

Documentation Updates

The title page of this document contains the following identifying information:

- Software Version number
- Document Release Date, which changes each time the document is updated
- Software Release Date, which indicates the release date of this version of the software

To check for recent updates or to verify that you are using the most recent edition of a document, go to:

https://www.microfocus.com/support-and-services/documentation

Support

Contact Information

Phone	A list of phone numbers is available on the Technical Support Page: https://softwaresupport.softwaregrp.com/support-contact-information
Support Web Site	https://softwaresupport.softwaregrp.com/
ArcSight Product Documentation	https://community.softwaregrp.com/t5/ArcSight-Product-Documentation/ct-p/productdocs

About this PDF Version of Online Help

This document is a PDF version of the online help. This PDF file is provided so you can easily print multiple topics from the help information or read the online help in PDF format. Because this content was originally created to be viewed as online help in a web browser, some topics may not be formatted properly. Some interactive topics may not be present in this PDF version. Those topics can be successfully printed from within the online help.

Contents

Configuration Guide for SmartConnector for ArcSight Common Event Format Hadoop 5
Product Overview 5
Common Event Format Implementation 7
Configuring Hadoop DFS API Security Settings 7
Preparing to install the SmartConnector 8
Installing and Configuring the SmartConnector 8
Modifying Parameters to Optimize Performance 9
Device Event Mapping to ArcSight Data Fields 10
Troubleshooting 10

Troubleshooting 12

Send Documentation Feedback 13

Configuration Guide for SmartConnector for ArcSight Common Event Format Hadoop

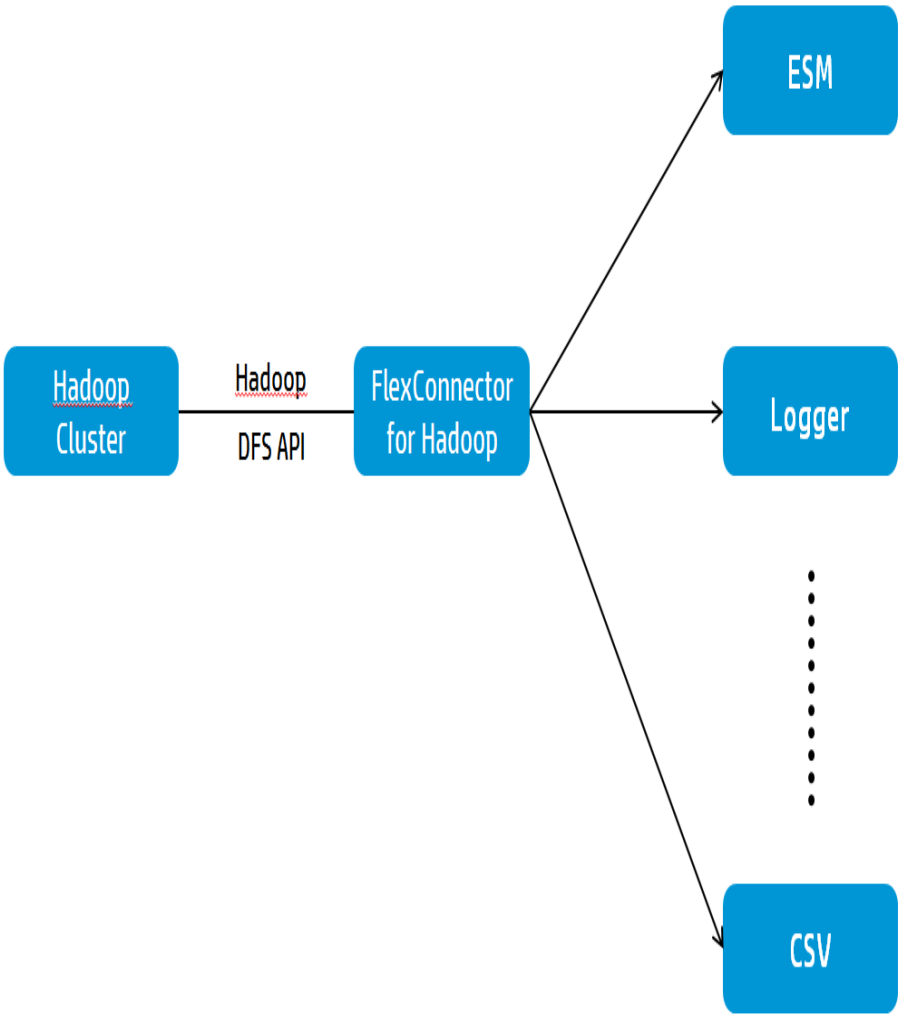
This guide provides information to install the SmartConnector for ArcSight Common Event Format (CEF) Hadoop and configure it for event collection.

Hadoop DFS version 2.5.1 is supported.

Product Overview

The Hadoop Distributed File System (HDFS) splits and stores large data files for processing across Hadoop machines in a cluster. This distributed file system provides high-throughput access to application data.

The SmartConnector for CEF Hadoop provides a configurable method to collect any event (or record) in CEF and stored in HDFS and forward the events to ESM or other destinations.



This SmartConnector is designed to collect data from static files that are either in compressed format `.gzip` or `.bz2` or in plain text. The source folder can contain files in all three file formats. The HDFS API determines the compression type automatically by using the file extension of the compressed file during datacollection.

The file compression type .lzo is not currently certified for data collection with this SmartConnector.

This SmartConnector can collect data from either local files or remote files. It collects data in batch mode, with no new events being written to the files.

It can collect files from a single folder that contains multiple files. However, it cannot collect data from subfolders.

By default, the SmartConnector checks for new files to collect every `3600000 msec` or after it is done processing files from the previous collection, whichever is earlier. To change the default monitoring interval, see Modifying Parameters to Optimize Performance.

If the connector stops collecting data for any reason, it starts collecting data from the point it left off when data collection resumes. Files that are processed by the connector are moved to a processed-files folder, and the extension `.processed` is appended to the processed files. For more information, see Configuring Hadoop DFS API Security Settings.

Common Event Format Implementation

The Common Event Format (CEF) standard format, developed by ArcSight, lets vendors and their customers quickly integrate their product information into ESM. CEF is an open log management standard that simplifies log management, letting third parties create their own device schema that are compatible with a standard that is used industry-wide for normalizing security events. Technology companies and customers can use the standardized CEF format to facilitate data collection and aggregation, for later analysis by an enterprise management system.

For more information about CEF, see the *Implementing ArcSight Common Event Format (CEF)* Guide. It defines the CEF protocol and provides details about how to implement the standard. It details the header and predefined extensions used within the standard as well as how to create user defined extensions. It also includes a list of CEF mappings as well as supported date formats.

Configuring Hadoop DFS API Security Settings

You must change certain default security properties for the Hadoop cluster to allow the SmartConnector for CEF Hadoop to collect data. These changes must be made for the Name node, as it acts as a Master node. Make sure that the following properties are configured as specified because they are checked before other access control checks:

The value for `hadoop.security.authorization` property in the `${HADOOP_CONF_DIR}/core-site.xml` file must be set to false as shown in the following XML block:

```
<property>
   <name>hadoop.security.authorization</name>
   <value>false</value>
   <description>Service level authorization params.</description>
   </property>
```

The value for `security.client.protocol.acl` property in the `${HADOOP_CONF_DIR}/hadoop-policy.xml` file must be ***** (an asterisk) as shown in the XML block:

```
<property>
   <name>security.client.protocol.acl</name>
   <value>*</value>
   </property>
```

The value for `dfs.permissions=false` property in `${HADOOP_CONF_DIR}/hdfs-site.xml.` must be set to false.

This property creates a processed-files folder inside the configured base folder from which the files are read and moves the files after they are processed by the connector. If a `Processed` folder does not exist, the connector creates one. If due to issues such as permissions issues aconnector cannot create the folder, then it logs an error message and leaves the processed files in the base folder.

The processed files are appended with `.processed`.

Preparing to install the SmartConnector

Before you install any SmartConnectors, make sure that the Micro Focus ArcSight products with which the connectors will communicate have already been installed correctly (such as ArcSight ESM or ArcSight Logger).

For complete product information, refer to the *Administrator's Guide to ArcSight Platform*, available on ArcSight Documentation.

If you are adding a connector to the ArcSight Management Center, see the *ArcSight Management Center Administrator's Guide* available on ArcSight Documentation for instructions.

Before installing the SmartConnector, make sure that the following are available:

- Local access to the machine where the SmartConnector is to be installed
- Administrator passwords

Installing and Configuring the SmartConnector

1. Start the installation wizard.
2. Follow the instructions in the wizard to install the core software.
3. Specify the relevant Global Parameters, when prompted.

4. Select **ArcSight Common Event Format Hadoop** and click **Next**.
5. Specify the following parameters, then click **Next**.

Parameter	Description
Hadoop Cluster IP and Port	Enter the IP address and port number of the Name Node (also known as the Master Node).
Core Site File Path	Enter the file path to the Hadoop Core Site.
HDFS Site File Path	Enter the file path to the Hadoop Distributed File System Site.
Log File Path	Enter the file path to the Hadoop log file.
Log File Pattern	Enter a pattern for data file names. Using the default value (event.*), the connector will look for log files starting with "event."

6. Select a destination and configure parameters.
7. Specify a name for the connector.
8. If you have selected ArcSight Manager as the destination, the certificate import window for the ArcSight Manager is displayed. Select **Import the certificate to the connector from destination** and click **Next**. (If you select **Do not import the certificate to connector from destination**, the connector installation will end.) The certificate is imported and the **Add connector Summary** window is displayed.
9. Select whether you want to run the connector as a service or in the standalone mode.
10. Complete the installation.
11. Run the SmartConnector.

For instructions about upgrading the connector or modifying parameters, see SmartConnector Installation and User Guide.

Modifying Parameters to Optimize Performance

After SmartConnector installation, you can use the agent.properties file to modify values for parameters to optimize connector performance.

1 From the `$ARCSIGHT_HOME\current\user\agent` directory open the file `agent.properties` in a pure ASCII text editor.

2 In the `agent.properties` file, modify the values for following parameters as needed:

Parameter	Default Value	Description
filecheckinterval	3600000 msec	Time interval in milliseconds at which the connector retrieves events from the Hadoop cluster.
fileupdatewaitinterval	10000 msec	Time interval in milliseconds to wait before starting to process a new file. The file must not have been modified in the last 10 seconds as a confirmation that the file transfer and writing is complete and the file is ready for processing.
processedfolderpath	/user/hadoop/processed	Path to the processed folder on the Hadoop cluster to move the files after they are processed.

4 Save the `agent.properties` file.

5 Restart the SmartConnector.

Device Event Mapping to ArcSight Data Fields

For device mappings for a product, refer to the vendor CEF documentation.

Information from vendors is formatted according to the CEF standard and sent to the ArcSight SmartConnector, which translates the data into an ArcSight event.

Troubleshooting

Java exception error - 'Failed to locate the winutils binary in the hadoop binary path'

This error can sometimes happen when you are running a connector in a Windows environment.

Microsoft technical support recommends that you download the compiled `winutils.exe` program from the following link, and save it to the `C:\hadoop\winutils\bin` directory: http://social.msdn.microsoft.com/Forums/windowsazure/en-US/28a57efb-082b-424b-8d9e-731b1fe135de/please-read-if-experiencing-job-failures?forum=hdinsight

Alternatively, add the `winutilpath` parameter with the path to the utility to agent.properties file in the `$ARCSIGHT_HOME\current\user\agent` directory. For example:

```
agents[0].winutilpath=c\:\\hadoop\\winutils\\
```

Java exception error about missing permission to move to the processed file

You might receive this message when the connector does not have permission to rename and move the file from the folder specified in the Log File Path parameter .

Make sure that the relevant account has full read/write permission, so that the connector can read, rename, or move it out to the processed log file path. The same rule applies to the processed log file path.

To change the permission, use a command such as `hadoop dfs -chmod a+w`.

Troubleshooting

Why am I getting a Java exception error - 'Failed to locate the winutils binary in the hadoop binary path'?

This error can sometimes happen when you are running a connector in a Windows environment.

Microsoft technical support recommends that you download the compiled `winutils.exe` program from the following link, and save it to the `C:\hadoop\winutils\bin` directory: http://social.msdn.microsoft.com/Forums/windowsazure/en-US/28a57efb-082b-424b-8d9e-731b1fe135de/please-read-if-experiencing-job-failures?forum=hdinsight

Alternatively, you can fix the problem by editing the `agent.properties` file (which can be found at `$ARCSIGHT_HOME\current\user\agent`) and adding the `winutilpath` parameter to enter the current path to the utility; for example:

```
agents[0].winutilpath=c\:\\hadoop\\winutils\\
```

Why am I getting a Java exception error about missing permission for moving to the processed file?

You could receive this message when the connector does not have permission for renaming and moving the file from the path you identified in the parameter **Log File Path** configured during the connector installation process (or **logfilepath** parameter specified in the `agent.properties` file). Make sure the folder this path specifies has full read/write permission for the relevant account (you can change it with a command, such as `hadoop dfs -chmod a+w`). The connector then can read the file, rename it, and move it out to the processed log file path. The same rule applies to the processed log file path.
please confirm that when customer used MySQL JDBC driver 5.1.38, they had issue to receivee events. And the workaround is to apply older driver 5.0.8, after that connector is able to received events.

Send Documentation Feedback

If you have comments about this document, you can contact the documentation team by email. If an email client is configured on this computer, click the link above and an email window opens with the following information in the subject line:

Feedback on SmartConnector for ArcSight Common Event Format Hadoop (Micro Focus Security ArcSight Connectors 8.3.0)

Just add your feedback to the email and click send.

If no email client is available, copy the information above to a new message in a web mail client, and send your feedback to Documentation-Feedback@microfocus.com.

We appreciate your feedback!