Cloudera Edge Management 1.2.2

MiNiFi Agent Quick Start

Date published: 2020-06-22
Date modified: 2020-11-23

https://docs.cloudera.com/

Legal Notice

Contents

MiNiFi agent quick start

Learn how to install and start using MiNiFi Java and C++ agents quickly.

Overview of Apache MiNiFi

Apache MiNiFi is an Apache NiFi project, designed to collect data at its source.

MiNiFi is developed with the following objectives in mind:

- Small and lightweight footprint
- Central agent management
- Data provenance generation
- NiFi integration for follow-on dataflow management and chain of custody information

Before you begin installing MiNiFi agent

Before you begin your MiNiFi agent installation, review the operating system and JDK support.

Operating System Support

Operating System	Version
RHEL/CentOS	7.0, 7.1, 7.2, 7.3, 7.4, 7.5, 7.6, 7.7
Debian	9
Ubuntu	16.04, 18.04
Windows	8, 10, Server 2012, Server 2012 R2, Server 2016, Server 2019

JDK Support

JDK	Version
OpenJDK	JDK8
Oracle JDK	JDK8

You can find how to download links for the following MiNiFi softwares in the *Obtaining the CEM Software* section:

- MiNiFi Java Agent
- MiNiFi C++ Agent
- MiNiFi Toolkit

Related Information

Obtaining the CEM software

Installing and starting MiNiFi Java

You have several options for installing and starting MiNiFi Java. Learn how to install and configure MiNiFi Java on Linux or on Windows.

Installing MiNiFi Java on Linux

Learn how to install MiNiFi Java on RHEL/CentOS, Ubuntu, Debian, and SLES.

About this task

To install MiNiFi Java on RHEL/CentOS, Ubuntu, Debian, SLES, complete the following steps:

Procedure

1. Download the tar.gz or zip file for the MiNiFi Java agent.

```
wget {java.tar.gz}
```

2. To install the MiNiFi Java agent, extract the file to your desired home directory.

Installing MiNiFi Java as a service on Linux

Learn how to install MiNiFi Java as a service.

Procedure

1. Navigate to the MiNiFi Java installation directory.
2. Enter:

```
bin/minifi.sh install
```

You can also specify a custom name for your MiNiFi Java installation, by specifying that name during your install command. For example, to install MiNiFi Java as a service and name it as dataflow, enter:

```
bin/minifi.sh install dataflow
```

Starting MiNiFi Java on Linux

After you download and install MiNiFi Java on Linux, you need to start MiNiFi Java.

About this task

You can start MiNiFi Java in the foreground, background, or as a service on Linux.

Procedure

1. Launching MiNiFi Java in the foreground: From a terminal window, navigate to the MiNiFi Java installation directory.
2. Enter:

```
bin/minifi.sh run
```

3. Launching MiNiFi Java in the background: From a terminal window, navigate to the MiNiFi Java installation directory.
4. Enter:

```
bin/minifi.sh start
```

5. Launch MiNiFi Java as a service: From a terminal window, enter:

```
sudo service minifi start
```

Installing MiNiFi Java on Windows

Learn how to install MiNiFi Java using a Windows MSI.

About this task

Before you begin your MiNiFi Java installation, ensure that you meet the following requirements:

- Install JDK 8.0 64 bit.
- Install Java to C:/java instead of C:/Program Files. Recent Windows versions mark everything in C:\Program Files as read only.
- Set the JAVA_HOME environment variable using the 8.3 style name conventions. For example: C:\Program\jdk1.8.0.
- Ensure JAVA_HOME is pointing to a 64-bit JRE/JDK.
- Ensure the Domain user has administrator privilege.
- Ensure that your system meets the minimum memory requirement for Windows which is 4GB.

Procedure

1. Extract the MiNiFi Java MSI files from the repo location https://archive.cloudera.com/web/CEM/windows/1.x/updates/<version>/minifi-<version>-<build-number>.msi to the location where you want to run the application.
2. Execute the MSI.

Configuring the MiNiFi Java MSI

The MSI file adds the Windows service for MiNiFi Java. The service is configured to be executed by either a local user in the computer, or a domain user in the Active Directory. Learn how to configure the MiNiFi Java MSI.

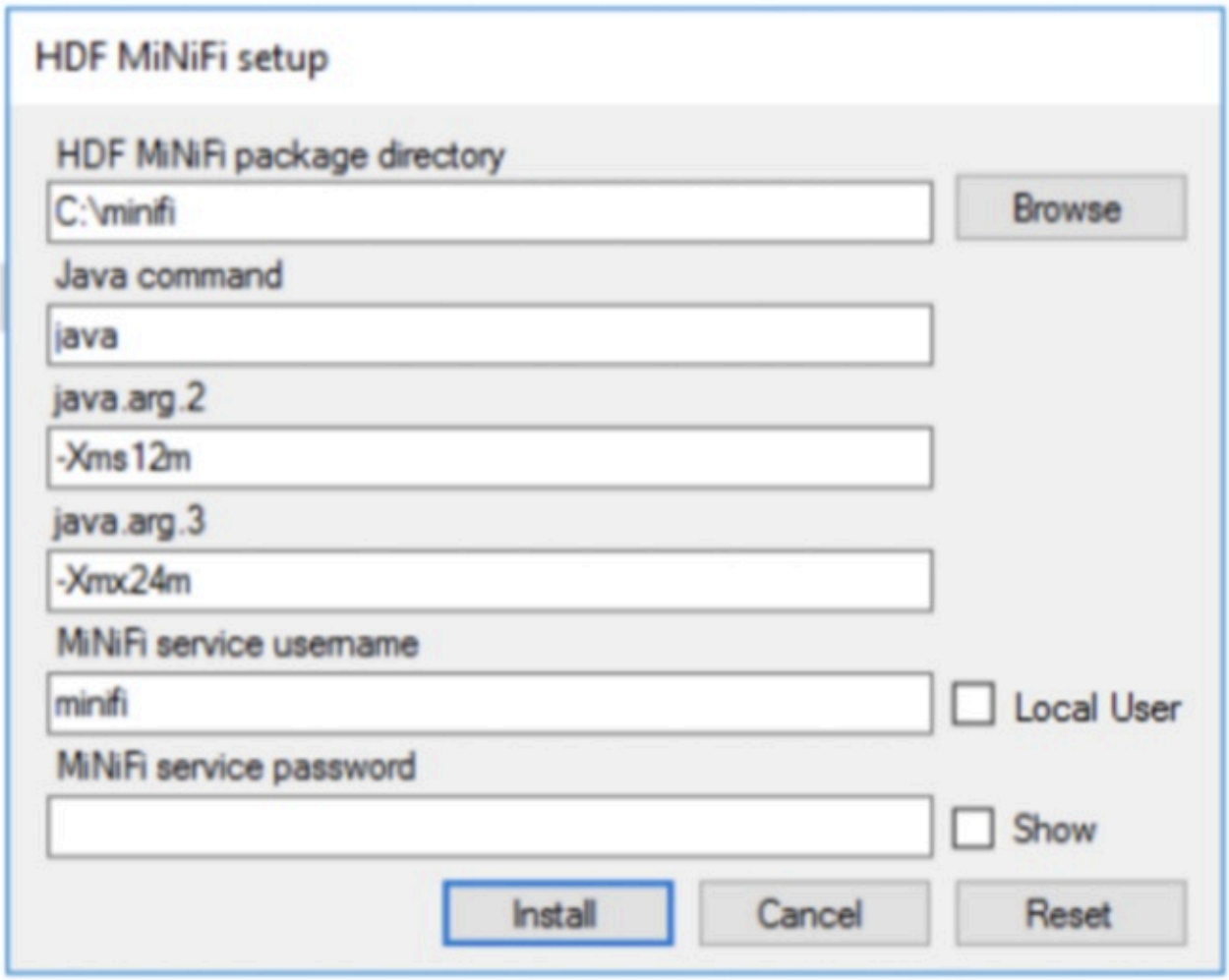
Using a Local User for MiNiFi Java Windows Service

There is no prerequisite to use a local user for the Windows service. The installer automatically sets up the user.

The installer also grants the following privileges to the specified user:

- SeCreateSymbolicLinkPrivilege
- SeServiceLogonRight

If the computer is a part of a domain, then the Local User checkbox appears in the MiNiFi setup window. Check the Local User checkbox to specify that a local user is used to execute the installed service.



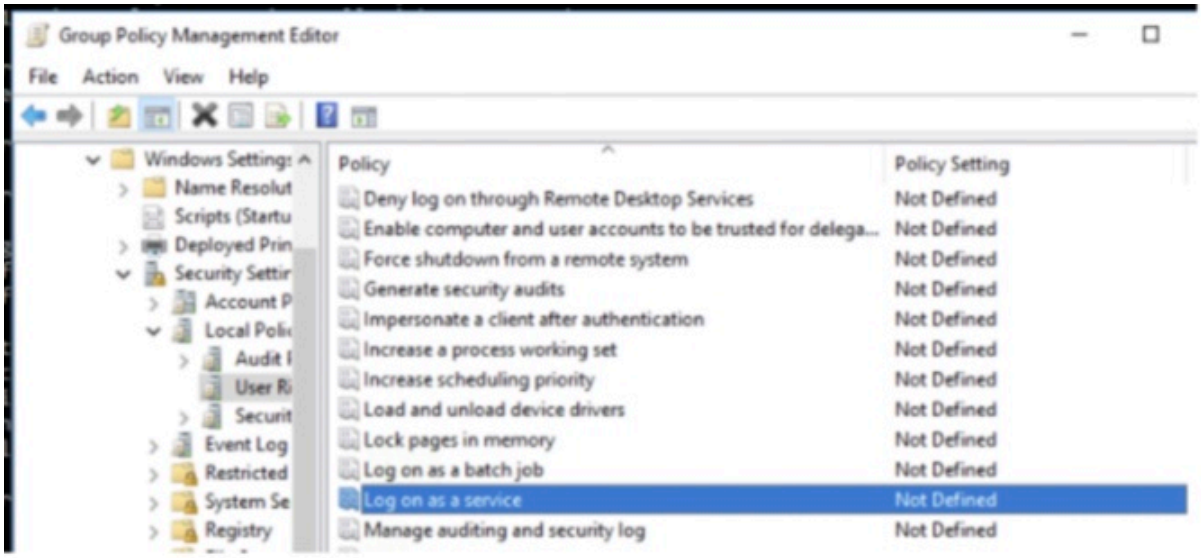
If a user specified in MiNiFi service username does not exist, the installer creates one with the specified MiNiFi service password. If the user already exists, the installer updates its password with the specified password.

Using a Domain User for MiNiFi Java Windows Service

Before you begin

- The computer must be part of the domain.
- The specified user must exist in the domain, and a correct password must be provided.
- ActiveDirectory PowerShell module must be available.

1. In the Group Policy Management Editor, set permission to Log on as a service.



2. Navigate to a machine where MiNiFi Java is installed and enter the following command:

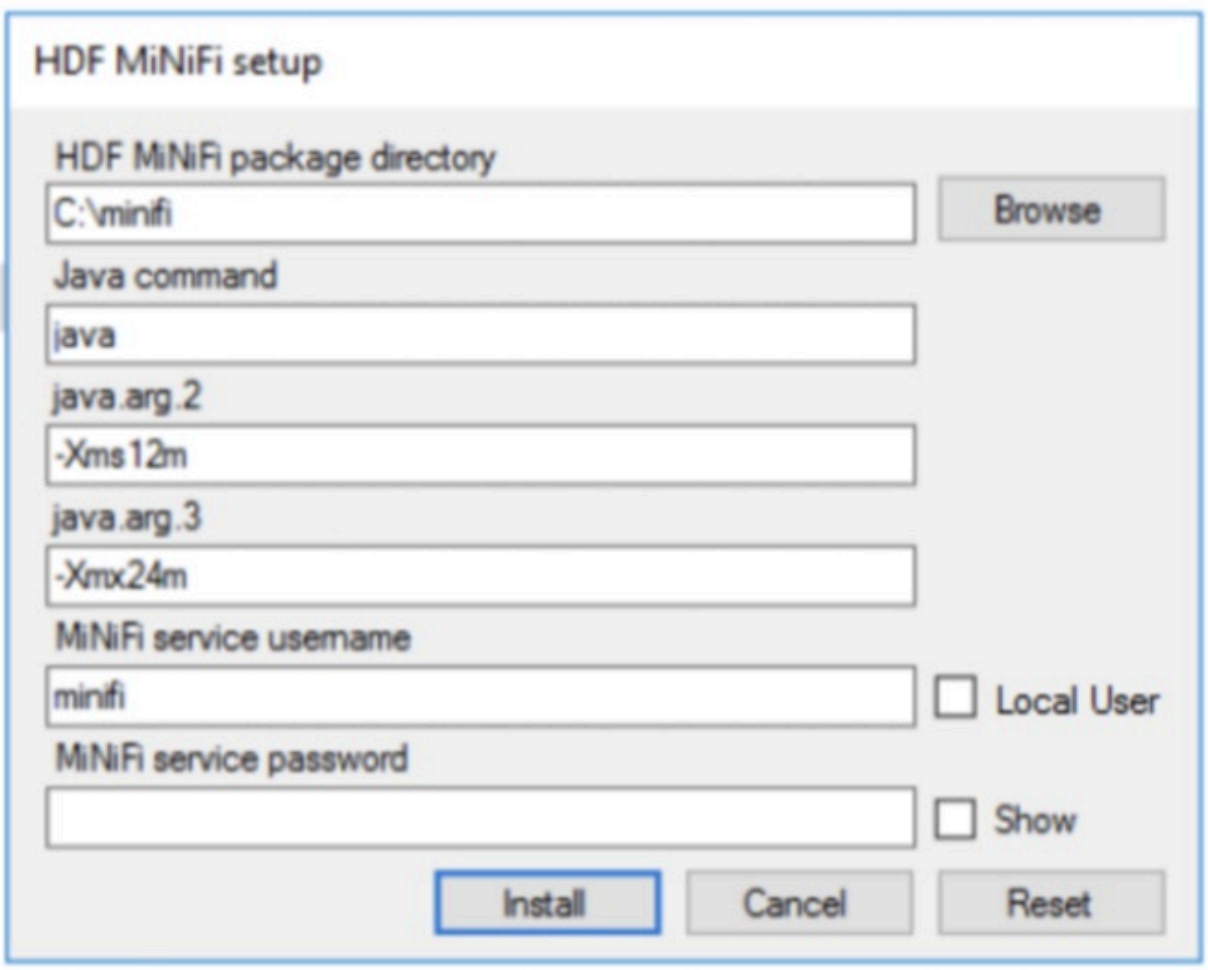
```
gpupdate
```

   The gpupdate command is a machine-wide command and can be executed from any directory on the MiNiFi machine.

3. Install the ActiveDirectory PowerShell module by entering the following in the PowerShell console:

```
Add-WindowsFeature RSAT-AD-PowerShell
```

4. In the MiNiFi setup window, uncheck the Local User checkbox, and then click Install.



Configuring C2 properties for MiNiFi Java

After you install the MiNiFi Java agent, update the configuration file.

About this task

The configuration file for MiNiFi Java is conf/bootstrap.conf.

Procedure

1. From the MiNiFi home directory, open the bootstrap.conf file.
2. Configure the Agent Class so that you can logically group MiNiFi instances according to their functionality. Specify the agent class:

```
nifi.c2.agent.class={AGENT_CLASS}
```

3. Configure the Agent ID. If you do not specify an Agent ID, MiNiFi generates a unique ID per agent instance.

```
nifi.c2.agent.identifier={AGENT_ID}
```

4. Configure your EFM Server endpoint:

```
nifi.c2.rest.url=http://{EFM_SERVER_IP}:10080/efm/api/c2-protocol/heartbeat
```

```
nifi.c2.rest.url.ack=http://{EFM_SERVER_IP}:10080/efm/api/c2-protocol/ac
knowledge
```

5. Configure the heartbeat interval:

```
nifi.c2.agent.heartbeat.period={HEARTBEAT_INTERVAL}
```

Starting MiNiFi Java on Windows

After you download and install MiNiFi Java, you can start MiNiFi Java in the foreground or as a service on Windows.

About this task

Launching MiNiFi Java in the foreground:

Procedure

1. From a command prompt window, navigate to the MiNiFi Java installation directory.
2. Enter the following command to launch MiNiFi Java in the foreground:

```
bin\run-minifi.bat
```

What to do next

Launching MiNiFi Java as a service:

You can start or stop the installed MiNiFi Java service from the Windows Service Manager.

Installing and starting MiNiFi C++

You have several options for installing and starting MiNiFi C++. Learn how to install and configure MiNiFi C++ on Linux or on Windows.

Installing MiNiFi C++ on Linux

Learn how to install MiNiFi C++ on RHEL/CentOS, Ubuntu, Debian, and SLES.

About this task

To install MiNiFi C++ on RHEL/CentOS, Ubuntu, Debian, SLES, complete the following steps:

Procedure

1. Download the tar.gz or zip file for the MiNiFi C++ agent.

```
wget {cpp.tar.gz}
```

2. To install the MiNiFi C++ agent, extract the file to your desired home directory.

Installing MiNiFi C++ as a service on Linux

Learn how to install MiNiFi C++ as a service.

Procedure

1. Navigate to the MiNiFi C++ installation directory.

2. Enter:

```
bin/minifi.sh install
```

You can also specify a custom name for your MiNiFi C++ installation, by specifying that name during your install command. For example, to install MiNiFi C++ as a service and name it as dataflow, enter:

```
bin/minifi.sh install dataflow
```

Starting MiNiFi C++ on Linux

After you download and install MiNiFi C++ on Linux, you need to start MiNiFi C++.

About this task

You can start MiNiFi C++ in the foreground, background, or as a service on Linux.

Procedure

1. Launching MiNiFi C++ in the foreground: From a terminal window, navigate to the MiNiFi C++ installation directory.
2. Enter:

```
bin/minifi.sh run
```

3. Launching MiNiFi C++ in the background: From a terminal window, navigate to the MiNiFi C++ installation directory.
4. Enter:

```
bin/minifi.sh start
```

5. Launch MiNiFi C++ as a service: From a terminal window, enter:

```
sudo service minifi start
```

Installing MiNiFi C++ on Windows

Learn how to install MiNiFi C++ using a Windows MSI.

About this task

Before you begin your MiNiFi C++ installation, ensure that you meet the following requirements:

- Ensure the Domain user has administrator privilege.
- Ensure that your system meets the minimum memory requirement for Windows which is 4GB.

Procedure

1. Extract the MiNiFi C++ MSI files from the repo location https://archive.cloudera.com/web/CEM/windows/1.x/updates/<version>/nifi-minifi-cpp-<version>-<build-number>.msi to the location where you want to run the application.
2. Execute the MSI.

Configuring the MiNiFi C++ MSI

The MSI file adds the Windows service for MiNiFi C++. The service is configured to be executed by either a local user in the computer, or a domain user in Active Directory. Learn how to configure the MiNiFi C++ MSI.

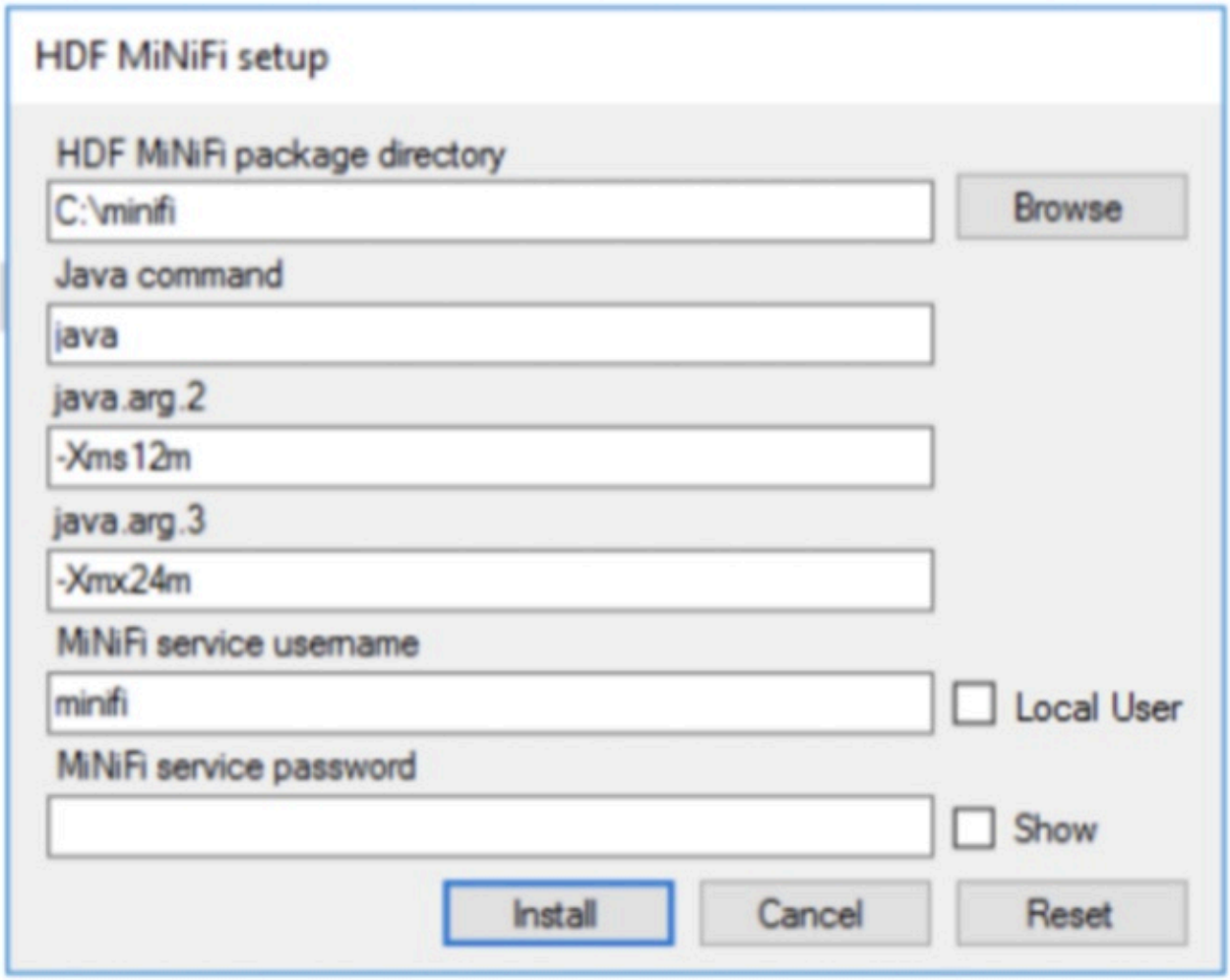
Using a Local User for MiNiFi C++ Windows Service

There is no prerequisite to use a local user for the Windows service. The installer automatically sets up the user.

The installer also grants the following privileges to the specified user:

- SeCreateSymbolicLinkPrivilege
- SeServiceLogonRight

If the computer is a part of a domain, then the Local User checkbox appears in the MiNiFi setup window. Check the Local User checkbox to specify that local user is used to execute the installed service.



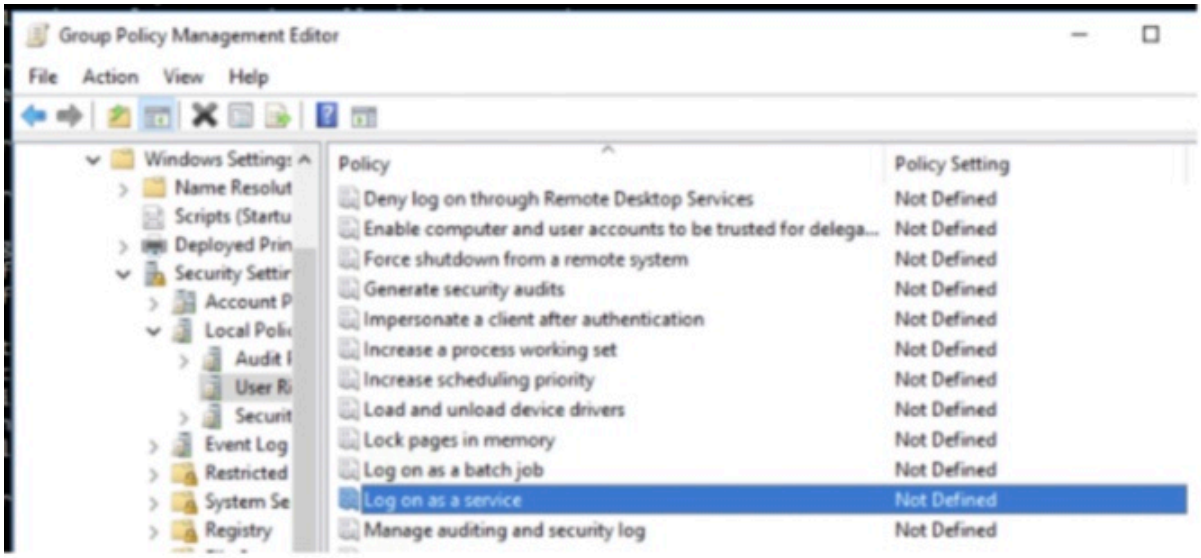
If a user specified in MiNiFi service username does not exist, the installer creates one with the specified MiNiFi service password. If the user already exists, the installer updates its password with the specified password.

Using a Domain User for MiNiFi C++ Windows Service

Before you begin

- The computer must be part of the domain.
- The specified user must exist in the domain, and a correct password must be provided.
- ActiveDirectory PowerShell module must be available.

1. In the Group Policy Management Editor, set permission to Log on as a service.



2. Navigate to a machine where MiNiFi C++ is installed and enter the following command:

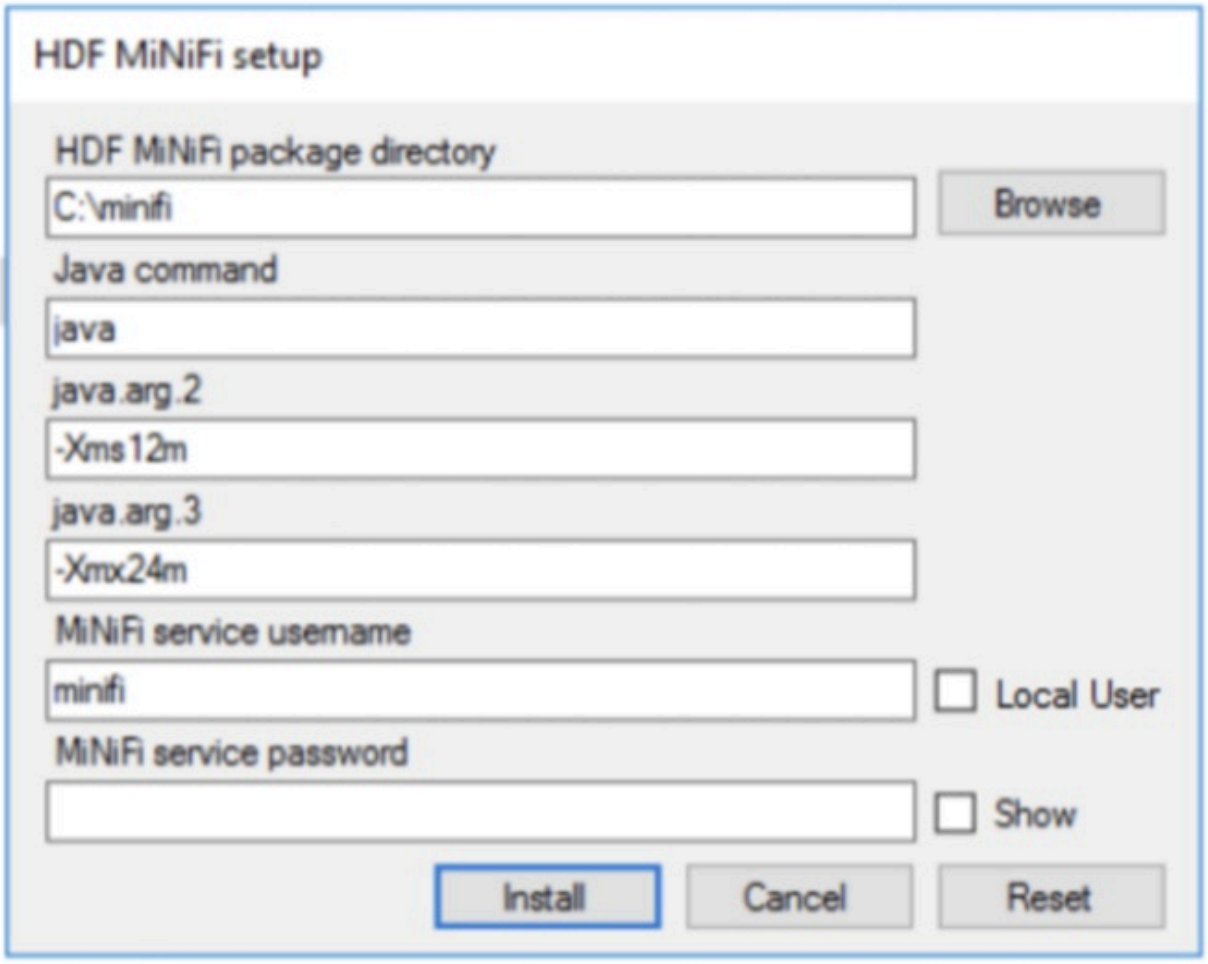
```
gpupdate
```

   The gpupdate command is a machine-wide command and can be executed from any directory on the MiNiFi machine.

3. Install the ActiveDirectory PowerShell module by entering the following in the PowerShell console:

```
Add-WindowsFeature RSAT-AD-PowerShell
```

4. In the MiNiFi setup window, uncheck the Local User checkbox then click Install.



Configuring C2 properties for MiNiFi C++

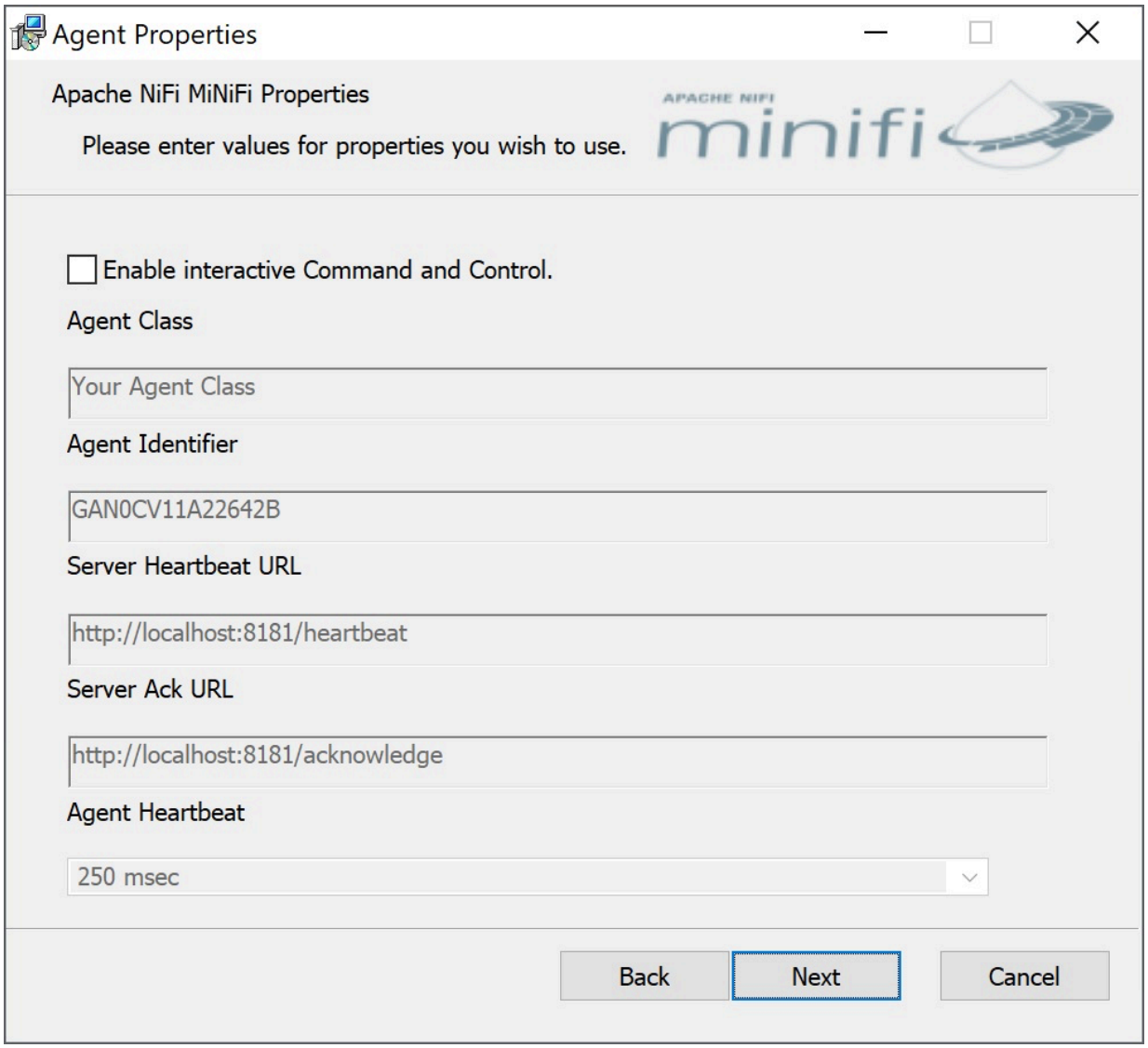
After you install the MiNiFi C++ agent, update the configuration file.

About this task

The configuration file for MiNiFi C++ is conf/minifi.properties.

Procedure

1. From the MiNiFi home directory, open the minifi.properties file.



2. Configure the Agent Class so that you can logically group MiNiFi instances according to their functionality. Specify the agent class:

```
nifi.c2.agent.class={AGENT_CLASS}
```

3. Configure the Agent ID. If you do not specify an Agent ID, MiNiFi generates a unique ID per agent instance.

```
nifi.c2.agent.identifier={AGENT_ID}
```

4. Configure your EFM Server endpoint:

```
nifi.c2.rest.url=http://{EFM_SERVER_IP}:10080/efm/api/c2-protocol/heartb
eat
nifi.c2.rest.url.ack=http://{EFM_SERVER_IP}:10080/efm/api/c2-protocol/ac
knowledge
```

5. Configure the heartbeat interval:

```
nifi.c2.agent.heartbeat.period={HEARTBEAT_INTERVAL}
```

6. Configure metrics:

```
nifi.c2.agent.protocol.class=RESTSender
```

Starting MiNiFi C++ on Windows

After you download and install MiNiFi C++, you can start MiNiFi C++ in the foreground or as a service on Windows.

About this task

Launching MiNiFi C++ in the foreground:

Procedure

1. From a command prompt window, navigate to the MiNiFi C++ installation directory.
2. Enter the following command to launch MiNiFi C++ in the foreground:

```
bin\run-minifi.bat
```

What to do next

Launching MiNiFi C++ as a service:

You can start or stop the installed MiNiFi C++ service from the Windows Service Manager.

Working with MiNiFi dataflows

When you are working with a MiNiFi dataflow, you should design it, add any additional configuration that your environment or use case requires, and then deploy your dataflow.

MiNiFi is not designed to accommodate substantial mid-dataflow configuration.

Setting up your MiNiFi dataflow

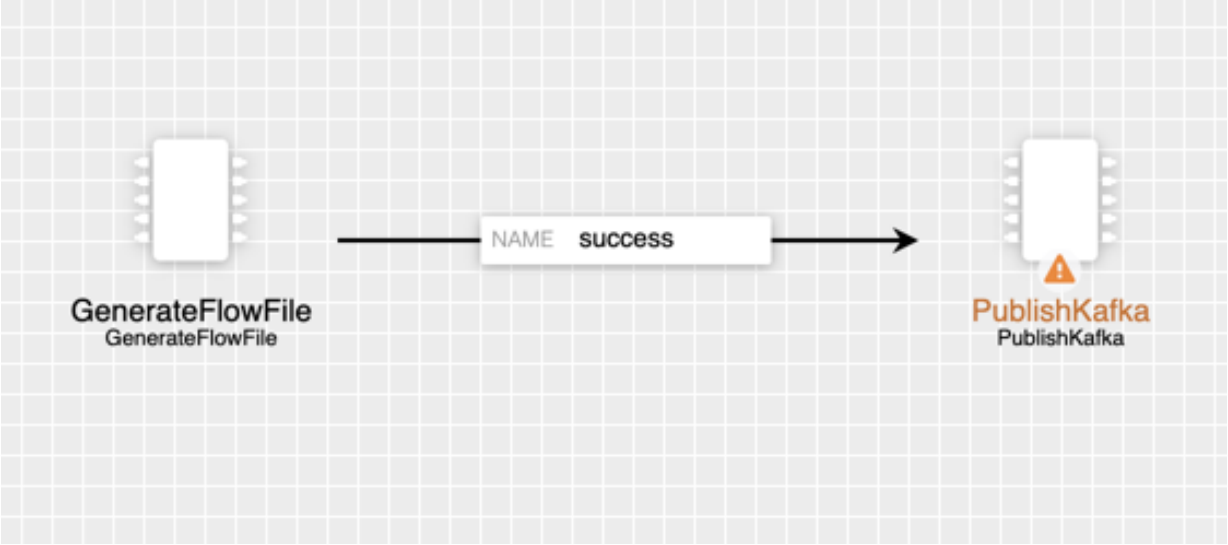
You can set up your MiNiFi dataflow using the Cloudera Edge Management (CEM) UI. To create a simple dataflow, you need to add processors, configure the processor properties, connect the processors, and publish the dataflow.

About this task

Procedure

1. Drag the GenerateFlowFile and PublishKafka processors to the canvas and connect them.

   The dataflow appears on the canvas as shown in the following image:



2. Configure the GenerateFlowFile processor and click APPLY to apply the changes.
3. Configure the PublishKafka processor and click APPLY to apply the changes.

   You need to set the following properties in the PublishKafka processor to enable SSL:

   - Security CA
   - Security Cert
   - Security Pass Phrase
   - Security Private Key
   - Security Protocol - Set it to ssl

   Apart from the above properties, you need to set the Known Brokers and Topic Name properties.

   Note: You must also enable certs for Kafka at the host where MiNiFi agent is running.

4. Publish the flow by selecting Publish from the ACTIONS drop-down.

Using processors not packaged with MiNiFi

You should be familiar with the processors that you can use out of the box.

You can find the supported processors by MiNiFi Java in *MiNiFi Java agent processor support*.

You can find the supported processors by MiNiFi C++ in *MiNiFi C++ agent processor support*.

MiNiFi is able to use the following processors out of the box:

- AttributesToJSON
- Base64EncodeContent
- CompressContent
- ControlRate
- ConvertCharacterSet
- ConvertJSONToSQL
- DetectDuplicate
- DistributeLoad
- DuplicateFlowFile

- EncryptContent
- EvaluateJsonPath
- EvaluateXPath
- EvaluateXQuery
- ExecuteProcess
- ExecuteSQL
- ExecuteStreamCommand
- ExtractText
- FetchDistributedMapCache
- FetchFile
- FetchSFTP
- GenerateFlowFile
- GetFTP
- GetFile
- GetHTTP
- GetJMSQueue
- GetJMSTopic
- GetSFTP
- HandleHttpRequest
- HandleHttpResponse
- HashAttribute
- HashContent
- IdentifyMimeType
- InvokeHTTP
- ListFile
- ListSFTP
- ListenHTTP
- ListenRELP
- ListenSyslog
- ListenTCP
- ListenUDP
- LogAttribute
- MergeContent
- ModifyBytes
- MonitorActivity
- ParseSyslog
- PostHTTP
- PutDistributedMapCache
- PutEmail
- PutFTP
- PutFile
- PutJMS
- PutSFTP
- PutSQL
- PutSyslog
- QueryDatabaseTable
- ReplaceText
- ReplaceTextWithMapping
- RouteOnAttribute
- RouteOnContent

- RouteText
- ScanAttribute
- ScanContent
- SegmentContent
- SplitContent
- SplitJson
- SplitText
- SplitXml
- TailFile
- TransformXml
- UnpackContent
- ValidateXml

Perform the following steps:

1. Set up your dataflow as described above.
2. Copy the desired NAR file into the MiNiFi lib directory.
3. Restart your MiNiFi instance.

Note: Currently only the StandardSSLContextService is supported as a controller service. It is created automatically if the Security Properties section is set and can be referenced in the processor configuration using the ID SSL-Context-Service.

Related Information

MiNiFi Java agent processor support
MiNiFi C++ agent processor support

Securing your dataflow

You can secure your MiNiFi dataflow using keystore or trust store SSL protocols, however, this information is not automatically generated. You need to generate your security configuration information yourself.

About this task

To run a MiNiFi dataflow securely, modify the Security Properties section in theconfig.yml file.

Procedure

1. Create your dataflow template as discussed in the *Setting up Your Dataflow* section.
2. Move it to minifi.conf and rename as config.yml.
3. Manually modify the Security Properties section of config.yml.

```
Security Properties:
keystore:
keystore type:
keystore password:
key password:
truststore:
truststore type:
truststore password:
ssl protocol: TLS
Sensitive Props:
key:
algorithm: PBEWITHMD5AND256BITAES-CBC-OPENSSL
provider: BC
```

Related Information

Setting up your MiNiFi dataflow

Managing MiNiFi

Apart from working with dataflows, you can also perform some management tasks using MiNiFi.

Monitoring status using MiNiFi

You can use the minifi.sh flowStatus option to monitor a range of aspects of your MiNiFi operational and dataflow status.

You can use the flowStatus option to get information dataflow component health and functionality, a MiNiFi instance, or system diagnostics.

FlowStatus accepts the following flags and options:

- Processors
  - health
  - bulletins
  - status
- Connections
  - health
  - stats
- remoteProcessGroups
  - health
  - bulletins
  - status
  - authorizationIssues
  - inputPorts
- controllerServices
  - health
  - bulletins
- provenancereporting
  - health
  - bulletins
- instance
  - health
  - bulletins
  - status
- Systemdiagnostics
  - heap
  - processorstats
  - contentrepositoryusage
  - flowfilerepositoryusage
  - garbagecollection

For example, this query gets the health, stats, and bulletins for the TailFile processors:

```
minifi.sh flowStatus processor:TailFile:health,stats,bulletins
```

Note: Currently the script only accepts one high level option at a time. Any connections, remote process groups, or processors names that contain ":", ";" or "," causes parsing errors when querying.

CEM can export time series metrics to several metric storage providers. The recommended metrics store service is Prometheus. Prometheus integrates with Grafana for time series metric visualization. With Prometheus and Grafana, you can store and visualize metrics for CEM. For details on how to enable and set Prometheus and Grafana, see *Monitoring Metrics with Grafana*.

Related Information

Monitoring metrics in CEM with Grafana

Loading a new dataflow for MiNiFi

You can load a new dataflow for a MiNiFi instance to run.

Procedure

1. Create a new config.yml file with the new dataflow.
2. Replace the existing config.yml in minifi/conf with the new file.
3. Restart MiNiFi.

Stopping MiNiFi

You can stop MiNiFi at any time.

Procedure

1. From a terminal window, navigate to the MiNiFi installation directory.
2. Enter:

```
bin/minifi.sh stop
```

3. To stop MiNiFi running as a service, from a terminal window, enter:

```
sudo service minifi stop
```