



Integrating Box with File Access Manager

Version: 8.4

Revised: March 27, 2023

Copyright and Trademark Notices

Copyright © 2023 SailPoint Technologies, Inc. All Rights Reserved.

All logos, text, content, including underlying HTML code, designs, and graphics used and/or depicted on these written materials or in this Internet website are protected under United States and international copyright and trademark laws and treaties, and may not be used or reproduced without the prior express written permission of SailPoint Technologies, Inc.

"SailPoint," "SailPoint & Design," "SailPoint Technologies & Design," "Identity Cube," "Identity IQ," "IdentityAI," "IdentityNow," "SailPoint Predictive Identity" and "SecurityIQ" are registered trademarks of SailPoint Technologies, Inc. None of the foregoing marks may be used without the prior express written permission of SailPoint Technologies, Inc. All other trademarks shown herein are owned by the respective companies or persons indicated.

SailPoint Technologies, Inc. makes no warranty of any kind with regard to this manual or the information included therein, including, but not limited to, the implied warranties of merchantability and fitness for a particular purpose. SailPoint Technologies shall not be liable for errors contained herein or direct, indirect, special, incidental or consequential damages in connection with the furnishing, performance, or use of this material.

Patents Notice. <https://www.sailpoint.com/patents>

Restricted Rights Legend. All rights are reserved. No part of this document may be published, distributed, reproduced, publicly displayed, used to create derivative works, or translated to another language, without the prior written consent of SailPoint Technologies. The information contained in this document is subject to change without notice.

Use, duplication or disclosure by the U.S. Government is subject to restrictions as set forth in subparagraph (c) (1) (ii) of the Rights in Technical Data and Computer Software clause at DFARS 252.227-7013 for DOD agencies, and subparagraphs (c)(1) and (c)(2) of the Commercial Computer Software Restricted Rights clause at FAR 52.227-19 for other agencies.

Regulatory/Export Compliance. The export and re-export of this software is controlled for export purposes by the U.S. Government. By accepting this software and/or documentation, licensee agrees to comply with all U.S. and foreign export laws and regulations as they relate to software and related documentation. Licensee will not export or re-export outside the United States software or documentation, whether directly or indirectly, to any Prohibited Party and will not cause, approve or otherwise intentionally facilitate others in so doing. A Prohibited Party includes: a party in a U.S. embargoed country or country the United States has named as a supporter of international terrorism; a party involved in proliferation; a party identified by the U.S. Government as a Denied Party; a party named on the U.S. Department of Commerce's Entity List in Supplement No. 4 to 15 C.F.R. § 744; a party prohibited from participation in export or re-export transactions by a U.S. Government General Order; a party listed by the U.S. Government's Office of Foreign Assets Control as ineligible to participate in transactions subject to U.S. jurisdiction; or any party that licensee knows or has reason to know has violated or plans to violate U.S. or foreign export laws or regulations. Licensee shall ensure that each of its software users complies with U.S. and foreign export laws and regulations as they relate to software and related documentation.

Contents

- Capabilities** **3**
- Box Connector Prerequisites** **4**
 - Box User Permissions 4
 - Communications Requirements 4
 - Software Requirements 4
- Box Connector Installation Flow Overview** **5**
- Collecting Data Stored in an External Application** **6**
 - Installation Locations 7
 - Box Connector Operation Principles 7
 - Permissions Collection Operation Principles 7
- Adding a Box Application** **9**
 - Select Wizard Type 9
 - General Details 9
 - Connection Details 10
 - Configuring and Scheduling the Permissions Collection 10
 - Selecting and Scheduling the Data Classification Settings 19
 - Data Privacy 20
 - Configuring Activity Monitoring 20
 - Configuring Data Enrichment Connectors 21
- Installing Services: Activity Monitor and Collectors** **23**
- Verifying the Box Connector Installation** **26**
 - Installed Services 26
 - Log Files 26
 - Monitored Activities 26
 - Permissions Collection 26

Capabilities

This connector enables you to use File Access Manager to access and analyze data stored in Box and do the following:

- Analyze the structure of your stored data.
- Monitor user activity in the resources.
- Classify the data being stored.
- Verify user permissions on the resources, and compare them against requirements.

See the File Access Manager documentation for a full description.

Box Connector Prerequisites

Box User Permissions

During the OAuth authorization process, a Box for Business Team Admin user must grant the SailPoint Box Application access to the data on Box.

If the "Disable unpublished apps by default" is set to True, the user will be unable to obtain the authentication code when configuring Box applications.

Communications Requirements

Requirement	Source	Destination	Port
File Access Manager Message Broker	Permissions Collector / Data Classification Collector	RabbitMQ	5671
File Access Manager Access	Activity Monitor / Permissions Collector	File Access Manager Servers	8000-8008
Permissions Collection / Data Classification	Permissions Collector / Data Classification	Box API	https
Activity Audit	Activity Monitor	Box API	https

Software Requirements

File Access Manager requires the latest ASP.NET Core 6.0.x Hosting Bundle. This bundle consists of .NET Runtime and ASP .NET Core Runtime. You can download the latest 6.0.x Hosting Bundle version from [here](#).

Box Connector Installation Flow Overview

To install the Box connector:

1. Configure all the prerequisites.
2. Add a new Box application in the File Access Manager website.
3. Install the relevant services:
 - Activity Monitor

Box currently does not support the Cloud-Ready architecture for permissions collection and data classification. Permission collection and data classification tasks will run on the central engine services associated with the application, regardless of whether these services have one or more collectors associated with the central engine.

Collecting Data Stored in an External Application

Terminology:

Connector

The collection of features, components and capabilities that comprise File Access Manager support for an endpoint.

Collector

The “Agent” component or service in a Data Classification and or Permission Collection architecture.

Engine

The core service counterpart of this architecture.

Identity Collector

A logical component used to fetch identities from an identity store and holds the configuration, settings for that identity store, and the relations between these identities.

The identity collector has no “physical” manifest.

- The actual work is done by the Collector Synchronizer.

The list below describes the high level installation process required to collect and analyze data from an external application. Most of these should already be set up in your File Access Manager installation. See the server Installation guide for further details.

Install a Data Classification central engine

One or more central engines, installed using the server installer

Install a Permission Collection central engine

One or more central engines, installed using the server installer

Create an Application in File Access Manager

From the Business Website. The application is linked to central engines listed above.

Add an Activity Monitor

To collect activities for this application - run the Collector Installation Manager and add an application under Activity Monitoring.

Installation Locations

Activity Monitor – installed remotely on a File Access Manager monitor application server, which can be a server joined to any domain, including a domain different from the monitored domain.

Box Connector Operation Principles

- File Access Manager Connector for Box uses the Box Content API for event monitoring, identity, and permissions collection.
- The Box Content API uses the OAuth 2.0 authorization protocol to authenticate and authorize API requests.
- SailPoint SecurityIQ for Box Connector is a registered Box App, which requires a short authorization process to use the Box API during the definition of the Box application.
- After the initial authorization process, File Access Manager handles the OAuth token management automatically and refreshes the token if needed.

Permissions Collection Operation Principles

- File Access Manager Box Permissions Collection task uses Box Content API to retrieve information from the Box application.
- File Access Manager creates a Box Identity Collector automatically at the end of the “Add New Application” wizard, which collects the Users and Groups from Box.

Users will only display in the Box Resource Tree if they are an owner of a resource.

- By default, permissions are analyzed on the folder level, but can also be analyzed on the file level. If the latter is the case, the system will only display uniquely managed files in the Business Resource Tree.

In contrast to other application types, to improve performance, Box permissions are also fetched from the target application during the Crawl task.

The permissions will only display in the client after the permission collection task has run, since they must be analyzed. If the crawler was unable to fetch the permissions, the permission collection task will fetch them.

Adding a Box Application

In order to integrate with Box, we must first create an application entry in File Access Manager. This entry includes the identification, connection details, and other parameters necessary to create the link.

To add an application, use the **New Application Wizard**.

1. Navigate to *Admin > Applications*
2. Click **Add New** to open the wizard.

Select Wizard Type

1. Click **Standard Application**
2. Click **Next** to open the **General Details** page.

General Details

Application Type

Box

Application Name

Logical name of the application

Description

Description of the application

Tags

Select tags for the application from the dropdown menu, and / or type a new name, and press **Enter** to create a new tag. The dropdown list of tags filters out matching tags as you type and displays up to 50 tags.

The **tags** replace the **Logical container** field that was used when creating applications in releases before 8.2

Event Manager Server

This option is available if there are more than one event manager servers configured in the system.

Select an event manager from the drop down menu.

Click **Next**. to open the Connection Details page.

Connection Details

Management Account ID

The account ID of the AWS management account - This is required for collecting user details and permissions from different accounts.

Use Dedicated IAM User

Use this to select the login method. Leave unchecked to use the recommended method of EC2 login.

Check this box to use a dedicated IAM use account for login.

If selecting a Dedicated IAM User method, fill in the following fields:

- ***Access Key Id***

The IAM user programmatic username of the File Access Manager user that was created in the pre-requisites.

- ***Secret Access Key***

The IAM user programmatic password.

Click **Next**.

Configuring and Scheduling the Permissions Collection

Permissions can be analyzed to determine the application permissions of an out-of-the-box application, provided you have defined an identity store for File Access Manager to use in its analysis, and you have run a crawl for the application.


The permission collector is a software component responsible for analyzing the permissions in an application.

The Central Permission Collector Service is responsible for running the Permission Collector and Crawler tasks.

If the “File Access Manager Central Permission Collector” wasn’t installed during the installation of the server, this configuration setting will be disabled.

Users should always run the crawl task before running a permission collection task.

To configure the Permission Collection

- Open the edit screen of the required application.
 - a. Navigate to **Admin > Applications**.
 - b. Scroll through the list, or use the filter to find the application.
 - c. Click the edit icon  on the line of the application.
- Press **Next** till you reach the **Crawler & Permissions Collection** settings page.

The actual entry fields vary according to the application type.

When entering this page in edit mode, you can navigate between the various configuration windows using the **Next** and **Back** buttons.

Central Permissions Collection Service

Select a central permission collection service from the dropdown list. You can create permissions collection services as part of the service installation process. See section "Services Configuration" in the File Access Manager Administrator Guide for further details.

Analyze all Objects in S3 Bucket

If checked – collect and analyze files in the buckets, and not only buckets and folders.

Default is unchecked.

Analyze ACL Permissions

Click to fetch and analyze ACL-type Permissions..

If checked, ACLs will be collected for business resources, which will impact the performance of the Permission Collector. For cases with a large number of resources, skipping the ACL permission fetch can improve the service run time considerably.

This option is checked by default.

If ACL is not supported by your server, make sure this field is unchecked.

Scheduling a Task

Create a Schedule

Click on this option to view the schedule setting parameters.

Schedule Task Name

A name for this scheduling task

When creating a new schedule, the system generates a default name in the following format:

{appName} - {type} Scheduler

You can override or keep this name suggestion.

Schedule

Select a scheduling frequency from the dropdown menu.

- **Schedule Types and Intervals**

Once

Single execution task runs.

Run After

Create dependency of tasks. The task starts running only upon successful completion of the first task.

Hourly

Set the start time.

Daily

Set the start date and time.

Weekly

Set the day(s) of the week on which to run.

Monthly

The start date defines the day of the month on which to run a task.

Quarterly

A monthly schedule with an interval of 3 months.

Half Yearly

A monthly schedule with an interval of 6 months.

Yearly

A monthly schedule with an interval of 12 months.

Date and time fields

Fill in the scheduling times. These fields differ, depending upon the scheduling frequency selected.


Active check box

Check this to activate the schedule.

Click **Next**.

Configuring and Scheduling the Crawler

To set or edit the Crawler configuration and scheduling

- Open the edit screen of the required application.
 - a. Navigate to **Admin > Applications**.
 - b. Scroll through the list, or use the filter to find the application.
 - c. Click the edit icon  on the line of the application.
- Press **Next** till you reach the **Crawler & Permissions Collection** settings page.

The actual entry fields vary according to the application type.

Calculate Resource Size

Determine when, or at what frequency, File Access Manager calculates the resources' size.

Select one of the following:

- Never
- Always
- Second crawl and on (This is the default)

Exclude CloudTrail Logs

Check this box to exclude CloudTrail logs from being crawled and analyzed. There could be a very large number of these log files, and scanning them will have a negative impact on performance.

The default is checked.

Create a Schedule

Click to open the schedule panel.


Setting the Crawl Scope

There are several options to set the crawl scope:

- Setting explicit list of resources to include and / or exclude from the scan.
- Creating a regex to define resources to exclude.

Including and Excluding Paths by List

To set the paths to include or exclude in the crawl process for an application

- Open the edit screen of the required application.
 - a. Navigate to **Admin > Applications**.
 - b. Scroll through the list, or use the filter to find the application.
 - c. Click the edit icon  on the line of the application.
- Press **Next** till you reach the **Crawler & Permissions Collection** settings page.

The actual entry fields vary according to the application type.

1. Scroll down to the Crawl configuration settings.
2. Click **Advanced Crawl Scope Configuration** to open the scope configuration panel.
3. Click Include / Exclude Resources to open the input fields.
4. To add a resource to a list, type in the full path to include / exclude in the top field and click **+** to add it to the list.
5. To remove a resource from a list, find the resource from the list, and click the x icon on the resource row.

When creating exclusion lists, excludes take precedence over includes.

Excluding Paths by Regex for AWS S3 Buckets

The AWS Path Structure in File Access Manager

File Access Manager uses a path name in the following structure:

Path Structure: Root/[OU]/[Account]/[Bucket Path]/[Folder]/[Filename]

Component structure: Root/[OU]/[OU2]/[Account name](#[Account ID])/s3.[region].[bucket name]/[folder]/[file name]

Example: Root/Example-OU/Example-Account(#420269343516)/s3.north-east-17.HR3InputDataBucket/Prospects/CVs/SueSmithPM.Docx

Root

All paths start with "Root"

OU

The organizational unit. This could be empty, or include a string of one or more OUs, according to the BR hierarchical structure.

Account

Since account names are not unique under an organization, this string includes the account ID and the account name


[Account name] ([Account ID])

Bucket Path

The bucket section of the path starts with "s3." and includes the region

s3.[region].[bucket]

To set filters of paths to exclude in the crawl process for an application using regex.

- Open the edit screen of the required application.
 - a. Navigate to **Admin > Applications**.
 - b. Scroll through the list, or use the filter to find the application.
 - c. Click the edit icon  on the line of the application.

- Press **Next** till you reach the **Crawler & Permissions Collection** settings page.

The actual entry fields vary according to the application type.

1. Click **Exclude Paths by Regex** to open the configuration panel.
2. Type in the paths to exclude by Regex, See regex examples in the section below. Since the system does not collect BRs that match this Regex, it also does not analyze them for permissions.

Crawler Regex Exclusion Examples - AWS S3 Buckets

The following are examples of crawler Regex exclusions:

Exclude all drives which start with one or more user names:

Starting with John.Doe

Regex: ^Users\John\.Doe@.*

Starting with John.Doe or Jane.Doe

Regex: ^Users\(John|Jane\)\.Doe@.*

Personal/admin3@501.sailpointtechnologies.com

Include ONLY drives which start with one or more user names:

Starting with John.Doe

Regex: ^(?!Users\John\.Doe@.*)*

Starting with John.Doe or Jane.Doe

Regex: ^(?!Users\(John|Jane\)\.Doe@.*)*

Narrow down the selection:

Include ONLY the C\$ drive shares: \\server_name\C\$

Regex: ^(?!\\\\server_name\\C\\$(\\$|\\\.*)).*

Include ONLY one folder under a share: \\server\share\folderA

Regex: ^(?!\\\\server_name\\share\\$(\\$|\\folderA\$|\\folderA\\\.*)).*

Include ONLY all administrative shares

Regex: `^(?!\\server_name\[a-zA-Z]\$($|)).*`

To write a backslash or a Dollar sign, add a backslash before it as an escape character.

To add a condition in a single command, use a pipe character “|” .

Excluding Top Level Resources

Use the top level exclusion screen to select top level roots to exclude from the crawl. This setting is done per application.

To exclude top level resources from the crawl process

1. Open the application screen

Admin > Applications

2. Find the application to configure and click the drop down menu on the application line. Select **Exclude Top Level Resources** to open the configuration panel.

3. **Run Task**

The Run Task button triggers a task that runs a short detection scan to detect the current top level resources.

Before running the task for the first time, the message above this button is:

"Note: Run task to detect the top-level resources"

If the top level resource list has changed in the application while you are on this screen, press this button to retrieve the updated structure.

Once triggered, you can see the task status in

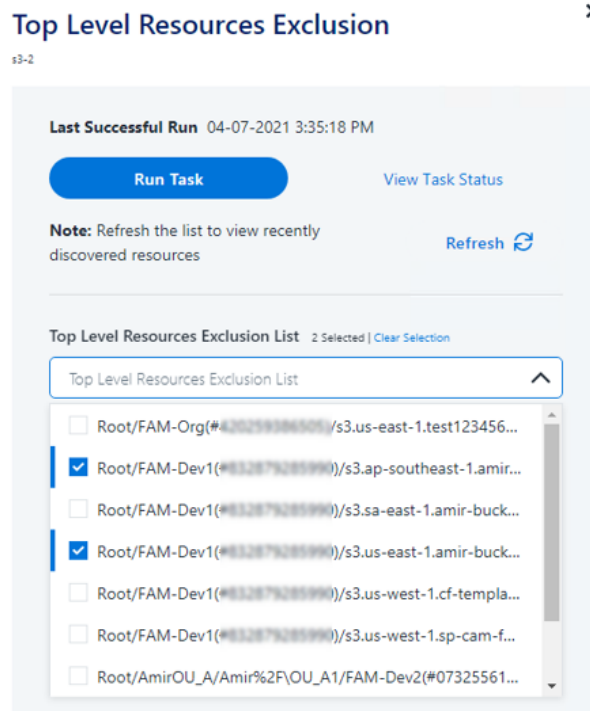
Settings > Task Management > Tasks

This will only work if the user has access to the task page

When the task has completed, press **Refresh** to update the page with the list of top level resources.

4. Click the top level resource list, and select top level resources to exclude.
5. Click **Save** to save the change.

- To refresh the list of top level resources, run the task again. Running the task will not clear the list of top level



resources to exclude.

Special Consideration for Long File Paths in Crawl

If you need to support long file paths above 4,000 characters for the crawl, set the flag

excludeVeryLongResourcePaths

in the Permission Collection Engine App.config file to true.

By default this value will be commented out and set to false.

This key ensures, when enabled, that paths longer than 4000 characters are excluded from the applications' resource discovery (Crawl), to avoid issues while storing them in the SQLServer database.

When enabled, business resources with full paths longer than 4000 characters, and everything included in the hierarchical structure below them, will be excluded from the crawl, and will not be collected by File Access Manager. This scenario is extremely rare.

You should not enable exclusion of long paths, unless you experience an issue.

Background

File Access Manager uses a hashing mechanism to create a unique identifier for each business resource stored in the File Access Manager database. The hashing mechanism in SQLServer versions 2014 and earlier, is unable to process (hash) values with 4,000 or more characters.

Though resources with paths of 4000 characters or longer are extremely rare, File Access Manager is designed to handle that limitation.

Identifying the Problem

When using an SQL Server database version 2014 and earlier

The following error message in the Permission Collection Engine log file:

```
System.Data.SqlClient.SqlException (0x80131904): String or binary data would be truncated.
```

In all other cases, this feature should not be enabled.

Setting the Long Resource Path Key


The Permission Collection Engine App.config file is `RoleAnalyticsServiceHost.exe.config`, and can be found in the folder

`%SailPoint_Home%\FileAccessManager\[Permission Collection instance]\`

Search for the key **`excludeVeryLongResourcePaths`** and correct it as described above.

Selecting and Scheduling the Data Classification Settings

To associate an application with a data classification service, and set the schedule

- Open the edit screen of the required application
 - a. Navigate to **Admin > Applications**
 - b. Scroll through the list, or use the filter to find the application
 - c. Click the edit icon  on the line of the application

- Press **Next** till you reach the **Data Classification** settings page.

The actual entry fields vary according to the application type

Central Data Classification Service

Associate the application with a Central Data Classification Service. This service is responsible for running the Data Classification tasks.

If the “Central Data Classification” wasn’t installed during the installation of the server, this field is disabled.

Disabling Data Classification

To disable data classification, delete the entry from the central data classification field.

Disabling data classification can also be achieved by setting the scheduler to be inactive (which is the default setting for data classification).

Create a Schedule

This option is enabled only if a central data classification service is selected.

See [Configuring and Scheduling the Permissions Collection](#)

See the chapter “Data Classification” in the File Access Manager Administrator Guide for more information

Click **Next** or **Finish**.

Data Privacy

A user can associate the application with a Central Data Classification Engine Service. This engine will be responsible for executed Data Privacy tasks.

Though using different processes for each, the Data Classification engine service is in charge for both Data Privacy and Data Classification discovery tasks.

You may choose the same service for both, or use a different one for each, to run them in parallel.

The fields on the Data Privacy step are the same as the Data Classification step.

Configuring Activity Monitoring

Configure the activity monitoring process frequency.

Polling Interval (sec)

Activity fetching interval [in seconds]. Default is set to 60 seconds,

Report Interval (sec)

Activity Monitor Health reporting interval [in seconds]. Default is set to 60 seconds.

Local Buffer Size (MB)

Local buffer size for activities [in MB]. Default is set to 200MB.

This cyclic buffer is used to store activities on the Application Monitor's machine in case of network errors that prevent the activities from being sent.

Activity Data Retention Period

By default, this feature is disabled.

When selecting the Clear Activity Data option, a user is able to provide a time frame (1 to 100) in either months or years for all activity to be retained. Once that time period is met, all data will be removed.

A user can also select to backup the data before it is deleted by selecting the Backup Events Before Clearing option.

The Backup Before Clearing Option will only be enabled if the backup option is set during the system installation. If a user has not selected the backup option during the installation nor provided a backup path, this option will not be enabled.

Activity Data Retention Period

Activity data will be retained for the specified period. Following that time period, activities will be cleared.

Clear Activity Data

How long do you want to keep activity data? *

12 Month(s)

Check this option to backup activity data before it is cleared.

Backup Events Before Clearing

Configuring Data Enrichment Connectors

The Data Enrichment Connectors (DEC) configuration enables us to select data enrichment sources. These can be used to add information from other sources about identities.

An enrichment source could be a local HR database that is used to combine users' job descriptions or departments to the information stored in the identity store.

Select the data enrichment connectors to enrich monitored activities from the Available DECs text box.

Use the > or >> arrows to move the selected DECs to the Current DECs text box.

The user can select multiple DECs. Simply select each desired DEC.

You can create a new DEC in the Administrative Client (Applications>Configuration>ActivityMonitoring>DataEnrichmentConnectors).

After creating a new DEC, click **Refresh** to refresh the dropdown list.

The chapter Connectors of the File Access Manager Administrator Guide provides more information on Data Enrichment Connectors, including what they are, how to configure them, and how they fit in the Activity Flow.

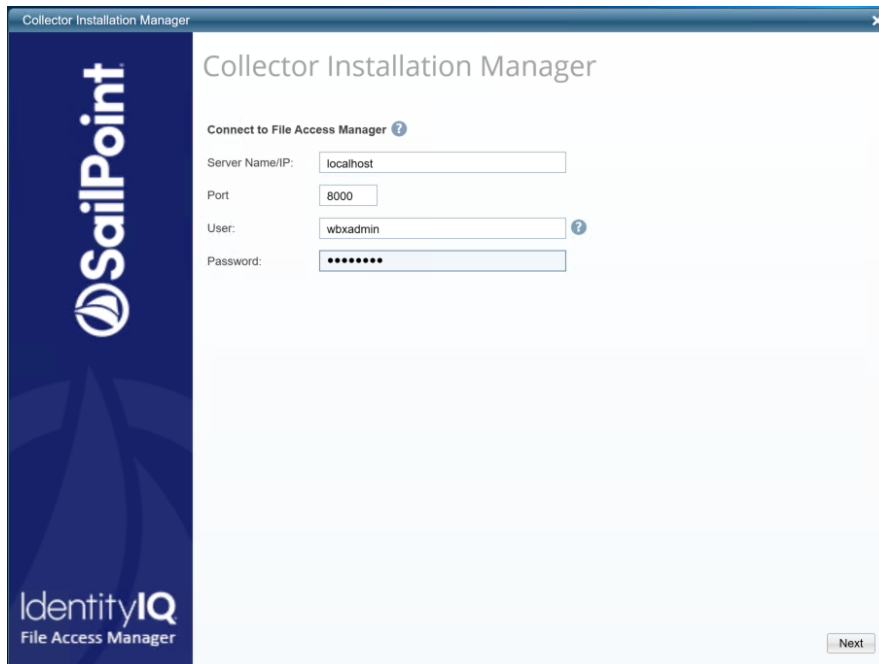
Installing Services: Activity Monitor and Collectors

The Collector Installation Manager is part of the File Access Manager installation package. This tool is used to install the activity monitor, permission collector, and data classification collector.

1. Run the **Collector Installation Manager** as an Administrator.

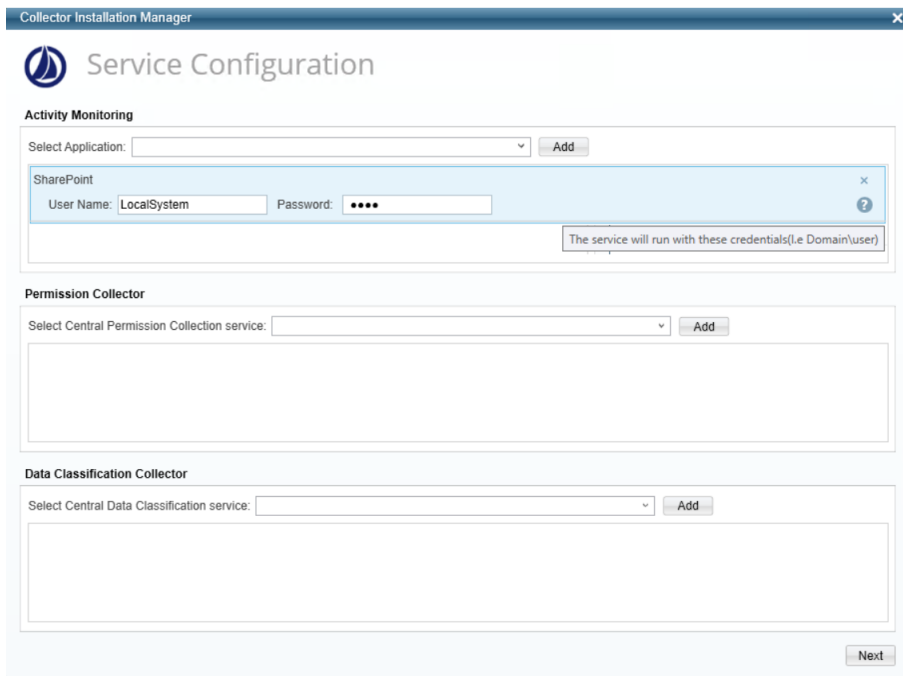
The installation files are in the installation package under the folder Collectors.

The Collector Installation Manager window displays.



2. Enter the credentials to connect to File Access Manager.
 - a. ServerName/IP should be pointed to the Agent Configuration Manager service server.
 - b. An File Access Manager user with Collector Manager permission (permission to install collectors). For Active Directory authentication, use the format domain\username.
3. Click **Next**.

The Service Configuration window displays.



4. If you are installing the Activity Monitor, select the application, and click **Add**.
5. When installing a SharePoint Activity Monitor, you will be prompted for service account credentials. This service account will be used by the Activity Monitor service to run the service and authenticate against the SharePoint IIS servers to fetch the logs (“Log on as”). Make sure the service account provided has local administrator privileges on the local server (hosting the Activity Monitor service) and can access the activity logs on the IIS servers.
6. If you are installing the Permission Collector, select the Central Permission Collector to which to connect this service, and click **Add**.
7. If you are installing the Data Classification, select the Central Classification Collector to which to connect this service, and click **Add**.
8. Click **Next**.

The Installation Folder window displays.

If this is the first time you are installing collectors on this machine, you will be prompted to select an installation folder, in which all future collectors will also be installed.

9. Browse and select the location of the target folder for installation.

10. Browse and select the location of the folder for system logs.
11. Click **Next**.
12. The system begins installing the selected components.
13. Click **Finish**.

The Finish button is displayed after all the selected components have been installed.

The File Access Manager Administrator Guide provides more information on the collector services.

Verifying the Box Connector Installation

Installed Services

Verify that the services installed for the connector are available and active. Using windows Service manager, or other tool, look for the File Access Manager services, and see that they are running.

for example:

- File Access Manager Central Activity Monitor - <Application_Name>
- File Access Manager Central Permissions Collection - <Application_Name>

Log Files

Check the log files listed below for errors

- "%SAILPOINT_HOME_LOGS%\PermissionCollection_<Service_Name>.log"
- "%SAILPOINT_HOME_LOGS%\DataClassification_<Service_Name>.log"
- "%SAILPOINT_HOME_LOGS%\BOX-<Application_Name>.log"

Monitored Activities

1. Simulate activities on Box.
2. Wait a minute (approximately).
3. Verify that the activities display in the File Access Manager website under
Forensics > Activities

Permissions Collection

1. Run the Crawler and Permissions Collector tasks (*Settings > Task Management > Scheduled Tasks*)

2. Verify that:

- The tasks completed successfully
- Business resources were created in the resource explorer (*Admin > Applications > [application column] > Manage Resources*)
- Permissions display in the Permission Forensics page (*Forensics > Permissions*)