Integrating DFS with File Access Manager

Copyright and Trademark Notices

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

Contents

Connector Overview

Connector Operation Principles

- File Access Manager Windows DFS application differs from other File Access Manager connectors in that it does not actively monitor activities, collect permissions, or classify data. Instead, it acts as a logical representation of multiple physical applications. It fetches data by mapping DFS logical shares to their corresponding physical target applications shares.

- The crawler service creates mapping between DFS applications and physical applications.

> Windows DFS applications only supports domain-based DFS namespaces.

> An application must be configured in File Access Manager for each DFS domain.

Terminology

DFS (Distributed File System) refers to a virtual arrangement of distributed Microsoft servers as a single resources tree.

For a full DFS Namespace overview, see

https://docs.microsoft.com/en-us/windows-server/storage/dfs-namespaces/dfs-overview.

DFS Namespace

A virtual view of shared folders on servers provided by DFS. A DFS namespace consists of a root and many links and targets. The namespace starts with a root that maps to one or more root targets. Below the root are links that map to their own targets.

Domain-based DFS namespace

A DFS namespace whose configuration information is stored in Active Directory.

DFS Link (folder with targets)

A component in a DFS path that lies below the root and maps to one or more link targets.

DFS Link Target (folder target)

The mapping destination of a link. A link target can be any UNC path, such as a shared folder or another DFS path.

Monitored Activities

Any activities on shares that are targets of DFS links are "tagged" with an additional field with the logical DFS path and an indication that they are DFS-related. This allows activities to be queried via a DFS application and business resources and to have its DFS logical path displayed.

Any activity type monitored by a physical application, mapped to the DFS application, can also be displayed via the DFS application.

Permissions Collection and Data Classification

DFS are logical resources that only point to the physical folders in which actual data are located. Windows DFS applications do not have Permissions Collection nor Data Classification services.

To display permissions and data classification results, DFS applications redirect their link folders to their mapped targets and display the results collected by their physical applications.

DFS Link Targets Priority

Some DFS links may point to multiple physical shares that are assumed to be replicated. If so, File Access Manager selects prioritized results from one or more physical shares, which is different for each of the following scenarios:

Activities

When a link has multiple targets, all physical target resources are queried for activities, since activities are not necessarily replicated consistently across shares.

Permissions

When a link has multiple targets, the target with the most recent File Access Manager permission analysis is selected.

Data Classification

When a link has multiple targets, the target with the most recent File Access Manager Data Classification analysis is selected.

Manual Matching of Unknown Target Host Names

During a DFS Crawl, the Crawler tries to match target host names to the host names of existing applications in the File Access Manager database.

When the Crawler is unable to match specific hosts, it attempts to match hosts via DNS lookups, and to find valid matching alias names (for example, a host name displayed as an IP address).

If a search cannot find host names or cannot match host name aliases to an existing host in the File Access Manager database, it is possible to configure matching hosts manually.

1. Create an *.xml file with the following structure:

```xml
<?xml version="1.0"?>

<mappings>

<key name="hostA">AlternateHostA</key>

<key name="hostB">172.66.12.12</key>

</mappings>
```

   In the example above, "hostA" is a host name of a link target to be matched manually. "AlternateHostA" is the host name to which "hostA" will be matched.

   > "AlternateHostA" should be a host name of an existing application in File Access Manager.

2. Add the following key to the DFS Permissions Collector's service "app.config".

```
"<add key="dfsMappings" value="C:\myMappings.xml"/>"
```

3. Replace "C:\myMappings.xml" with the path that points to the configuration file.

4. Restart the DFS Permissions Collector service.

Capabilities

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

- Analyze the structure of your stored data.

- Monitor user activity in the resources.

- Manage access fulfillment - automated granting and revoking of access - according to rules set in File Access Manager.

See the File Access Manager documentation for a full description.

Prerequisites

Make sure your system fits the descriptions below before starting the installation.

Communications Requirements

Requirement	Source	Destination	Port
Database Access	Permissions Collector	File Access Manager database	Per the specific DB definitions
LDAP for authenticating			389
RPC			135 + Dynamic ports range
Required for collecting DFS details			139, 445
DNS			53 (UDP and TCP)

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 .

Permissions

When the Crawler service runs on a server, which is part of the domain that hosts the domain-based DFS namespaces, neither a user nor a password configuration is required.

If the server is not part of the DFS domain, then the following requirements must be met:

- The server must be able to resolve the DFS domain.

- The Crawler service must be configured with a standard domain user on that DFS domain in order to gain access using impersonation.

DFS Installation Flow Overview

To install the DFS connector:

1. Configure all the prerequisites.

2. Add a new DFS application in the Business Website.

3. Install the relevant services:

   - Activity Monitor - This is the activity collection engine, used by all connectors that support activity monitoring.

Collecting Data Stored in an External Application

Connector / Collector terminology:

Connector

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

Collector

The "Agent" component or service in a 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 It 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 Permission Collection central engine

One or more central engines, installed using the server installer

Create an Application in File Access Manager

In the File Access Manager website, (*Admin > Applications*). The application is linked to central engines listed above.

Crawler Responsibilities

- The crawler works with native API calls that communicate with the domain controller and the DFS namespace servers.

- The crawler:

    - Creates the DFS resources tree

    - Creates mapping between the DFS links resources and physical applications resources.

    - The crawler can only map DFS links to physical shares in the File Access Manager database. Therefore, before the DFS crawler runs, the shares in the physical applications must have already been found. If a physical share is not found for a DFS link, the crawler will issue an "unfound target" warning in the task details.

Adding a DFS Application

In order to integrate with DFS, 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

> Windows DFS

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

Complete the Connection Details fields:

Domain Name

> The user defined in the prerequisites

Username

> The user defined in the prerequisites

Password

> The user defined in the prerequisites

The Windows DFS application supports only crawling, therefore, the scheduling tab will only contain the crawler scheduling window.

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.

Create a Schedule

> Click to open the schedule panel. See Scheduling a Task

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 [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

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 Example

The following are examples of crawler Regex exclusions:

Exclude all shares which start with one or more shares names:

> Example: Starting with \\[domain_name]\[namespace]\shareName
>
> Regex: ^\\\\domain_name\\namespace\\shareName$
>
> Real Example: Exclude myShare share
>
> Path: \\myDomain\myNamespace\myShare
>
> Regex: ^\\\\myDomain\\myNamespace\\myShare$

> Example: Starting with \\[domain_name]\[namespace]\shareName or \\[domain_name]\[namespace]\OtherShareName
>
> Regex: ^\\\\domain_name\\namespace\\(shareName|OtherShareName)$

Include ONLY shares which start with one or more shares names:

> Example: Starting with \\[domain_name]\[namespace]\shareName
>
> Regex: ^(?!\\\\domain_name($|\\namespace($|\\shareName($|\\.*)))).*

> Real Examples: Include only myShare share
>
> Path: ^(?!\\\\myDomain($|\\myNamespace($|\\myShare($|\\.*)))).*
>
> Regex: ^\\\\myDomain\\myNamespace\\myShare$

> Starting with \\server_name*shareName* or \\server_name*OtherShareName*
>
> `Regex:` `^(?!\\\\\server_name\\(`*shareName*`|`*OtherShareName*`)($|\\.*)).*`

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 yo u 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.

6. 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 ealier

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.

Installing Services: Collector Installation

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.
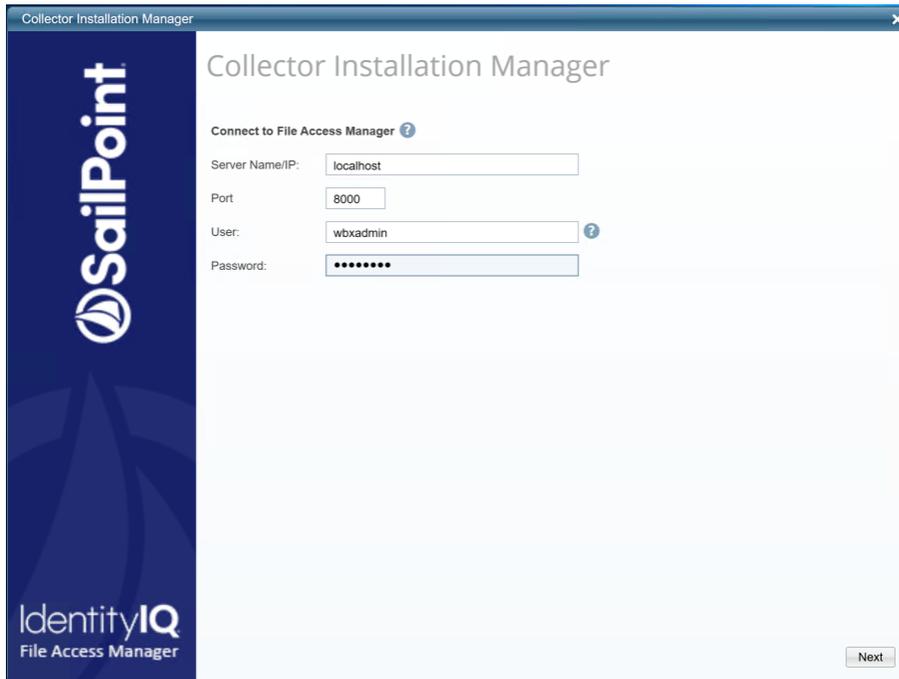
2. Enter the credentials to connect to File Access Manager.

   a. ServerName/IP should be pointed to the Agent Configuration Manager service server.

   b. A File Access Manager user with Collector Manager permission (permission to install collectors). For Active Directory authentication, use the format domain\username.

3. Click **Next** to open the Service Configuration window.
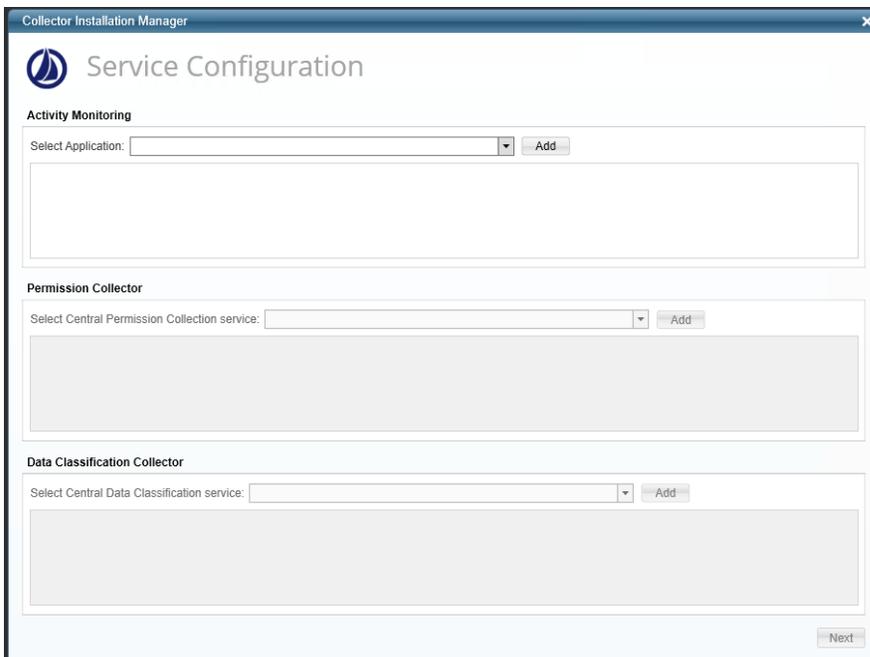
4. If you are installing the Permission Collector, select the Central Permission Collector to which to connect this service. Click **Add**.

5. If you are installing the Data Classification, select the Central Classification Collector to which to connect this service. Click **Add**.

6. 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. All future collectors will be installed in this folder.

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

8. Browse and select the location of the folder for system logs.

9. Click **Next**.

10. The system begins installing the selected components.

11. 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 DFS Connector Installation

Crawler

1. Run the Crawler task (*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*).

   - DFS is mapped to physical resources.

Monitored Activities

1. Assure that activities are received for the physical application, mapped to the DFS share.

2. Run the Crawler task, and verify that the DFS resource explorer has the DFS share and folder to be used for the activity simulations.

3. Simulate activities on physical DFS shares.

4. Wait a for approximately one minute.

5. Query activities in the File Access Manager website by selecting the DFS application <Application_Name>.

6. Verify that the activities display in the web Client.

   > The activity details should contain a "Logical Paths" field with the DFS logical paths.

Permissions Collection

1. Run a Permissions Collector task on a physical application with DFS target shares.

2. Verify the following:

- The task completed successfully.

- Permissions display in the Permissions Forensics window for the physical DFS target shares.

- Permissions display in the Permissions Forensics window for the DFS links which points to the above physical target shares

Data Classification

1. Run a Data Classification task on a physical application with DFS target shares.

2. Verify that:

   - The task completed successfully.

   - Data Classification results display in the File Access Manager website Data Classification Forensics window for the physical DFS target shares.

   - Data Classification results display in the File Access Manager website Data Classification Forensics window for the DFS links that point to the above physical target shares.

Troubleshooting

Check the issues below for common problems and suggested ways of handling them.

Manual Matching of Unknown Target Host Names

During a DFS Crawl, the Crawler tries to match target host names to the host names of existing applications in the File Access Manager database.

When the Crawler is unable to match specific hosts, it attempts to match hosts via DNS lookups, and to find valid matching alias names (for example, a host name displayed as an IP address).

If a search cannot find host names or cannot match host name aliases to an existing host in the File Access Manager database, it is possible to configure matching hosts manually.

To manually configure matching hosts, perform the following steps:

1. Create an *.xml file with the following structure:

   ```
   <?xml version="1.0"?>

   <mappings>

   <key name="hostA">AlternateHostA</key>

   <key name="hostB">172.66.12.12</key>

   </mappings>
   ```

   In the above example, "hostA" is a host name of a link target to be matched manually. "AlternateHostA" is the host name to which "hostA" will be matched.

   > "AlternateHostA" should be a host name of an existing application in File Access Manager.

2. Add the following key to the DFS Permissions Collector's service "app.config".

   ```
   "<add key="dfsMappings" value="C:\myMappings.xml"/>"
   ```

3. Replace "C:\myMappings.xml" with the path that points to the configuration file.

4. Restart the DFS Permissions Collector service.

Activities not Displayed in the Website

If activities are not shown in the Business Website:

- Verify that all prerequisites were set.

- Check if the shares on the physical applications have activities.

- If the physical shares do not have activities, this indicates that the monitor on the physical application is not working properly. Please refer to the relevant physical application Connector Troubleshooting guide.

If the shares on the physical applications have activities, but the DFS shares do not have activities, this

indicates that the Windows DFS crawler did not map the DFS shares to physical application shares properly.