AIP Scanner

The Azure Information Protection scanner is a tool for automatic labeling and classification of files and documents from on-premises file shares and SharePoint servers. Due to the requirements of labels with automatic processing, it requires an Azure Information Protection Premium P2 license.

The scanner runs as a service on Windows Server and lets you discover, classify, and protect files on the following data stores:

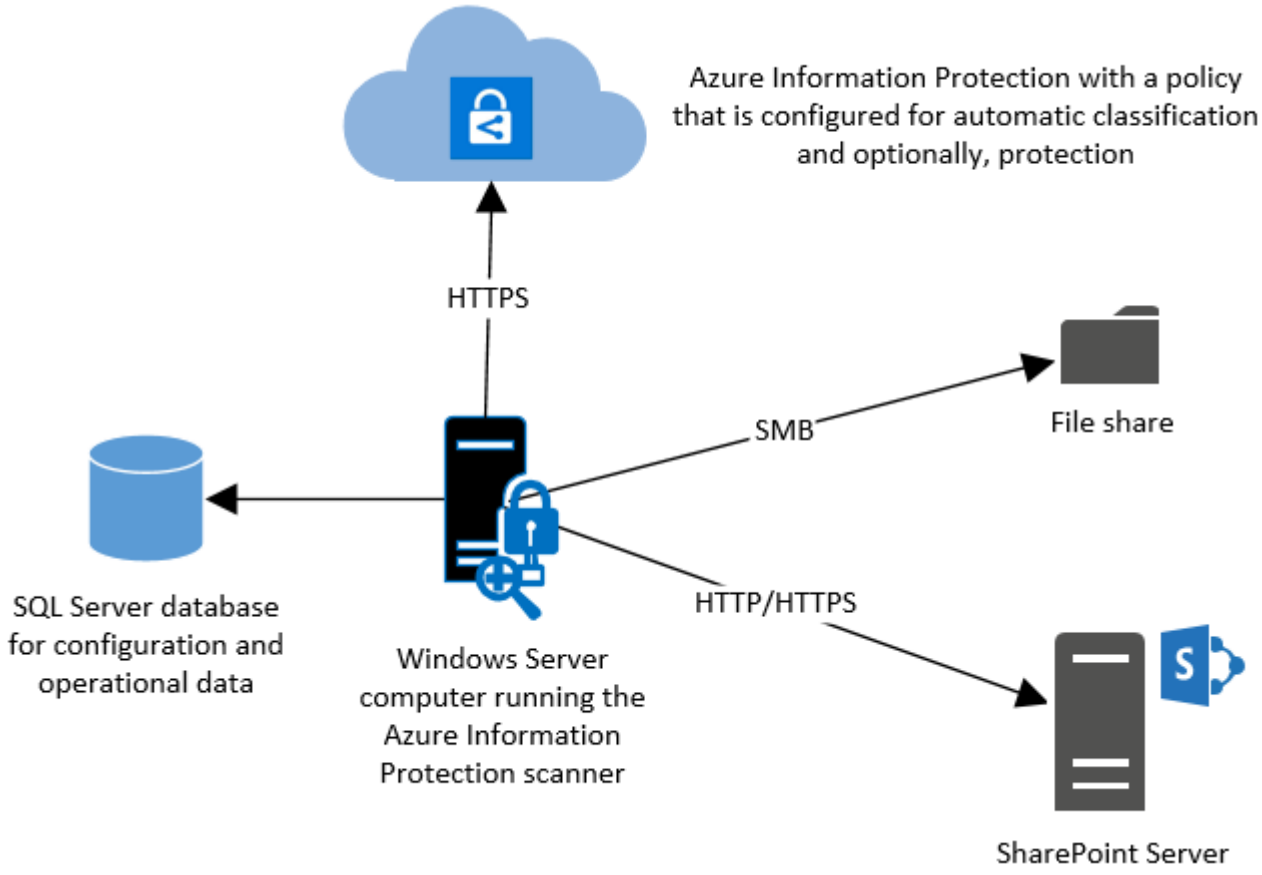
- Local folders on the Windows Server computer that runs the scanner.
- UNC paths for network shares that use the Server Message Block (SMB) or Common Internet File System (CIFS) protocol.
- Sites and libraries for SharePoint Server 2016 and SharePoint Server 2013.

To scan and label files on cloud repositories, use Cloud App Security.

The Azure Information Protection scanner can word as an extension to a Windows Server File Classification Infrastructure (FCI).

How does the Azure Information Protection scanner work?

The scanner is installed on a Windows Server with access to the on-premise environment containing the file shares and SharePoint libraries to label. It can inspect any files that Windows can index by using iFilters that are installed on the computer. Then, to determine if the files need labeling, the scanner uses the Microsoft 365 built-in data loss prevention (DLP) sensitivity information types and pattern detection, or Microsoft 365 regex patterns. Because the scanner uses the Azure Information Protection client, it can classify and protect the same file types.



You can also run the scanner in discovery mode only, where no labels are applied to files and only reports are generated. You can then use the reports to check which labels would be applied or not to file shares and libraries.

Note: The scanner does not discover and label in real time. It runs once or cycling in task jobs and systematically crawls through files on data stores that you specify.

Prerequisites for the Azure Information Protection scanner

The following table identifies the prerequisites for the Azure Information Protection scanner.

Requirement	More information
Windows Server computer to run the scanner service	• Windows Server 2016 or Windows Server 2012 R2 • Minimum 4 processors and 4 GB of RAM
SQL Server to store the scanner configuration	• SQL Server 2012 is the minimum version (Express / Standard / Enterprise)

	• Local or remote instance • Sysadmin role to install the scanner
Service account to run the scanner service	In addition to running the scanner service, this account authenticates to Azure AD and downloads the Azure Information Protection policy. This account must therefore be an Active Directory account that is synchronized to Azure AD, with the following additional requirements: • Log on locally allowed • Log on as a service allowed • Read/Write permissions to the data (file) repositories • To relabel or remove labels: super user role
The Azure Information Protection client is installed on the Windows Server computer	You must install the full client for the scanner. Do not install the client with just the PowerShell module.
Configured labels that apply automatic classification, and optionally, protection	See the previous lesson for configuring labels and policies with automatic conditions for file and document labeling.

Installation and Configuration of the scanner

The installation and configuration of the scanner is completed using PowerShell cmdlets, using the AADRM modules from the Azure Information Protection client and within the Azure Portal.

The following three configuration steps must be performed to install the service itself, create the application registration to access the Azure service, and set up the service with the application registration keys.

Step 1: Install the scanner on the Windows server

1. Prepare the environment to meet the prerequisites.
2. Open an elevated PowerShell and connect to the AADRM service using the following cmdlet:Connect-AadrmService
3. Activate the Super User feature using the following cmdlet:Enable-AadrmSuperUserFeature
4. Configure the synced Azure AD user as a super user using the following cmdlet:Add-AadrmSuperUser -EmailAddress <mail address of the account>
5. Install Azure Information Protection scanner using the following cmdlet:Install-AIPScanner
6. Enter the credentials to create the SQL database and for the context the service shall run in (<domain>\<username>).
7. Enter the credentials for accessing the Azure Information Protection service.
8. Verify that the service is now installed by using **Administrative Tools** > **Services**. Select the **Azure Information Protection Scanner** service and check if the correct user is configured on the **Log On**

Step 2: Configure the scanner in the Azure Portal

From the same Windows Server computer, or from your desktop, sign in to the Azure portal to create two Azure AD applications that are needed to specify an access token for authentication. After an initial interactive sign-in, this token lets the scanner run non-interactively.

You must perform the following steps to create and configure the Azure AD applications:

1. In a new browser window, sign in the **Azure Portal**.
2. For the Azure AD tenant that you use with Azure Information Protection, navigate to Azure **Active Directory** > **App registrations**.
3. Select **New application registration** to create your Web app /API application. On the **Create label**, specify the following values, and then click **Create**:
    - **Name:** AIPOnBehalfOf (must be unique per tenant)
    - **Application Type:** Web app /API

    - **Sign-on URL:** http://localhost
4. Select the **AIPOnBehalfOf** application that you've just created and copy the value for the **Application ID** to a notepad, then close this blade.
5. Select the **Settings** blade and select **Keys**. Add a new key by specifying a **description** and your choice of **duration** (1 year, 2 years, or never expires).
    - **Description:** AIPClient
    - **Expires:** Never expires
6. Then select **Save** and copy the string for the **Value** that is displayed. **Caution:** After closing the blade the Password value cannot be retrieved anymore.
7. Back on the **App registrations** blade, select **New application registration** to create your native application. On the **Create label**, specify the following values, and then click **Create**:
    - **Name:** AIPClient
    - **Application Type:** Native
    - **Sign-on URL:** http://localhost
8. Select the **AIPClient** application that you've just created and copy the value for the **Application ID** to a notepad, then close this blade.
9. On the **Settings** blade, select **Required permissions.**
10. On the **Required permissions** blade, click **Add**, and then click **Select an API**. In the search box, type **AIPOnBehalfOf**. Select this value in the list box, and then click **Select**.
11. On the **Enable Access**blade, select **AIPOnBehalfOf**, click **Select**, and then click **Done**.
12. Back on the **Required permissions** blade, select **Grant Permissions**, click **Yes** to confirm, and then close this blade.

You've now completed the configuration of the two apps, and you have the values that you need to run the **Set-AIPAuthentication** cmdlet with the parameters.

WebAppId	**Application ID** of **AIPOnBehalfOf** application
WebAppKey	Password **Value** of the **AIPOnBehalfOf** application key
NativeAppId	**Application ID** of **AIPClient** application

Step 3: Configure the scanner on the Windows Server

Configure the scanner service with the application registration keys to access your tenant's Azure Information Protection service.

1. Run **Windows PowerShell** as a different user and enter the credentials of the scanner service account.
2. Acquire application access privileges for the user context by running the following cmdlet with the previously created application Ids and key:Set-AIPAuthentication ' -webAppId **<WebAppId>** ' -webAppKey **<WebAppKey>** ' -nativeAppId **<NativeAppId>**
3. When executing the command, a logon window appears. Enter the credentials of the Azure AD synced account with super user access to the Azure Information Protection service.

When the response of the cmdlet is a token acquired for application access, the operation was successful, and the scanner is now able to relabel and unprotect any on-premises content automatically.