User Guide

AWS Application Discovery Service

AWS Application Discovery Service: User Guide

Table of Contents

What is AWS Application Discovery Service?

AWS Application Discovery Service helps you plan your migration to the AWS cloud by collecting usage and configuration data about your on-premises servers and databases. Application Discovery Service is integrated with AWS Migration Hub and AWS Database Migration Service Fleet Advisor. Migration Hub simplifies your migration tracking as it aggregates your migration status information into a single console. You can view the discovered servers, group them into applications, and then track the migration status of each application from the Migration Hub console in your home Region. You can use DMS Fleet Advisor to assess migrations options for database workloads.

All discovered data is stored in your AWS Migration Hub home Region. Therefore, you must set your home Region in the Migration Hub console or with CLI commands before performing any discovery and migration activities. Your data can be exported for analysis in Microsoft Excel or AWS analysis tools such as Amazon Athena and Amazon Quick.

Using Application Discovery Service APIs, you can export the system performance and utilization data for your discovered servers. Input this data into your cost model to compute the cost of running those servers in AWS. Additionally, you can export data about the network connections that exist between servers. This information helps you determine the network dependencies between servers and group them into applications for migration planning.

> **Note**
>
> Your home Region must be set in AWS Migration Hub before you begin the process of discovery, because your data will be stored in your home Region. For more information about working with a home Region, see Home Region.

Application Discovery Service offers three ways of performing discovery and collecting data about your on-premises servers:

- **Agentless discovery** can be performed by deploying the Application Discovery Service Agentless Collector (Agentless Collector) (OVA file) through your VMware vCenter. After Agentless Collector is configured, it identifies virtual machines (VMs) and hosts associated with vCenter. Agentless Collector collects the following static configuration data: Server hostnames, IP addresses, MAC addresses, disk resource allocations, database engine versions, and database schemas.

Additionally, it collects the utilization data for each VM and database providing the average and peak utilization for metrics such as CPU, RAM, and Disk I/O.

- **Agent-based discovery** can be performed by deploying the AWS Application Discovery Agent (Discovery Agent) on each of your VMs and physical servers. The agent installer is available for Windows and Linux operating systems. It collects static configuration data, detailed time-series system-performance information, inbound and outbound network connections, and processes that are running.
- **File-based import** allows you to import details of your on-premises environment directly into Migration Hub without using the Agentless Collector or Discovery Agent, so you can perform migration assessment and planning directly from your imported data. The data ingested is dependent on the data provided.

Application Discovery Service integrates with application discovery solutions from AWS Partner Network (APN) partners. These third-party solutions can help you import details about your on-premises environment directly into Migration Hub, without using any agentless collector or discovery agent. Third-party application discovery tools can query AWS Application Discovery Service, and they can write to the Application Discovery Service database using the public API. In this way, you can import data into Migration Hub and view it, so that you can associate applications with servers and track migrations.

VMware Discovery

If you have virtual machines (VMs) that are running in the VMware vCenter environment, you can use the Agentless Collector to collect system information without having to install an agent on each VM. Instead, you load this on-premises appliance into vCenter and allow it to discover all of its hosts and VMs.

Agentless Collector captures system performance information and resource utilization for each VM running in the vCenter, regardless of what operating system is in use. However, it cannot “look inside” each of the VMs, and as such, cannot figure out what processes are running on each VM nor what network connections exist. Therefore, if you need this level of detail and want to take a closer look at some of your existing VMs in order to assist in planning your migration, you can install the Discovery Agent on an as-needed basis.

Also, for VMs hosted on VMware, you can use both the Agentless Collector and Discovery Agent to perform discovery simultaneously. For details regarding the exact types of data each discovery tool will collect, see Using the VMware vCenter Agentless Collector data collection module.

Database discovery

If you have database and analytics servers in your on-premises environment, then you can use the Agentless Collector to discover and inventory these servers. You can then collect performance metrics for each database server without the need to install Agentless Collector on each computer in your environment.

The Agentless Collector database and analytics data collection module captures metadata and performance metrics that provide insight into your data infrastructure. The database and analytics data collection module uses LDAP in Microsoft Active Directory to gather information about the OS, database, and analytics servers in your network. Then, the data collection module periodically runs queries to collect actual utilization metrics of CPU, memory, and disk capacity for the databases and analytics servers. For details regarding the collected metrics, see Data collected by the database and analytics module.

After Agentless Collector completes data collection from your environment, you can use the AWS DMS console for further analysis and for planning your migration. For example, to choose an optimal migration target in the AWS Cloud, you can generate target recommendations for your source databases. For more information, see Using the database and analytics data collection module.

Compare Agentless Collector and Discovery Agent

The following table provides a quick comparison of the data collection methods that Application Discovery Service supports.

	Agentless Collector	Discovery Agent	Migration Hub template	RVTools export
Supported server types				
VMware virtual machine	Yes	Yes	Yes	Yes
Physical server	No	Yes	Yes	Yes
Deployment				
Per server	No	Yes	N/A	No

	Agentless Collector	Discovery Agent	Migration Hub template	RVTools export
Per vCenter	Yes	No	N/A	Yes
Per data center on the same network	No	No	N/A	No
Collected data				
Server profile (static configuration) data	Yes	Yes	Yes	Yes
Server utilization metrics from Hypervisor (CPU, RAM, etc.)	Yes	Yes	Yes	No
Server utilization metrics from server (CPU, RAM, etc.)	Yes	Yes	Yes	No
Server network connections (TCP only)	Yes	Yes	No	No
Running processes	No	Yes	No	No
Collection interval	-60 minutes	-15 seconds	Single snapshot	Single snapshot
Server data use cases				
View server data in Migration Hub	Yes	Yes	Profile only	No

	Agentless Collector	Discovery Agent	Migration Hub template	RVTools export
Generate Amazon EC2 recommendation based on server profile	Yes	Yes	Yes	Yes
Generate Amazon EC2 recommendation based on utilization data	Yes	Yes	Yes	No
Export of latest utilization snapshot data	Yes	Yes	Yes	No
Export of time series utilization data	No	Yes	No	No
Network data use cases				
Visualization in Migration Hub	Yes	Yes	No	No
Export to Amazon Athena for further exploration	No	Yes	No	No
Export to CSV file	No	Yes	No	No
Database use cases				

	Agentless Collector	Discovery Agent	Migration Hub template	RVTools export
Database server profile (static configuration) data	Yes	No	No	No
Supported database engines	Oracle, SQL Server, MySQL, PostgreSQL	None	None	None
Database schema complexity and duplicates	Yes	No	No	No
Database schema objects	Yes	No	No	No
Platform support				
Supported operating systems	Any OS running in VMware center v5.5 or newer versions	Any Linux or Windows server	Any Linux or Windows server	Any Linux server, Windows server, or VMware v5.5 or newer versions

Assumptions

To use Application Discovery Service, the following is assumed:

- You have signed up for AWS. For more information, see Setting up Application Discovery Service.
- You have selected a Migration Hub home Region. For more information, see the documentation regarding home Regions.

Here's what to expect:

- The Migration Hub home Region is the only Region where Application Discovery Service stores your discovery and planning data.
- Discovery agents, connectors, and imports can be used in your selected Migration Hub home Region only.
- For a list of AWS Regions where you can use Application Discovery Service, see the Amazon Web Services General Reference.

AWS Application Discovery Service availability change

After careful consideration, we decided to close AWS Application Discovery Service to new customers starting November 7, 2025. If you would like to use the service, sign up prior to that date. Existing customers can continue to use the service as normal.

This topic provides information about the availability change and guidance for transitioning to AWS Transform.

Service availability details

Application Discovery Service will stop accepting new customers starting November 7, 2025. AWS Transform is our next-generation, agentic AI service that provides similar capabilities and enhanced VM discovery and assessment capabilities. Existing Application Discovery Service customers can continue using the service to complete their ongoing discovery projects, which typically have a 4-month lifecycle. The service's core functionality for discovering and collecting data about on-premises servers and applications is now available in AWS Transform with improved features, requiring no customer effort for transition.

Until November 7, 2025, we will continue to maintain Application Discovery Service's security and reliability. While we will not be adding new features to the service, we remain committed to providing security updates and maintaining service availability to ensure your ongoing migration projects continue to run smoothly. Our focus is on ensuring a stable environment for existing customers to complete their in-flight migration initiatives while preparing for the enhanced capabilities available in AWS Transform.

AWS Transform transition

AWS Transform is our recommended solution that brings together all Application Discovery Service capabilities while introducing powerful new features. It provides comprehensive discovery and assessment capabilities through agent-based and agentless collectors, with enhanced VMware environment analysis. The service enables automated application dependency mapping and wave planning, while offering improved discovery logic with what-if analysis and cost estimates. With advanced features including integrated storage and database discovery, consolidated 3P tool integration, and comprehensive VM configuration analysis, AWS Transform is designed to make customers' migration assessment and planning process more efficient and successful.

Transitioning to AWS Transform is straightforward, with no data migration required. Existing discovery projects in Application Discovery Service will continue to function normally until completion. When customers are ready to start new discovery projects, they can begin using AWS Transform directly - all the discovery and assessment capabilities from Application Discovery Service are available there with enhanced features. To begin using AWS Transform please refer to the Getting Started Guide. AWS Support team is available through the AWS Support console to assist with AWS Transform access or questions about ongoing discovery projects.

Frequently asked questions

What does this mean for the service (are you going to shut the service down)?

Application Discovery Service will stop accepting new customers starting November 7, 2025. The service will continue to operate for existing customers to complete their ongoing migration projects.

How will existing customers be impacted?

Existing customers will not experience any disruption to their current migration projects. They can continue using Application Discovery Service as normal until their projects are completed. All ongoing projects will remain accessible, and security updates will continue to be deployed to maintain service reliability.

On November 7, 2025, the customer is having issues, how can they escalate?

Customers experiencing issues can contact AWS Support through their AWS Support console. The AWS Support team is available to assist with any service-related questions or concerns.

What alternatives can customers explore?

AWS Transform is the recommended alternative service. Launched in 2025, AWS Transform includes similar Application Discovery Service capabilities. Offers enhanced features with improved VMware environment analysis and automated dependency mapping. Provides integrated storage and database discovery capabilities and comprehensive assessment tools. No special tooling required when transitioning to AWS Transform.

How can customers migrate off of Application Discovery Service?

No formal migration process is required. Existing projects can continue in Application Discovery Service until completion. For new projects, customers can start directly in AWS Transform, which

provides all the familiar capabilities of Application Discovery Service with enhanced features. No data migration is needed, and AWS Support is available to assist with the transition.

If you have additional questions, please contact us through AWS Support or read our FAQs.

Setting up Application Discovery Service

Before you use AWS Application Discovery Service for the first time, complete the following tasks:

Create IAM users

Sign in to the Migration Hub console and choose a home Region

Sign up for an AWS account

To get started with AWS, you need an AWS account. For information about creating an AWS account, see Getting started with an AWS account in the *AWS Account Management Reference Guide*.

Create IAM users

Topics

- Creating an IAM Non-Administrative User

Creating an IAM Non-Administrative User

When creating non-administrative IAM users, follow the security best practice Grant Least Privilege, granting users minimum permissions.

Use IAM managed policies to define the level of access to Application Discovery Service by non-administrative IAM users. For information about Application Discovery Service managed policies, see AWS managed policies for AWS Application Discovery Service.

To create a non-administrator IAM user

1. In AWS Management Console, navigate to the IAM console.
2. Create a non-administrator IAM user by following the instructions for creating a user with the console as described in Creating an IAM user in your AWS account in the *IAM User Guide*.

   While following the instructions in the *IAM User Guide*:

   - When on the step about the **Set permissions** page, choose the option to **Attach existing policies to user directly**. Then select a managed IAM policy for Application Discovery

Service from the list of policies. For information about Application Discovery Service managed policies, see AWS managed policies for AWS Application Discovery Service.

- When on the step about viewing the user's access keys (access key IDs and secret access keys), follow the guidance in the **Important** note about saving the user's new access key ID and secret access key in a safe and secure place.

3. After you create the user provide them with programmatic access as described in Support programmatic user access.

Sign in to the Migration Hub console and choose a home Region

You need to choose an AWS Migration Hub home Region in the AWS account that you're using for the AWS Application Discovery Service.

To choose a home Region

1. Using your AWS account, sign in to the AWS Management Console and open the Migration Hub console at https://console.aws.amazon.com/migrationhub/.
2. In the Migration Hub console navigation pane, choose **Settings** and the choose a home Region.

   Your Migration Hub data is stored in your home Region for purposes of discovery, planning, and migration tracking. For more information, see The Migration Hub Home Region.

AWS Application Discovery Agent

The AWS Application Discovery Agent (Discovery Agent) is software that you install on on-premises servers and VMs targeted for discovery and migration. Agents capture system configuration, system performance, running processes, and details of the network connections between systems. Agents support most Linux and Windows operating systems, and you can deploy them on physical on-premises servers, Amazon EC2 instances, and virtual machines.

> **Note**
>
> Before you deploy the Discovery Agent, you must choose a Migration Hub home Region. You must register your agent in your home Region.

The Discovery Agent runs in your local environment and requires root privileges. When you start the Discovery Agent, it connects securely with your home region and registers with Application Discovery Service.

- For example, if `eu-central-1` is your home Region, it registers `arsenal-discovery.eu-central-1.amazonaws.com` with Application Discovery Service.
- Or substitute your home Region as needed for all other Regions except us-west-2.
- If `us-west-2` is your home Region, it registers `arsenal.us-west-2.amazonaws.com` with Application Discovery Service.

How it works

After registration, the agent starts collecting data for the host or VM where it resides. The agent pings the Application Discovery Service at 15-minute intervals for configuration information.

The collected data includes system specifications, times series utilization or performance data, network connections, and process data. You can use this information to map your IT assets and their network dependencies. All of these data points can help you determine the cost of running these servers in AWS and also plan for migration.

Data is transmitted securely by the Discovery Agents to Application Discovery Service using Transport Layer Security (TLS) encryption. Agents are configured to upgrade automatically when new versions become available. You can change this configuration setting if desired.

Tip

Before downloading and beginning Discovery Agent installation, be sure to read through all of the required prerequisites in Prerequisites for Discovery Agent

Data collected by Discovery Agent

AWS Application Discovery Agent (Discovery Agent) is software that you install on on-premises servers and VMs. Discovery Agent collects system configuration, times series utilization or performance data, process data, and Transmission Control Protocol (TCP) network connections. This section describes the data that's collected.

Table legend for Discovery Agent collected data:

- The term host refers to either a physical server or a VM.
- Collected data is in measurements of kilobytes (KB) unless stated otherwise.
- Equivalent data in the Migration Hub console is reported in megabytes (MB).
- The polling period is in intervals of approximately 15 seconds and is sent to AWS every 15 minutes.
- Data fields denoted with an asterisk (*) are only available in the `.csv` files that are produced from the agent's API export function.

Data field	Description
agentAssignedProcessId*	Process ID of processes discovered by the agent
agentId	Unique ID of agent
agentProvidedTimeStamp*	Date and time of agent observation *(mm/dd/yyyy hh:mm:ss am/pm)*
cmdLine*	Process entered at the command line
cpuType	Type of CPU (central processing unit) used in host

Data field	Description
destinationIp*	IP address of device to which packet is being sent
destinationPort*	Port number to which the data/request is to be sent
family*	Protocol of routing family
freeRAM (MB)	Free RAM and cached RAM that can be made immediately available to applications, measured in MB
gateway*	Node address of network
hostName	Name of host data was collected on
hypervisor	Type of hypervisor
ipAddress	IP address of the host
ipVersion*	IP version number
isSystem*	Boolean attribute to indicate if a process is owned by the OS
macAddress	MAC address of the host
name*	Name of the host, network, metrics, etc. data is being collected for
netMask*	IP address prefix that a network host belongs to
osName	Operating system name on host
osVersion	Operating system version on host
path	Path of the command sourced from the command line

Data field	Description
sourceIp*	IP address of the device sending the IP packet
sourcePort*	Port number from which the data/request originates from
timestamp*	Date and time of reported attribute logged by agent
totalCpuUsagePct	Percentage of CPU usage on host during polling period
totalDiskBytesReadPerSecond (Kbps)	Total kilobits read per second across all disks
totalDiskBytesWrittenPerSecond (Kbps)	Total kilobits written per second across all disks
totalDiskFreeSize (GB)	Free disk space expressed in GB
totalDiskReadOpsPerSecond	Total number of read I/O operations per second
totalDiskSize (GB)	Total capacity of disk expressed in GB
totalDiskWriteOpsPerSecond	Total number of write I/O operations per second
totalNetworkBytesReadPerSecond (Kbps)	Total amount of throughput of bytes read per second
totalNetworkBytesWrittenPerSecond (Kbps)	Total amount of throughput of bytes written per second
totalNumCores	Total number of independent processing units within CPU
totalNumCpus	Total number of central processing units
totalNumDisks	The number of physical hard disks on a host

Data field	Description
totalNumLogicalProcessors*	Total number of physical cores times the number of threads that can run on each core
totalNumNetworkCards	Total count of network cards on server
totalRAM (MB)	Total amount of RAM available on host
transportProtocol*	Type of transport protocol used

Prerequisites for Discovery Agent

The following are the prerequisites and the tasks that you must perform before you can successfully install the AWS Application Discovery Agent (Discovery Agent).

- You must set an AWS Migration Hub home region before you begin installing Discovery Agent.
- If you have a *1.x* version of the agent installed, it must be removed before installing the latest version.
- If the host that the agent is being installed on runs Linux, then verify that the host at least supports the Intel i686 CPU architecture (also known as the P6 micro architecture).
- Generate access keys needed to install Discovery Agent.
- Verify that your operating system (OS) environment is supported:

  Linux

  Amazon Linux 2012.03, 2015.03

  Amazon Linux 2 (9/25/2018 update and later)

  Ubuntu 12.04, 14.04, 16.04, 18.04, 20.04

  Red Hat Enterprise Linux 5.11, 6.10, 7.3, 7.7, 8.1

  CentOS 5.11, 6.9, 7.3

  SUSE 11 SP4, 12 SP5, 15 SP5

  Windows

  Windows Server 2003 R2 SP2

  Windows Server 2008 R1 SP2, 2008 R2 SP1

  Windows Server 2012 R1, 2012 R2

Windows Server 2016

Windows Server 2019

Windows Server 2022

- If outbound connections from your network are restricted, you'll need to update your firewall settings. Agents require access to `arsenal` over TCP port 443. They don't require any inbound ports to be open.

  For example, if your home Region is `eu-central-1`, you'd use `https://arsenal-discovery.eu-central-1.amazonaws.com:443`
- Access to Amazon S3 in your home region is required for auto-upgrade to function.
- Create an AWS Identity and Access Management (IAM) user in the console and attach the existing `AWSApplicationDiscoveryAgentAccess` IAM managed policy. This policy allows the user to perform necessary agent actions on your behalf. For more information about managed policies, see AWS managed policies for AWS Application Discovery Service.
- Check the time skew from your Network Time Protocol (NTP) servers and correct if necessary. Incorrect time synchronization causes the agent registration call to fail.

> **Note**
>
> The Discovery Agent has a 32-bit agent executable, which works on 32-bit and 64-bit operating systems. The number of installation packages needed for deployment is reduced by having a single executable. This executable agent works for Linux and for Windows OS. It is addressed in their respective installation sections that follow.

Installing Discovery Agent

This page covers how to install the Discovery Agent on Linux and Microsoft Windows.

Install Discovery Agent on Linux

Complete the following procedure on Linux. Be sure that your Migration Hub home region has been set before you begin this procedure.

> **Note**
>
> If you are using a non-current Linux version, see Considerations with older Linux platforms.

To install AWS Application Discovery Agent in your data center

1. Sign in to your Linux-based server or VM and create a new directory to contain your agent components.
2. Switch to the new directory and download the installation script from either the command line or the console.

   a. To download from the command line, run the following command.

   ```
   curl -o ./aws-discovery-agent.tar.gz https://s3.region.amazonaws.com/aws-discovery-agent.region/linux/latest/aws-discovery-agent.tar.gz
   ```

   b. To download from the Migration Hub console, do the following:

      i. Sign in to the AWS Management Console and open the Migration Hub console at https://console.aws.amazon.com/migrationhub/.

      ii. In the left navigation page, under **Discover**, choose **Tools**.

      iii. In the **AWS Discovery Agent** box, choose **Download agents**, then choose **Download for Linux**. Your download begins immediately.

3. Verify the cryptographic signature of the installation package with the following three commands:

   ```
   curl -o ./agent.sig https://s3.region.amazonaws.com/aws-discovery-agent.region/linux/latest/aws-discovery-agent.tar.gz.sig
   ```

   ```
   curl -o ./discovery.gpg https://s3.region.amazonaws.com/aws-discovery-agent.region/linux/latest/discovery.gpg
   ```

   ```
   gpg --no-default-keyring --keyring ./discovery.gpg --verify agent.sig aws-discovery-agent.tar.gz
   ```

The agent public key (`discovery.gpg`) fingerprint is `7638 F24C 6717 F97C 4F1B 3BC0 5133 255E 4DF4 2DA2`.

4. Extract from the tarball as shown following.

```
tar -xzf aws-discovery-agent.tar.gz
```

5. To install the agent, choose one of the following installation methods.

To...	Do this...
Install Discovery Agent	To install the agent, run the agent install command as shown in the following example. In the example, replace *your-home-region* with the name of your home region, *aws-access-key-id* with your access key id, and *aws-secret-access-key* with your secret access key. `sudo bash install -r your-home-region -k aws-access-key-id -s aws-secret-access-key` By default, agents automatically download and apply updates as they become available . We recommend using this default configura tion. However, if you don't want agents to download and apply updates automatic ally, include the `-u false` parameter when running the agent install command.
(Optional) Install Discovery Agent and configure a non-transparent proxy	To configure a non-transparent proxy, add the following parameters to the agent install command:

To...	Do this...
	• **-e** The proxy password. • **-f** The proxy port number. • **-g** The proxy scheme. • **-i** The proxy username. The following is an example of the agent install command using the non-transparent proxy parameters. `sudo bash install -r your-home-region -k aws-access-key-id -s aws-secret-access-key -d myproxy.mycompany.com -e mypassword -f proxy-port-number -g https -i myusername` If your proxy doesn't require authentication, then leave out the `-e` and `-i` parameters. The example install command uses `https`, if your proxy uses HTTP, specify `http` for the `-g` parameter value.

6. If outbound connections from your network are restricted, you'll need to update your firewall settings. Agents require access to `arsenal` over TCP port 443. They don't require any inbound ports to be open.

   For example, if your home Region is `eu-central-1`, you'd use `https://arsenal-discovery.eu-central-1.amazonaws.com:443`

Considerations with older Linux platforms

Some older Linux platforms such as SUSE 10, CentOS 5, and RHEL 5 are either at end of life or only minimally supported. These platforms can suffer from out-of-date cipher suites that prevent the agent update script from downloading installation packages.

Curl

The Application Discovery agent requires `curl` for secure communications with the AWS server. Some old versions of `curl` are not able to communicate securely with a modern web service.

To use the version of `curl` included with the Application Discovery agent for all operations, run the installation script with the `-c true` parameter.

Certificate Authority Bundle

Older Linux systems might have an out-of-date Certificate Authority (CA) bundle, which is critical to secure internet communication.

To use the CA bundle included with the Application Discovery agent for all operations, run the installation script with the `-b true` parameter.

These installation script options can be used together. In the following example command, both of the script parameters are passed to the installation script:

```
sudo bash install -r your-home_region -k aws-access-key-id -s aws-secret-access-key -c
 true -b true
```

Install Discovery Agent on Microsoft Windows

Complete the following procedure to install an agent on Microsoft Windows. Be sure that your Migration Hub home region has been set before you begin this procedure.

To install AWS Application Discovery Agent in your data center

1. Download the Windows agent installer *but do not double-click to run the installer within Windows*.

> **Important**
>
> Do not double-click to run the installer within Windows as it will fail to install. *Agent installation only works from the command prompt*. (If you already double-clicked on the installer, you must go to **Add/Remove Programs** and uninstall the agent before continuing on with the remaining installation steps.)

If the Windows agent installer doesn't detect any version of the Visual C++ x86 runtime on the host, it automatically installs the Visual C++ x86 2015–2019 runtime before installing the agent software.

2. Open a command prompt as an administrator and navigate to the location where you saved the installation package.
3. To install the agent, choose one of the following installation methods.

To...	Do this...
Install Discovery Agent	To install the agent, run the agent install command as shown in the following example. In the example, replace *your-home-region* with the name of your home region, *aws-access-key-id* with your access key ID, and *aws-secret-access-key* with your secret access key. Optionally, you can set the agent installation location by specifying the folder path *C:\install-location* for the INSTALLLOCATION parameter. For example, `INSTALLLOCATION="` *C:\install-location* `"`. The resulting folder hierarchy will be [INSTALLLOCATION path]\AWS Discovery. By default, the install location is the `Program Files` folder. Optionally, you can use `LOGANDCONFIGLOCATION` to override the default directory (ProgramData) for the agent logs folder and configuration file. The resulting folder hierarchy is [*LOGANDCONFIGLOCATION path*]`\AWS Discovery`.

To...	Do this...
	`.\AWSDiscoveryAgentInstaller.exe REGION=" your-home-region " KEY_ID="aws-access-key-id " KEY_SECRET=" aws-secret-access-key " /quiet` By default, agents automatically download and apply updates as they become available . We recommend using this default configuration. However, if you don't want agents to download and apply updates automatically, include the following parameter when running the agent install command: `AUTO_UPDATE=false` **Warning** Disabling auto-upgrades will prevent the latest security patches from being installed.

To...	Do this...
(Optional) Install Discovery Agent and configure a non-transparent proxy	To configure a non-transparent proxy, add the following public properties to the agent install command: • **PROXY_HOST** – The name of the proxy host • **PROXY_SCHEME** – The proxy scheme • **PROXY_PORT** – The proxy port number • **PROXY_USER** – The proxy user name • **PROXY_PASSWORD** – The proxy user password The following is an example of the agent install command using the non-transparent proxy properties. `.\AWSDiscoveryAgentInstaller.exe REGION=" your-home-region " KEY_ID="aws-access-key-id " KEY_SECRET=" aws-secret-access-key " PROXY_HOST=" myproxy.mycompany.com " PROXY_SCHEME="https" PROXY_PORT=" proxy-port-number " PROXY_USER=" myusername " PROXY_PASSWORD=" mypassword " /quiet` If your proxy doesn't require authentication, then omit the `PROXY_USER` and `PROXY_PASSWORD` properties. The example install command uses `https`. If your proxy uses HTTP, specify `http` for the `PROXY_SCHEME` value.

4. If outbound connections from your network are restricted, you must update your firewall settings. Agents require access to `arsenal` over TCP port 443. They don't require any inbound ports to be open.

   For example, if your home Region is `eu-central-1`, you'd use the following: `https://arsenal-discovery.eu-central-1.amazonaws.com:443`

Package signing and automatic upgrades

For Windows Server 2008 and later, Amazon cryptographically signs the Application Discovery Service agent installation package with an SHA256 certificate. For SHA2-signed autoupdates on Windows Server 2008 SP2, ensure that hosts have a hotfix installed to support SHA2 signature authentication. Microsoft's latest support hotfix helps support SHA2 authentication on Windows Server 2008 SP2.

> **Note**
>
> The hotfixes for SHA256 support for Windows 2003 are no longer publicly available from Microsoft. If these fixes are not already installed in your Windows 2003 host, manual upgrades are necessary.

To perform upgrades manually

1. Download the Windows Agent Updater.
2. Open command prompt as an administrator.
3. Navigate to the location where the updater was saved.
4. Run the following command.

```
AWSDiscoveryAgentUpdater.exe /Q
```

Managing the Discovery Agent process

This page covers how to manage the Discovery Agent on Linux and Microsoft Windows.

Manage the Discovery Agent process on Linux

You can manage the behavior of the Discovery Agent at the system level using the `systemd`, `Upstart`, or `System V init` tools. The following tabs outline the commands for the supported tasks in each of the respective tools.

systemd

Management Commands for the Application Discovery Agent

Task	Command
Verify that an agent is running	`sudo systemctl status aws-discovery-daemon.service`
Start an agent	`sudo systemctl start aws-discovery-daemon.service`
Stop an agent	`sudo systemctl stop aws-discovery-daemon.service`
Restart an agent	`sudo systemctl restart aws-discovery-daemon.service`

Upstart

Management commands for the Application Discovery Agent

Task	Command
Verify that an agent is running	`sudo initctl status aws-discovery-daemon`
Start an agent	`sudo initctl start aws-discovery-daemon`
Stop an agent	`sudo initctl stop aws-discovery-daemon`
Restart an agent	`sudo initctl restart aws-discovery-daemon`

System V init

Management commands for the Application Discovery Agent

Task	Command
Verify that an agent is running	`sudo /etc/init.d/aws-discovery-daemon status`
Start an agent	`sudo /etc/init.d/aws-discovery-daemon start`
Stop an agent	`sudo /etc/init.d/aws-discovery-daemon stop`
Restart an agent	`sudo /etc/init.d/aws-discovery-daemon restart`

Manage the Discovery Agent process on Microsoft Windows

You can manage the behavior of the Discovery Agent at the system level through the Windows Server Manager Services console. The following table describes how.

Task	Service Name	Service Status/Action
Verify that an agent is running	AWS Discovery Agent AWS Discovery Updater	Started
Start an agent	AWS Discovery Agent AWS Discovery Updater	Choose **Start**
Stop an agent	AWS Discovery Agent AWS Discovery Updater	Choose **Stop**
Restart an agent	AWS Discovery Agent AWS Discovery Updater	Choose **Restart**

Uninstalling Discovery Agent

This page covers how to uninstall the Discovery Agent on Linux and Microsoft Windows.

Uninstall Discovery Agent on Linux

This section describes how to uninstall Discovery Agent on Linux.

To uninstall an agent if you're using the yum package manager

- Use the following command to uninstall an agent if using yum.

```
rpm -e --nodeps aws-discovery-agent
```

To uninstall an agent if you're using the apt-get package manager

- Use the following command to uninstall an agent if using apt-get.

```
apt-get remove aws-discovery-agent:i386
```

To uninstall an agent if you're using the zypper package manager

- Use the following command to uninstall an agent if using zypper.

```
zypper remove aws-discovery-agent
```

Uninstall Discovery Agent on Microsoft Windows

This section describes how to uninstall Discovery Agent on Microsoft Windows.

To uninstall a discovery agent on Windows

1. Open the Control Panel in Windows.
2. Choose **Programs**.
3. Choose **Programs and Features**.
4. Select **AWS Discovery Agent**.

5. Choose **Uninstall**.

> **Note**
>
> If you choose to reinstall the agent after uninstalling it, run the following command with the `/repair` and `/norestart` options.
>
> ```
> .\AWSDiscoveryAgentInstaller.exe REGION="your-home-region" KEY_ID="aws-access-key-id" KEY_SECRET="aws-secret-access-key" /quiet /repair /norestart
> ```

To uninstall a discovery agent on Windows using the command line

1. Right-click **Start**.
2. Choose **Command Prompt**.
3. Use the following command to uninstall a discovery agent on Windows.

```
wmic product where name='AWS Discovery Agent' call uninstall
```

> **Note**
>
> If the `.exe` file is present on the server, you can uninstall the agent completely from the server by using the following command. If you use this command to uninstall, you don't need to use the `/repair` and `/norestart` options when you reinstall the agent.
>
> ```
> .\AWSDiscoveryAgentInstaller.exe /quiet /uninstall
> ```

Starting and stopping Discovery Agent data collection

After the Discovery Agent is deployed and configured, if data collections stops you can restart it. You can start or stop data collection through the console by following the steps in Starting and stopping data collectors in the AWS Migration Hub console, or by making API calls through the AWS CLI. Before starting be sure to generate access keys needed to manage the Discovery Agent.

To install the AWS CLI and start or stop data collection

1. If you have not already done so, install the AWS CLI appropriate to your OS type (Windows or Mac/Linux). See the AWS Command Line Interface User Guide for instructions.
2. Open the Command prompt (Windows) or Terminal (MAC/Linux).
   a. Type `aws configure` and press Enter.
   b. Enter your AWS Access Key ID and AWS Secret Access Key.
   c. Enter your home Region for the Default Region Name, for example *us-west-2*. (We are assuming that `us-west-2` is your home Region in this example.)
   d. Enter `text` for Default Output Format.
3. To find the ID of the agent you want to stop or start data collection for, type the following command:

```
aws discovery describe-agents
```

4. To start data collection by the agent, type the following command:

```
aws discovery start-data-collection-by-agent-ids --agent-ids <agent ID>
```

   To stop data collection by the agent, type the following command:

```
aws discovery stop-data-collection-by-agent-ids --agent-ids <agent ID>
```

Troubleshooting Discovery Agent

This page covers troubleshooting the Discovery Agent on Linux and Microsoft Windows.

Troubleshooting Discovery Agent on Linux

If you encounter problems while installing or using the Discovery Agent on Linux, consult the following guidance about logging and configuration. When helping to troubleshoot potential issues with the agent or its connection to the Application Discovery Service, AWS Support often requests these files.

- **Log files**

Log files for Discovery Agent are located in the following directory.

```
/var/log/aws/discovery/
```

Log files are named to indicate whether they are generated by the main daemon, the automatic upgrader, or the installer.

- **Configuration files**

  Configuration files for Discovery Agent version 2.0.1617.0 or newer are located in the following directory.

  ```
  /etc/opt/aws/discovery/
  ```

  Configuration files for versions of Discovery Agent before 2.0.1617.0 are located in the following directory.

  ```
  /var/opt/aws/discovery/
  ```

- For instructions on how to remove older versions of the Discovery Agent, see Prerequisites for Discovery Agent.

Troubleshooting Discovery Agent on Microsoft Windows

If you encounter problems while installing or using the AWS Application Discovery Agent on Microsoft Windows, consult the following guidance about logging and configuration. AWS Supportoften requests these files when helping to troubleshoot potential issues with the agent or its connection to the Application Discovery Service.

- **Installation logging**

  In some cases, the agent install command appears to fail. For example, a failure can appear with the Windows Services Manager showing that the discovery services are not being created. In this case, add **/log install.log** to the command to generate a verbose installation log.

- **Operational logging**

On Windows Server 2008 and later, agent log files can be found under the following directory.

```
C:\ProgramData\AWS\AWS Discovery\Logs
```

On Windows Server 2003, agent log files can be found under the following directory.

```
C:\Documents and Settings\All Users\Application Data\AWS\AWS Discovery\Logs
```

Log files are named to indicate whether generated by the main service, automatic upgrades, or the installer.

- **Configuration file**

  On Windows Server 2008 and later, the agent configuration file can be found at the following location.

  ```
  C:\ProgramData\AWS\AWS Discovery\config
  ```

  On Windows Server 2003, the agent configuration file can be found at the following location.

  ```
  C:\Documents and Settings\All Users\Application Data\AWS\AWS Discovery\config
  ```

- For instructions on how to remove earlier versions of the Discovery Agent, see Prerequisites for Discovery Agent.

Application Discovery Service Agentless Collector

Application Discovery Service Agentless Collector (Agentless Collector) is an on-premises application that collects information through agentless methods about your on-premises environment, including server profile information (for example, OS, number of CPUs, amount of RAM), database metadata, utilization metrics, and data about network traffic among on-premises servers. You install the Agentless Collector as a virtual machine (VM) in your VMware vCenter Server environment using an Open Virtualization Archive (OVA) file.

Agentless Collector has a modular architecture, which allows for the use of multiple agentless collection methods. Agentless Collector provides modules for data collection from VMware VMs and from database and analytics servers. It also provides a module for collecting data about network traffic among your on-premises servers.

Agentless Collector supports data collection for AWS Application Discovery Service (Application Discovery Service) by collecting usage and configuration data about your on-premises servers and databases, as well as data about network traffic among your on-premises servers.

Application Discovery Service is integrated with AWS Migration Hub, a service that simplifies your migration tracking as it aggregates your migration status information into a single console. You can view the discovered servers, obtain Amazon EC2 recommendations, visualize network connections, group servers into applications, and then track the migration status of each application from the Migration Hub console in your home Region.

The Agentless Collector database and analytics data collection module is integrated with AWS Database Migration Service (AWS DMS). This integration helps plan your migration to the AWS Cloud. You can use the database and analytics data collection module to discover database and analytics servers in your environment and build an inventory of servers that you want to migrate to the AWS Cloud. This data collection module collects database metadata and actual utilization metrics of CPU, memory, and disk capacity. After you collect these metrics, you can use the AWS DMS console to generate target recommendations for your source databases.

Prerequisites for Agentless Collector

The following are the prerequisites for using Application Discovery Service Agentless Collector (Agentless Collector):

- One or more AWS accounts.

- An AWS account with the AWS Migration Hub home Region set, see Sign in to the Migration Hub console and choose a home Region. Your Migration Hub data is stored in your home Region for purposes of discovery, planning, and migration tracking.
- An AWS account IAM user that is set up to use the AWS managed policy `AWSApplicationDiscoveryAgentlessCollectorAccess`. To use the database and analytics data collection module, this IAM user must also use two customer managed IAM policies `DMSCollectorPolicy` and `FleetAdvisorS3Policy`. For more information, see Deploying Application Discovery Service Agentless Collector. The IAM user must be created in an AWS account with Migration Hub home Region set.
- VMware vCenter Server V5.5, V6, V6.5, 6.7 or 7.0.

> **Note**
>
> The Agentless Collector supports all of these versions of VMware, but we currently test against version 6.7 and 7.0.

- For VMware vCenter Server setup, make sure that you can provide vCenter credentials with Read and View permissions set for the System group.
- Agentless Collector requires outbound access over TCP port 443 to several AWS domains. For a list of these domains, see Configure firewall for outbound access to AWS domains.
- To use the database and analytics data collection module, create an Amazon S3 bucket in the AWS Region that you set as your Migration Hub home Region. The database and analytics data collection modules stores inventory metadata in this Amazon S3 bucket. For more information, see Creating a bucket in the *Amazon S3 User Guide*.
- Agentless Collector version 2 requires ESXi 6.5 or a later version.

Configure data perimeter for access to AWS service-owned resources

The Agentless Collector automatic update feature retrieves updates in the form of Docker images from an AWS service-owned Public ECR Repository. If you are using data perimeters to control access to Amazon ECR in your environment, you might need to explicitly allow access to the following to use the automatic update feature:

- Resource ARNs that require access: `arn:aws:ecr-public::446372222237:repository/6e5498e4-8c31-4f57-9991-13b4b992ff7b`
- Required permissions: `ecr-public:DescribeImages`

Configure firewall for outbound access to AWS domains

If outbound connections from your network are restricted, you must update your firewall settings to allow outbound access to the AWS domains that Agentless Collector requires. Which AWS domains require outbound access depend on if your Migration Hub home Region is US West (Oregon) Region, us-west-2, or some other Region.

The following domains require outbound access if your AWS account home Region is us-west-2:

- `arsenal-discovery.us-west-2.amazonaws.com` – The collector uses this domain to validate that it is configured with the required IAM user credentials. The collector also uses it for sending and storing collected data since the home Region is us-west-2.
- `migrationhub-config.us-west-2.amazonaws.com` – The collector uses this domain to determine which home Region the collector sends data to based on the provided IAM user credentials.
- `api.ecr-public.us-east-1.amazonaws.com` – The collector uses this domain to discover available updates.
- `public.ecr.aws` – The collector uses this domain for downloading the updates.
- `dms.`*`your-migrationhub-home-region`*`.amazonaws.com` – The collector uses this domain to connect to the AWS DMS data collector.
- `s3.amazonaws.com` – The collector uses this domain to upload data that is collected by the database and analytics data collection module to your Amazon S3 bucket.
- `sts.amazonaws.com` – The collector uses this domain to understand what account the collector has been configured with.

The following domains require outbound access if your AWS account home Region is not `us-west-2`:

- `arsenal-discovery.us-west-2.amazonaws.com` – The collector uses this domain to validate that it is configured with the required IAM user credentials.
- `arsenal-discovery.`*`your-migrationhub-home-region`*`.amazonaws.com` – The collector uses this domain for sending and storing collected data.
- `migrationhub-config.us-west-2.amazonaws.com` – The collector uses this domain to determine which home Region the collector should send data to based on the provided IAM user credentials.

- `api.ecr-public.us-east-1.amazonaws.com` – The collector uses this domain to discover available updates.
- `public.ecr.aws` – The collector uses this domain for downloading the updates.
- `dms.`*`your-migrationhub-home-region`*`.amazonaws.com` – The collector uses this domain to connect to the AWS DMS data collector.
- `s3.amazonaws.com` – The collector uses this domain to upload data that is collected by the database and analytics data collection module to your Amazon S3 bucket.
- `sts.amazonaws.com` – The collector uses this domain to understand what account the collector has been configured with.

When setting up Agentless Collector, you might receive errors such as **Setup failed – Check your credentials and try again** or **AWS cannot be reached. Please verify network settings**. These errors can be caused by a failed attempt by the Agentless Collector to establish an HTTPS connection to one of the AWS domains that it needs outbound access to.

If a connection to AWS cannot be established, Agentless Collector cannot collect data from your on-premises environment. For information about how to fix the connection to AWS, see Fixing Agentless Collector cannot reach AWS during setup.

Deploying Application Discovery Service Agentless Collector

To deploy Application Discovery Service Agentless Collector, you must first create an IAM user and download the collector. This page walks you through the steps to take to deploy a collector.

Create an IAM user for Agentless Collector

To use Agentless Collector, in the AWS account that you used in Sign in to the Migration Hub console and choose a home Region, you must create an AWS Identity and Access Management (IAM) user. Then, set up this IAM user to use the following AWS managed policy AWSApplicationDiscoveryAgentlessCollectorAccess. You attach this IAM policy when you create the IAM user.

To use the database and analytics data collection module, create two customer managed IAM policies. These policies provide access your Amazon S3 bucket and the AWS DMS API. For more information, see Create a customer managed policy in the *IAM User Guide*.

- Use the following JSON code to create the **DMSCollectorPolicy** policy.

  JSON

```
{
    "Version":"2012-10-17",
    "Statement": [{
        "Effect": "Allow",
        "Action": [
            "dms:DescribeFleetAdvisorCollectors",
            "dms:ModifyFleetAdvisorCollectorStatuses",
            "dms:UploadFileMetadataList"
        ],
        "Resource": "*"
    }]
}
```

- Use the following JSON code to create the **FleetAdvisorS3Policy** policy.

  JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "s3:GetObject*",
                "s3:GetBucket*",
                "s3:List*",
                "s3:DeleteObject*",
                "s3:PutObject*"
            ],
            "Resource": [
                "arn:aws:s3:::bucket_name",
                "arn:aws:s3:::bucket_name/*"
            ]
        }
    ]
}
```

  In the preceding example, replace *bucket_name* with the name of the Amazon S3 bucket that you created in the prerequisites step.

We recommend that you create a non-administrative IAM user to use with Agentless Collector. When creating non-administrative IAM users, follow the security best practice Grant Least Privilege, granting users minimum permissions.

To create a non-administrator IAM user to use with Agentless Collector

1. In AWS Management Console, navigate to the IAM console, using the AWS account that you used to set the home Region in Sign in to the Migration Hub console and choose a home Region.
2. Create a non-administrator IAM user by following the instructions for creating a user with the console as described in Creating an IAM user in your AWS account in the *IAM User Guide*.

   While following the instructions in the *IAM User Guide*:

   - When on the step about selecting the type of access, select **Programmatic access**. Note, while not recommended, only select **AWS Management Console access** if you plan to use the same IAM user credentials for accessing the AWS console.
   - When on the step about the **Set permission** page, choose the option to **Attach existing policies to user directly**. Then select the `AWSApplicationDiscoveryAgentlessCollectorAccess` AWS managed policy from the list of policies.

     Next, select the `DMSCollectorPolicy` and `FleetAdvisorS3Policy` customer managed IAM policies.
   - When on the step about viewing the user's access keys (access key IDs and secret access keys), follow the guidance in the **Important** note about saving the user's new access key ID and secret access key in a safe and secure place. You'll need these access keys in Configuring Agentless Collector.

     It's an AWS security best practice to rotate access keys. For information about rotating keys, see Rotate access keys regularly for use cases that require long-term credentials in the *IAM User Guide.*

Download the Agentless Collector

To set up the Application Discovery Service Agentless Collector (Agentless Collector), you must download and deploy the Agentless Collector Open Virtualization Archive (OVA) file. The Agentless

Collector is a virtual appliance that you install in your on-premises VMware environment. This step describes how to download the collector OVA file and the next step describes how to deploy it.

To download the collector OVA file and verify its checksum

1. Sign in to vCenter as a VMware administrator and switch to the directory where you want to download the Agentless Collector OVA file.
2. Download the OVA file from the following URL:

   Agentless Collector OVA
3. Depending on which hashing algorithm you use in your system environment, download either the MD5 or SHA256 to get the file containing the checksum value. Use the downloaded value to verify the `ApplicationDiscoveryServiceAgentlessCollector` file downloaded in the preceding step.
4. Depending on your variation of Linux, run the version appropriate MD5 command or SHA256 command to verify that the cryptographic signature of the `ApplicationDiscoveryServiceAgentlessCollector.ova` file matches the value in the respective MD5/SHA256 file that you downloaded.

   ```
   $ md5sum ApplicationDiscoveryServiceAgentlessCollector.ova
   ```

   ```
   $ sha256sum ApplicationDiscoveryServiceAgentlessCollector.ova
   ```

Deploy Agentless Collector

Application Discovery Service Agentless Collector (Agentless Collector) is a virtual appliance that you install in your on-premises VMware environment. This section describes how to deploy the Open Virtualization Archive (OVA) file that you downloaded in your VMware environment.

Agentless Collector virtual machine specifications

Agentless Collector version 2

- **Operating System** – Amazon Linux 2023
- **RAM** – 16 GB
- **CPU** – 4 cores
- **VMware requirements** – See VMware host requirements for running AL2023 on VMware

Agentless Collector version 1

- **Operating System** – Amazon Linux 2
- **RAM** – 16 GB
- **CPU** – 4 cores

The following procedure steps you through deploying the Agentless Collector OVA file in your VMware environment.

To deploy Agentless Collector

1. Sign in to vCenter as a VMware administrator.
2. Use one of the following ways to install the OVA file:
   - Use the UI: Choose **File**, choose **Deploy OVF Template**, select the collector OVA file you downloaded in the previous section, and then complete the wizard. Ensure the proxy settings in the server management dashboard are configured correctly.
   - Use the command line: To install the collector OVA file from the command line, download and use the VMware Open Virtualization Format Tool (ovftool). To download ovftool, select a release from the OVF Tool Documentation page.

     The following is an example of using the ovftool command line tool to install the collector OVA file.

     ```
     ovftool --acceptAllEulas --name=AgentlessCollector --datastore=datastore1
      -dm=thin ApplicationDiscoveryServiceAgentlessCollector.ova
      'vi://username:password@vcenterurl/Datacenter/host/esxi/'
     ```

     The following describe the *replaceable* values in the example

     - The name is the name that you want to use for your Agentless Collector VM.
     - The datastore is the name of the datastore in your vCenter.
     - The OVA file name is the name of the downloaded collector OVA file.
     - The username/password are your vCenter credentials.
     - The vcenterurl is the URL of your vCenter.
     - The vi path is the path to your VMware ESXi host.

3. Locate the deployed Agentless Collector in your vCenter. Right-click the VM, and then choose **Power**, **Power On**.
4. After a few minutes, the IP address of the collector displays in vCenter. You use this IP address to connect to the collector.

Accessing the Agentless Collector console

The following procedure describes how to access the Application Discovery Service Agentless Collector (Agentless Collector) console.

To access the Agentless Collector console

1. Open a web browser, and then type the following URL in the address bar: **`https://`***`<ip_address>`***`/`**, where *`<ip_address>`* is the IP address of the collector from Deploy Agentless Collector.
2. Choose **Get Started** the first time you access Agentless Collector. Thereafter, you'll be asked to **Log in**.

If you're accessing the Agentless Collector console for the first time, next you'll Configuring Agentless Collector. Otherwise, next you'll see The Agentless Collector dashboard.

Configuring Agentless Collector

Application Discovery Service Agentless Collector (Agentless Collector) is an Amazon Linux 2 based virtual machine (VM). The following section describes how to configure a collector VM on the Agentless Collector console's **Configure Agentless Collector** page.

To configure a collector VM on the Configure Agentless Collector page

1. For **Collector name**, enter a name for the collector to identify it. The name can contain spaces but it cannot contain special characters.
2. Under **Data synchronization**, enter the AWS access key and secret key for the AWS account IAM user to specify as the destination account to receive the data discovered by the collector. For information about the requirements for the IAM user, see Deploying Application Discovery Service Agentless Collector.

a. For **AWS access-key**, enter the access key of the AWS account IAM user that you're specifying as the destination account.

b. For **AWS secret-key**, enter the secret key of the AWS account IAM user that you are you're specifying as the destination account.

c. (Optional) If your network requires the use of a proxy to access AWS, enter the proxy host, proxy port, and, optionally, the credentials needed to authenticate with your existing proxy server.

3. Under **Agentless Collector password**, set up a password to use to authenticate access to Agentless Collector.

- Passwords are case-sensitive
- Passwords must be between 8 and 64 characters in length
- Passwords must contain at least one character from each of the following four categories:
  - Lowercase letters (a-z)
  - Uppercase letters (A-Z)
  - Numbers (0-9)
  - Non-alphanumeric characters (@$!#%*?&)
- Passwords cannot contain special characters other than the following ones: @$!#%*?&

a. For **Agentless Collector password**, enter a password to use to authenticate access to the collector.

b. For **Re-enter Agentless Collector password**, for verification, enter the password again.

4. Under **Other settings**, read the **License Agreement**. If you agree to accept it, select the check box.

5. To enable automatic updates for Agentless Collector, under **Other settings**, select **Automatically update Agentless Collector**. If you do not select this checkbox, you'll need to manually update Agentless Collector as described in Manually updating Application Discovery Service Agentless Collector.

6. Choose **Save configurations**.

The following topics describe optional collector configuration tasks.

Optional Configuration Tasks

(Optional) Configure a static IP address for the Agentless Collector VM

The following steps describe how to configure a static IP address for the Application Discovery Service Agentless Collector (Agentless Collector) VM. When first installed, the collector VM is configured to use the Dynamic Host Configuration Protocol (DHCP).

> **Note**
>
> The Agentless Collector supports IPv4. It does not support IPv6.

Agentless Collector version 2

To configure a static IP address for the collector VM

1. Collect the following network information from VMware vCenter:
   - **Static IP address** – An unsigned IP address in the subnet. For example, 192.168.1.138.
   - **CIDR netmask** – To get the CIDR netmask, check the IP address setting of the VMware vCenter host that hosts the collector VM. For example, /24.
   - **Default Gateway** – To get the default gateway, check the IP address setting of the VMware vCenter host that hosts the collector VM. For example, 192.168.1.1.
   - **Primary DNS** – To get the primary DNS, check the IP address setting of the VMware vCenter host that hosts the collector VM. For example, 192.168.1.1.
   - (Optional) **Secondary DNS**
   - (Optional) **Local domain name** – This allows the collector to reach the vCenter host URL without the domain name.
2. Open the collector's VM console and sign in as **`ec2-user`** using the password **`collector`** as shown in the following example.

```
username: ec2-user
password: collector
```

3. Disable the network interface, by entering the following command in the remote terminal.

```
sudo ip link set ens192 down
```

4. Update the interface configuration by using the following steps.

   a. Open 10-cloud-init-ens192.network in the vi editor by using the following command.

   ```
   sudo vi /etc/systemd/network/10-cloud-init-ens192.network
   ```

   b. Update the values, as shown in the following example, with the information that you collected in the **Collect network information** step.

   ```
   [Match]
   Name=ens192

   [Network]
   DHCP=no
   Address=static-ip-value/CIDR-netmask
   Gateway=gateway-value
   DNS=dnsserver-value
   ```

5. Update the Domain Name System (DNS) using the following steps.

   a. Open the `resolv.conf` file in vi using the following command.

   ```
   sudo vi /etc/resolv.conf
   ```

   b. Update the `resolv.conf` file in vi using the following command.

   ```
   search localdomain-name
   options timeout:2 attempts:5
   nameserver dnsserver-value
   ```

   The following example shows an edited `resolv.conf` file.

   ```
   search vsphere.local
   options timeout:2 attempts:5
   nameserver 192.168.1.1
   ```

6. Enable the network interface, by entering the following command.

```
sudo ip link set ens192 up
```

7. Reboot the VM as shown in the following example.

```
sudo reboot
```

8. Verify your network settings using the following steps.

   a. Check if the IP address is configured correctly, by entering the following commands.

```
ifconfig
ip addr show
```

   b. Check that the gateway was added correctly, by entering the following command.

```
route -n
```

   The output should be similar to the following example.

```
Kernel IP routing table
Destination     Gateway         Genmask         Flags Metric Ref    Use
 Iface
0.0.0.0         192.168.1.1     0.0.0.0         UG    0      0        0 eth0
172.17.0.0      0.0.0.0         255.255.0.0     U     0      0        0
 docker0
192.168.1.0     0.0.0.0         255.255.255.0   U     0      0
```

   c. Verify that you can ping a public URL, by entering the following command.

```
ping www.google.com
```

   d. Verify that you can ping the vCenter IP address or host name as shown in the following example.

```
ping vcenter-host-url
```

Agentless Collector version 1

To configure a static IP address for the collector VM

1. Collect the following network information from VMware vCenter:

   - **Static IP address** – An unsigned IP address in the subnet. For example, 192.168.1.138.
   - **Network mask** – To get the network mask, check the IP address setting of the VMware vCenter host that hosts the collector VM. For example, 255.255.255.0.
   - **Default Gateway** – To get the default gateway, check the IP address setting of the VMware vCenter host that hosts the collector VM. For example, 192.168.1.1.
   - **Primary DNS** – To get the primary DNS, check the IP address setting of the VMware vCenter host that hosts the collector VM. For example, 192.168.1.1.
   - (Optional) **Secondary DNS**
   - (Optional) **Local domain name** – This allows the collector to reach the vCenter host URL without the domain name.

2. Open the collector's VM console and sign in as **`ec2-user`** using the password **`collector`** as shown in the following example.

   ```
   username: ec2-user
   password: collector
   ```

3. Disable the network interface, by entering the following command in the remote terminal.

   ```
   sudo /sbin/ifdown eth0
   ```

4. Update the interface eth0 configuration using the following steps.

   a. Open ifcfg-eth0 in the vi editor using the following command.

   ```
   sudo vi /etc/sysconfig/network-scripts/ifcfg-eth0
   ```

   b. Update the interface values, as shown in the following example, with the information that you collect in the **Collect network information** step.

   ```
   DEVICE=eth0
   BOOTPROTO=static
   ONBOOT=yes
   ```

```
IPADDR=static-ip-value
NETMASK=netmask-value
GATEWAY=gateway-value
TYPE=Ethernet
USERCTL=yes
PEERDNS=no
RES_OPTIONS="timeout:2 attempts:5"
```

5. Update the Domain Name System (DNS) using the following steps.

   a. Open the `resolv.conf` file in vi using the following command.

   ```
   sudo vi /etc/resolv.conf
   ```

   b. Update the `resolv.conf` file in vi using the following command.

   ```
   search localdomain-name
   options timeout:2 attempts:5
   nameserver dnsserver-value
   ```

   The following example shows an edited `resolv.conf` file.

   ```
   search vsphere.local
   options timeout:2 attempts:5
   nameserver 192.168.1.1
   ```

6. Enable the network interface, by entering the following command.

   ```
   sudo /sbin/ifup eth0
   ```

7. Reboot the VM as shown in the following example.

   ```
   sudo reboot
   ```

8. Verify your network settings using the following steps.

   a. Check if the IP address is configured correctly, by entering the following commands.

   ```
   ifconfig
   ip addr show
   ```

   b. Check that the gateway was added correctly, by entering the following command.

```
route -n
```

The output should be similar to the following example.

```
Kernel IP routing table
Destination     Gateway         Genmask         Flags Metric Ref    Use
 Iface
0.0.0.0         192.168.1.1     0.0.0.0         UG    0      0        0 eth0
172.17.0.0      0.0.0.0         255.255.0.0     U     0      0        0
 docker0
192.168.1.0     0.0.0.0         255.255.255.0   U     0      0
```

c. Verify that you can ping a public URL, by entering the following command.

```
ping www.google.com
```

d. Verify that you can ping the vCenter IP address or host name as shown in the following example.

```
ping vcenter-host-url
```

(Optional) Reset the Agentless Collector VM back to using DHCP

The following steps describe how to reconfigure the Agentless Collector VM to use DHCP.

Agentless Collector version 2

To configure the collector VM to use DHCP

1. Disable the network interface by running the following command in the remote terminal.

```
sudo ip link set ens192 down
```

2. Update the interface configuration by using the following steps.

   a. Open the `10-cloud-init-ens192.network` file in the vi editor by using the following command.

```
sudo vi /etc/systemd/network/10-cloud-init-ens192.network
```

b. Update the values as shown in the following example.

```
[Match]
Name=ens192

[Network]
DHCP=yes

[DHCP]
ClientIdentifier=mac
```

3. Reset the DNS setting, by entering the following command.

```
echo "" | sudo tee /etc/resolv.conf
```

4. Enable the network interface, by entering the following command.

```
sudo ip link set ens192 up
```

5. Reboot the collector VM as shown in the following example.

```
sudo reboot
```

Agentless Collector version 1

To configure the collector VM to use DHCP

1. Disable the network interface by running the following command in the remote terminal.

```
sudo /sbin/ifdown eth0
```

2. Update the network configuration by using the following steps.

a. Open the `ifcfg-eth0` file in the vi editor using the following command.

```
sudo /sbin/ifdown eth0
```

b. Update the values in the `ifcfg-eth0` file as shown in the following example.

```
DEVICE=eth0
BOOTPROTO=dhcp
ONBOOT=yes
TYPE=Ethernet
USERCTL=yes
PEERDNS=yes
DHCPV6C=yes
DHCPV6C_OPTIONS=-nw
PERSISTENT_DHCLIENT=yes
RES_OPTIONS="timeout:2 attempts:5"
```

3. Reset the DNS setting by entering the following command.

```
echo "" | sudo tee /etc/resolv.conf
```

4. Enable the network interface by entering the following command.

```
sudo /sbin/ifup eth0
```

5. Reboot the collector VM as shown in the following example.

```
sudo reboot
```

(Optional) Configure the Kerberos authentication protocol

If your OS server supports the Kerberos authentication protocol, then you can use this protocol to connect to your server. To do so, you must configure the Application Discovery Service Agentless Collector VM.

The following steps describe how to configure the Kerberos authentication protocol on your Application Discovery Service Agentless Collector VM.

To configure the Kerberos authentication protocol on your collector VM

1. Open the collector's VM console and sign in as **`ec2-user`** using the password **`collector`** as shown in the following example.

```
username: ec2-user
password: collector
```

2. Open the `krb5.conf` configuration file in the `/etc` folder. To do so, you can use the following code example.

```
cd /etc
sudo nano krb5.conf
```

3. Update the `krb5.conf` configuration file with the following information.

```
[libdefaults]
    forwardable = true
    dns_lookup_realm = true
    dns_lookup_kdc = true
    ticket_lifetime = 24h
    renew_lifetime = 7d
    default_realm = default_Kerberos_realm

[realms]
 default_Kerberos_realm = {
     kdc = KDC_hostname
     server_name = server_hostname
     default_domain = domain_to_expand_hostnames
 }

[domain_realm]
 .domain_name = default_Kerberos_realm
 domain_name = default_Kerberos_realm
```

   Save the file and exit the text editor.

4. Reboot the collector VM as shown in the following example.

```
sudo reboot
```

Using the Agentless Collector Network Data Collection module

The Network Data Collection module makes it possible for you to discover dependencies among servers in your on-premises data center. This network data accelerates your migration planning by providing visibility into how applications communicate across servers.

The Network Data Collection module connects to the servers that the VMware vCenter module identifies, and analyzes source IP to destination IP/port traffic for those servers.

Topics

Setting up the Network Data Collection module

The Network Data Collection module collects network data for the server inventory that comes from the VMware vCenter module. Therefore, to use the Network Data Collection module, first set up the VMware vCenter module. For instructions, follow the guidance in the following topics:

1. the section called “Deploying a collector”
2. the section called “Accessing the collector console”
3. the section called “Configuring the collector”
4. the section called “Using the VMware data collection module”

To set up the Network Data Collection module

1. On the Agentless Collector dashboard, in the **Network Data Collection** section, choose **View network connections**.
2. On the **Network connections** page, choose **Edit collector**.
3. In the credentials section, enter at least one set of credentials. You can enter up to 10 sets of credentials. The first time the module attempts to collect data for a server, it tries all of the credentials until it finds a set of credentials that works; it then saves that set and uses it again in subsequent attempts. For information about setting up credentials, see the section called “Setting up credentials”.
4. In the **Data collection preferences** section, to automatically start collecting data when a server reboots, select **Start data collection automatically**.
5. If you haven't set up WinRM certificates, select **Disable WinRM certificate checks**.
6. Choose **Save**.
7. Collection happens on the servers every 15 seconds. To see the details of the collection attempts for a given server, select the checkbox to the left of the server in the **Servers** table.

Setting up credentials

The Network Data Collection module uses WinRM to collect data from Windows servers. It uses SNMPv2 and SNMPv3 to collect data from Linux servers.

WinRM credentials:

- Specify the username and password of a Windows account that has the following:
  - Read access to the `\root\standardcimv2` namespace
  - Read permissions for `MSFT_NetTCPConnection` class
  - Remote WMI access
- We recommend that you create a dedicated service account with minimal required permissions.
- Avoid using domain administrator or local administrator accounts.
- Port 5986 (HTTPS) must be open between collector and target servers.
- Avoid disabling WinRM certificate check. For information about setting up WinRM certificates, see the section called “Addressing self-signed certification problems when configuring WinRM certificates”.

SNMPv2 credentials:

- Provide a read-only community string that can access 1.3.6.1.2.1.6.13.* OID
- SNMPv3 is preferable to SNMPv2 because of the improved security in SNMPv3
- Port 161/UDP must be open between collector and target servers
- Use complex, non-default community strings
- Avoid common strings like "public" or "private"
- Treat community strings like passwords

SNMPv3 credentials

- Provide a username/password and auth/privacy details with read-only permission that can access 1.3.6.1.2.1.6.13.* OID.
- Port 161/UDP must be open between collector and target servers
- Enable both authentication and privacy
- Use strong authentication protocols (SHA preferred over MD5)

- Use strong encryption protocols (AES preferred over DES)
- Use complex passwords for both auth and privacy
- Use unique usernames (avoid common names)

General best practices for Credential Management

- Store credentials securely
- Regularly rotate all credentials
- Use password managers or secure vaults
- Monitor credential usage
- Follow the principle of least privilege and only grant the minimum necessary permissions needed

Network data collection attempts

When a new server is discovered, the collector attempts each configured credential for each IP address. After the collector finds a valid credential, it only uses that credential. After two consecutive failures, the collector attempts to collect networking data for a server after 30 minutes, 2 hours, 8 hours, and then 24 hours. After 6 failed attempts, the collector continues to try all configured credentials once every day. To resolve the issue, either edit the current credentials or add additional ones by choosing **Edit collector**, or make changes to the target server being monitored.

Server status in the Network Data Collection module

The following table explains the collection status values.

Status	Meaning
Collecting or Collected	The last collection attempt for network connections was successful.
Erroring or Errored	The last collection attempt for network connections failed due to either a networkin g or permissions problem. For additional information, select the checkbox to the left of the server that has the error.

Status	Meaning
Skipped	Servers for which no valid credentials were provided. Update or configure additional server credentials.
No data	Data collection for the server has not started. To start collecting data, choose **Start collector** .
Pending	Collection has been started but no collectio n attempts have been made. Wait a few minutes, and then refresh the list.

Using the VMware vCenter Agentless Collector data collection module

This section describes the Application Discovery Service Agentless Collector (Agentless Collector) VMware vCenter data collection module, which is used to collect server inventory, profile, and utilization data from your VMware VMs.

Topics

Setting up the Agentless Collector data collection module for VMware vCenter

This section describes how to set up the Agentless Collector VMware vCenter data collection module to collect server inventory, profile, and utilization data from your VMware VMs.

> **Note**
>
> Before starting vCenter setup, make sure you can provide vCenter credentials with Read and View permissions set for the System group.

To set up the VMware vCenter data collection module

1. On the **Agentless Collector** dashboard page, under **Data collection**, choose **Set up** in the **VMware vCenter** section.
2. On the **Set up VMware vCenter data collection** page, perform the following:
   a. Under **vCenter credentials**:
      i. For **vCenter URL/IP**, enter the IP address of your VMware vCenter Server host.
      ii. For **vCenter Username**, enter the name of a local or domain user that the collector uses to communicate with vCenter. For domain users, use the form *domain**username* or *username@domain*.
      iii. For **vCenter Password**, enter the local or domain user password.
   b. Under **Data collection preferences**:
      - To automatically start collecting data immediately following a successful setup, select **Start data collection automatically**.
   c. Choose **Set up**.

Next, you'll see the **VMware data collection details** page, which is described in the next topic.

Viewing VMware data collection details

The **VMware data collection details** page shows details about the vCenter you set up in Setting up the Agentless Collector data collection module for VMware vCenter.

Under **Discovered vCenter servers**, the vCenter you set up is listed with the following information about the vCenter:

- The IP address of the vCenter server.
- The number of servers in the vCenter.

- The status of the data collection.
- How long since the last update.

Choose **Remove vCenter server** to remove the displayed vCenter server and return you to the **Set up VMware vCenter data collection** page.

If you did not choose to start data collection automatically, you can start data collection by using the **Start data collection** button on this page. After data collection starts, the start button changes to **Stop data collection**.

If the **Collection status** column shows **Collecting**, data collection has started.

You view the collected data in the AWS Migration Hub console. If you're collecting data for a VMware vCenter server inventory, you can access data that appears in the console approximately 15 minutes after turning on data collection.

You can choose **View servers in Migration Hub** on this page to open the Migration Hub console, if your access to the internet is not blocked. Whether you choose this button or not, for information about how to access the Migration Hub console, see Viewing your collected data.

The following are the guidelines for recommended length of data collection according to migration planning activities:

- TCO (total cost of ownership) - 2 to 4 weeks
- Migration planning - 2 to 6 weeks

Controlling the scope of vCenter data collection

The vCenter user requires read-only permissions on each ESX host or VM to inventory using Application Discovery Service. Using the permission settings, you can control which hosts and VMs are included in the data collection. You can either allow all hosts and VMs under the current vCenter to be inventoried, or grant permissions on a case-by-case basis.

> **Note**
>
> As a security best practice, we recommend against granting additional, unneeded permissions to the vCenter user of the Application Discovery Service.

The following procedures describe configuration scenarios ordered from least granular to most granular. These procedures are for vSphere Client v6.7.0.2. The procedures for other versions of the client might be different, depending on which version of the vSphere client you are using.

To discover data about *all* ESX hosts and VMs under the current vCenter

1. In your VMware vSphere client, choose **vCenter** and then choose either **Hosts and Clusters** or **VMs and Templates**.
2. Choose a datacenter resource and then choose **Permissions**.
3. Choose the vCenter user and then choose the symbol to add, edit, or remove a user role.
4. Choose **Read-only** from the **Role** menu.
5. Choose **Propagate to children** and then choose **OK**.

To discover data about a *specific* ESX host and *all* of its child objects

1. In your VMware vSphere client, choose **vCenter** and then choose either **Hosts and Clusters** or **VMs and Templates**.
2. Choose **Related Objects**, **Hosts**.
3. Open the context (right-click) menu for the host name and choose **All vCenter Actions**, **Add Permission**.
4. Under **Add Permission**, add the vCenter user to the host. For **Assigned Role**, choose **Read-only**.
5. Choose **Propagate to children**, **OK**.

To discover data about a *specific* ESX host or child VM

1. In your VMware vSphere client, choose **vCenter** and then choose either **Hosts and Clusters** or **VMs and Templates**.
2. Choose **Related Objects**.
3. Choose **Hosts** (showing a list of ESX hosts known to vCenter) or **Virtual Machines** (showing a list of VMs across all ESX hosts).
4. Open the context (right-click) menu for the host or VM name and choose **All vCenter Actions**, **Add Permission**.
5. Under **Add Permission**, add the vCenter user to the host or VM. For **Assigned Role**, choose **Read-only**, .

6. Choose **OK**.

> **Note**
>
> If you chose **Propagate to children**, you can still remove the read-only permission from ESX hosts and VMs on a case-by-case basis. This option has no effect on inherited permissions applying to other ESX hosts and VMs.

Data collected by the Agentless Collector VMware vCenter data collection module

The following information describes the data that's collected by the Application Discovery Service Agentless Collector (Agentless Collector) VMware vCenter data collection module. For information about setting up data collection, see Setting up the Agentless Collector data collection module for VMware vCenter.

Table legend for Agentless Collector VMware vCenter collected data:

- Collected data is in measurements of kilobytes (KB) unless stated otherwise.
- Equivalent data in the Migration Hub console is reported in megabytes (MB).
- Data fields denoted with an asterisk (*) are available only in the .csv files that are produced from the Application Discovery Service API export function.

  The Agentless Collector supports data export using the AWS CLI. To export collected data using the AWS CLI, follow the instructions described under **Export System Performance Data for All Servers** on the page Export Collected Data in the *Application Discovery Service User Guide*.
- The polling period is in intervals of approximately 60 minutes.
- Data fields denoted with a double asterisk (**) currently return a *null* value.

Data field	Description
applicationConfigurationId*	ID of the migration application the VM is grouped under.

Data field	Description
avgCpuUsagePct	Average percentage of CPU usage over polling period.
avgDiskBytesReadPerSecond	Average number of bytes read from disk over polling period.
avgDiskBytesWrittenPerSecond	Average number of bytes written to disk over polling period.
avgDiskReadOpsPerSecond**	Average number of read I/O operations per second null.
avgDiskWriteOpsPerSecond**	Average number of write I/O operations per second.
avgFreeRAM	Average free RAM expressed in MB.
avgNetworkBytesReadPerSecond	Average amount of throughput of bytes read per second.
avgNetworkBytesWrittenPerSecond	Average amount of throughput of bytes written per second.
computerManufacturer	Vendor reported by the ESXi host.
computerModel	Computer model reported by the ESXi host.
configId	ID assigned by Application Discovery Service to the discovered VM.
configType	Type of resource discovered.
connectorId	ID of the virtual appliance.
cpuType	vCPU for a VM, actual model for a host.
datacenterId	ID of the vCenter.
hostId*	ID of the VM host.

Data field	Description
hostName	Name of host running the virtualization software.
hypervisor	Type of hypervisor.
id	ID of server.
lastModifiedTimeStamp*	Latest date and time of data collection before data export.
macAddress	MAC address of the VM.
manufacturer	Maker of the virtualization software.
maxCpuUsagePct	Max. percentage of CPU usage during polling period.
maxDiskBytesReadPerSecond	Max. number of bytes read from disk over polling period.
maxDiskBytesWrittenPerSecond	Max. number of bytes written to disk over polling period.
maxDiskReadOpsPerSecond**	Max. number of read I/O operations per second.
maxDiskWriteOpsPerSecond**	Max. number of write I/O operations per second.
maxNetworkBytesReadPerSecond	Max. amount of throughput of bytes read per second.
maxNetworkBytesWrittenPerSecond	Max. amount of throughput of bytes written per second.
memoryReservation*	Limit to avoid overcommitment of memory on VM.

Data field	Description
moRefId	Unique vCenter Managed Object Reference ID.
name*	Name of VM or network (user specified).
numCores	Number of CPU cores assigned to VM.
numCpus	Number of CPU sockets on the ESXi host.
numDisks**	Number of disks on VM.
numNetworkCards**	Number of network cards on VM.
osName	Operating system name on VM.
osVersion	Operating system version on VM.
portGroupId*	ID of group of member ports of VLAN.
portGroupName*	Name of group of member ports of VLAN.
powerState*	Status of power.
serverId	Application Discovery Service assigned ID to the discovered VM.
smBiosId*	ID/version of the system management BIOS.
state*	Status of the virtual appliance.
toolsStatus	Operational state of VMware tools
totalDiskFreeSize	Free disk space expressed in MB. Available for vCenter Server 7.0 and later versions.
totalDiskSize	Total capacity of disk expressed in MB.
totalRAM	Total amount of RAM available on VM in MB.
type	Type of host.

Data field	Description
vCenterId	Unique ID number of a VM.
vCenterName*	Name of the vCenter host.
virtualSwitchName*	Name of the virtual switch.
vmFolderPath	Directory path of VM files.
vmName	Name of the virtual machine.

Using the database and analytics data collection module

This section describes how to set up, configure, and use a database and analytics data collection module. You can use this data collection module to connect to your data environment and collect metadata and performance metrics from your on-premises databases and analytics servers. For information about the metrics that you can collect with this module, see Data collected by the Agentless Collector database and analytics data collection module.

Important

End of support notice: On May 20, 2026, AWS will end support for AWS Database Migration Service Fleet Advisor. After May 20, 2026, you will no longer be able to access the AWS DMS Fleet Advisor console or AWS DMS Fleet Advisor resources. For more information, see AWS DMS Fleet Advisor end of support.

At a high level, when using the database and analytics data collection module, you take the following steps.

1. Complete the prerequisite steps, configure your IAM user, and create the AWS DMS data collector.
2. Configure data forwarding to make sure that your data collection module can send the collected metadata and performance metrics to AWS.

3. Add your LDAP servers and use them to discover OS servers in your data environment. Alternatively, add your OS servers manually or use the Using the VMware data collection module.
4. Configure connection credentials to your OS servers and then use them to discover database servers.
5. Configure connection credentials to your database and analytics servers and then run the data collection. For more information, see Database and analytics data collection.
6. View collected data in the AWS DMS console and use it to generate target recommendations for a migration to the AWS Cloud. For more information, see Database and analytics data collection.

Topics

Supported OS, database, and analytics servers

The database and analytics data collection module in the Agentless Collector supports Microsoft Active Directory LDAP servers.

This data collection module supports the following OS servers.

- Amazon Linux 2
- CentOS Linux version 6 and higher
- Debian version 10 and higher
- Red Hat Enterprise Linux version 7 and higher
- SUSE Linux Enterprise Server version 12 and higher
- Ubuntu version 16.01 and higher
- Windows Server 2012 and higher
- Windows XP and higher

Also, the database and analytics data collection module supports the following database servers.

- Microsoft SQL Server version 2012 and up to 2019
- MySQL version 5.6 and up to 8
- Oracle version 11g Release 2 and up to 12c, 19c, and 21c
- PostgreSQL version 9.6 and up to 13

Creating the AWS DMS data collector

Your database and analytics data collection module uses an AWS DMS data collector to interact with the AWS DMS console. You can view the collected data in the AWS DMS console, or use it to determine the right-sized AWS target engine. For more information, see Using the AWS DMS Fleet Advisor Target Recommendations feature.

Before you create an AWS DMS data collector, create an IAM role that your AWS DMS data collector uses to access your Amazon S3 bucket. You created this Amazon S3 bucket when you completed the prerequisites in Prerequisites for Agentless Collector.

To create an IAM role for your AWS DMS data collector to access Amazon S3

1. Sign in to the AWS Management Console and open the IAM console at https://console.aws.amazon.com/iam/.
2. In the navigation pane, choose **Roles**, then choose **Create role**.
3. On the **Select trusted entity** page, for **Trusted entity type**, choose **AWS Service**. For **Use cases for other AWS services**, choose **DMS**.
4. Select the **DMS** check box and choose **Next**.
5. On the **Add permissions** page, choose **FleetAdvisorS3Policy** that you created before. Choose **Next**.
6. On the **Name, review, and create** page, enter **`FleetAdvisorS3Role`** for **Role name**, then choose **Create role**.
7. Open the role that you created, and choose the **Trust relationships** tab. Choose **Edit trust policy**.
8. On the **Edit trust policy** page, paste the following JSON into the editor, replacing the existing code.

JSON

```
{
 "Version":"2012-10-17",
 "Statement": [{
  "Sid": "",
  "Effect": "Allow",
  "Principal": {
   "Service": [
    "dms.amazonaws.com",
    "dms-fleet-advisor.amazonaws.com"
   ]
  },
  "Action": "sts:AssumeRole"
 }]
}
```

9. Choose **Update policy**.

Now, create a data collector in the AWS DMS console.

To create an AWS DMS data collector

1. Sign in to the AWS Management Console and open the AWS DMS console at https://console.aws.amazon.com/dms/v2/.
2. Choose the AWS Region that you set as your Migration Hub home Region. For more information, see Sign in to Migration Hub and choose a home Region.
3. In the navigation pane, choose **Data collectors** under **Discover**. The **Data collectors** page opens.
4. Choose **Create data collector**. The **Create data collector** page opens.
5. For **Name** in the **General configuration** section, enter a name of your data collector.
6. In the **Connectivity** section, choose **Browse S3**. Choose the Amazon S3 bucket that you created before from the list.
7. For **IAM role**, choose `FleetAdvisorS3Role` that you created before.
8. Choose **Create data collector**.

Configuring data forwarding

After you create the required AWS resources, configure data forwarding from the database and analytics data collection module to your AWS DMS collector.

To configure data forwarding

1. Open the Agentless Collector console. For more information, see Accessing the collector console.
2. Choose **View Database and analytics collector**.
3. On the **Dashboard** page, choose **Configure data forwarding** in the **Data forwarding** section.
4. For **AWS Region**, **IAM access key ID**, and **IAM secret access key**, your Agentless Collector uses the values that you configured before. For more information, see Sign in to Migration Hub and choose a home Region and Deploying a collector.
5. For **Connected DMS data collector**, choose your data collector that you created in the AWS DMS console.
6. Choose **Save**.

After you configure data forwarding, check the **Data forwarding** section on the **Dashboard** page. Make sure that your database and analytics data collection module displays **Connected** for **Access to DMS** and **Access to S3**.

Adding your LDAP and OS servers

The database and analytics data collection module uses LDAP in Microsoft Active Directory to gather information about the OS, database, and analytics servers in your network. *Lightweight Directory Access Protocol (LDAP)* is an open standard application protocol. You can use this protocol to access and maintain distributed directory information services over your IP network.

You can add an existing LDAP server into your database and analytics data collection module to automatically discover OS servers in your network. If you don't use LDAP, you can add OS servers manually.

To add an LDAP server to your database and analytics data collection module

1. Open the Agentless Collector console. For more information, see Accessing the collector console.

2. Choose **View Database and analytics collector**, then choose **LDAP servers** under **Discovery** in the navigation pane.
3. Choose **Add LDAP server**. The **Add LDAP server** page opens.
4. For **Hostname**, enter the hostname of your LDAP server.
5. For **Port**, enter the port number that is used for LDAP requests.
6. For **User name**, enter the user name that you use to connect to your LDAP server.
7. For **Password**, enter the password that you use to connect to your LDAP server.
8. (Optional) Choose **Verify connection** to make sure that you added your LDAP server credentials correctly. Alternatively, you can verify your LDAP server connection credentials later, from the list on the **LDAP servers** page.
9. Choose **Add LDAP server**.
10. On the **LDAP servers** page, select your LDAP server from the list and choose **Discover OS servers**.

> **Important**
>
> For OS discovery, the data collection module needs credentials for the domain server to run requests using the LDAP protocol.

The database and analytics data collection module connects to your LDAP server and discovers your OS servers. After the data collection module completes the OS servers discovery, you can see the list of discovered OS servers by choosing **View OS servers**.

Alternatively, you can add your OS servers manually or import the list of servers from a comma-separated values (CSV) file. Also, you can use the VMware vCenter Agentless Collector data collection module to discover your OS servers. For more information, see Using the VMware data collection module.

To add an OS server to your database and analytics data collection module

1. On the **Database and analytics collector** page, choose **OS servers** under **Discovery** in the navigation pane.
2. Choose **Add OS server**. The **Add OS server** page opens.
3. Provide your OS server credentials.

a. For **OS type**, choose the operating system of your server.

b. For **Hostname / IP**, enter the hostname or IP address of your OS server.

c. For **Port**, enter the port number that is used for remote queries.

d. For **Authentication type**, choose the authentication type that your OS server uses.

e. For **User name**, enter the user name that you use to connect to your OS server.

f. For **Password**, enter the password that you use to connect to your OS server.

g. Choose **Verify** to make sure that you added your OS server credentials correctly.

4. (Optional) Add multiple OS servers from a CSV file.

a. Choose **Bulk import OS servers from CSV**.

b. Choose **Download template** to save a CSV file that includes a template that you can customize.

c. Enter the connection credentials for your OS servers into the file according to the template. The following example shows how you can provide OS server connection credentials in a CSV file.

```
OS type,Hostname/IP,Port,Authentication type,Username,Password
Linux,192.0.2.0,22,Key-based authentication,USER-EXAMPLE,ANPAJ2UCCR6DPCEXAMPLE
Windows,203.0.113.0,,NTLM,USER2-EXAMPLE,AKIAIOSFODNN7EXAMPLE
```

Save your CSV file after you add credentials for all your OS servers.

d. Choose **Browse**, then choose your CSV file.

5. Choose **Add OS server**.

6. After you add credentials for all OS servers, select your OS servers and choose **Discover database servers**.

Discovering your database servers

This section guides you through the steps you must take to configure your operating system and database servers. Then, you'll discover your servers and have the option to add a database or analytics server manually.

For database discovery, you must create users for your source databases with the minimum permissions required for the data collection module. For more information, see Creating database users for AWS DMS Fleet Advisor in the *AWS DMS User Guide*.

Configuring set up

To discover the databases running on the previously added OS Servers, the data collection module requires access to the operating system and database servers. This page outlines the steps you need to take to make sure that your database is accessible at the port that you specified in connection settings. You'll also turn on the remote authentication on your database server and provide your data collection module with permissions.

Configure set up on Linux

Complete the following procedure to configure set up to discover database servers on Linux.

To configure Linux to discover database servers

1. Provide sudo access to the `ss` and `netstat` commands.

   The following code example grants sudo access to the `ss` and `netstat` commands.

   ```
   sudo bash -c "cat << EOF >> /etc/sudoers.d/username
   username ALL=(ALL) NOPASSWD: /usr/bin/ss
   username ALL=(ALL) NOPASSWD: /usr/bin/netstat
   EOF"
   ```

   In the preceding example, replace *`username`* with the name of the Linux user that you specified in OS server connection credentials.

   The preceding example uses the `/usr/bin/` path to the `ss` and `netstat` commands. This path might be different in your environment. To determine the path to the `ss` and `netstat` commands, run the `which ss` and `which netstat` commands.
2. Configure your Linux servers to allow running remote SSH scripts and allow the Internet Control Message Protocol (ICMP) traffic.

Configure set up on Microsoft Windows

Complete the following procedure to configure set up to discover database servers on Microsoft Windows.

To configure Microsoft Windows to discover database servers

1. Provide credentials with grants to run Windows Management Instrumentation (WMI) and WMI Query Language (WQL) queries and read the registry.
2. Add the Windows user that you specified in OS server connection credentials to the following groups: Distributed COM Users, Performance Log Users, Performance Monitor Users, and Event Log Readers. To do so, use the following code example.

```
net localgroup "Distributed COM Users" username /ADD
net localgroup "Performance Log Users" username /ADD
net localgroup "Performance Monitor Users" username /ADD
net localgroup "Event Log Readers" username /ADD
```

   In the preceding example, replace *username* with the name of the Windows user that you specified in OS server connection credentials.
3. Grant the required permissions for the Windows user that you specified in OS server connection credentials.

   - For **Windows Management and Instrumentation Properties**, choose **Local Launch** and **Remote Activation**.
   - For **WMI Control**, choose the **Execute Methods**, **Enable Account**, **Remote Enable**, and **Read Security** permissions for the `CIMV2`, `DEFAULT`, `StandartCimv2`, and `WMI` namespaces.
   - For **WMI plug-in**, run `winrm configsddl default` and then choose **Read** and **Execute**.
4. Configure your Windows host by using the following code example.

```
netsh advfirewall firewall add rule name="Open Ports for WinRM incoming traffic"
 dir=in action=allow protocol=TCP localport=5985, 5986 # Opens ports for WinRM
netsh advfirewall firewall add rule name="All ICMP V4" protocol=icmpv4:any,any
 dir=in action=allow # Allows ICPM traffic

Enable-PSRemoting -Force # Enables WinRM
Set-Service WinRM -StartMode Automatic # Allows WinRM service to run on host
 startup
Set-Item WSMan:\localhost\Client\TrustedHosts -Value {IP} -Force # Sets the
 specific IP from which the access to WinRM is allowed

winrm set winrm/config/service '@{Negotiation="true"}' # Allow Negosiate auth usage
```

```
winrm set winrm/config/service '@{AllowUnencrypted="true"}' # Allow unencrypted
 connection
```

Discovering a database server

Complete the following set of tasks to discover and add database servers on the console.

To start the discovery of your database servers

1. On the **Database and analytics collector** page, choose **OS servers** under **Discovery** in the navigation pane.
2. Select the OS servers that include your database and analytics servers, then choose **Verify connection** on the **Actions** menu.
3. For servers that have the **Connectivity** status of **Failed**, edit the connection credentials.
   a. Select a single server or multiple servers when they have identical credentials, then choose **Edit** on the **Actions** menu. The **Edit OS server** page opens.
   b. For **Port**, enter the port number that is used for remote queries.
   c. For **Authentication type**, choose the authentication type that your OS server uses.
   d. For **User name**, enter the user name that you use to connect to your OS server.
   e. For **Password**, enter the password that you use to connect to your OS server.
   f. Choose **Verify connection** to make sure that you updated your OS server credentials correctly. Next, choose **Save**.
4. After you update credentials for all OS servers, select your OS servers and choose **Discover database servers**.

The database and analytics data collection module connects to your OS servers and discovers the supported database and analytics servers. After the data collection module completes the discovery, you can see the list of discovered database and analytics servers by choosing **View database servers**.

Alternatively, you can add your database and analytics servers to inventory manually. Also, you can import the list of servers from a CSV file. You can skip this step if you already added all your database and analytics servers to the inventory.

To add a database or analytics server manually

1. On the **Database and analytics collector** page, choose **Data collection** in the navigation pane.
2. Choose **Add database server**. The **Add database server** page opens.
3. Provide your database server credentials.

   a. For **Database engine**, choose the database engine of your server. For more information, see Supported OS, database, and analytics servers.

   b. For **Hostname / IP**, enter the hostname or IP address of your database or analytics server.

   c. For **Port**, enter the port where your server runs.

   d. For **Authentication type**, choose the authentication type that your database or analytics server uses.

   e. For **User name**, enter the user name that you use to connect to your server.

   f. For **Password**, enter the password that you use to connect to your server.

   g. Choose **Verify** to make sure that you added your database or analytics server credentials correctly.
4. (Optional) Add multiple servers from a CSV file.

   a. Choose **Bulk import database servers from CSV**.

   b. Choose **Download template** to save a CSV file that includes a template that you can customize.

   c. Enter the connection credentials for your database and analytics servers into the file according to the template. The following example shows how you can provide database or analytics server connection credentials in a CSV file.

```
Database engine,Hostname/IP,Port,Authentication type,Username,Password,Oracle
 service name,Database,Allow public key retrieval,Use SSL,Trust server
 certificate
Oracle,192.0.2.1,1521,Login/Password authentication,USER-
EXAMPLE,AKIAI44QH8DHBEXAMPLE,orcl,,,,
PostgreSQL,198.51.100.1,1533,Login/Password authentication,USER2-
EXAMPLE,bPxRfiCYEXAMPLE,,postgre,,TRUE,
MSSQL,203.0.113.1,1433,Login/Password authentication,USER3-
EXAMPLE,h3yCo8nvbEXAMPLE,,,,,TRUE
MySQL,2001:db8:4006:812:ffff:200e,8080,Login/Password authentication,USER4-
EXAMPLE,APKAEIVFHP46CEXAMPLE,,mysql,TRUE,TRUE,
```

Save your CSV file after you add credentials for all your database and analytics servers.

d. Choose **Browse**, then choose your CSV file.

5. Choose **Add database server**.
6. After you add credentials for all OS servers, select your OS servers and choose **Discover database servers**.

After you add all your database and analytics servers into the data collection module, add them to the inventory. The database and analytics data collection module can connect to the servers from the inventory and collects metadata and performance metrics.

To add your database and analytics servers to the inventory

1. On the **Database and analytics collector** page, choose **Database servers** under **Discovery** in the navigation pane.
2. Select the database and analytics servers, for which you want to collect metadata and performance metrics.
3. Choose **Add to inventory**.

After you add all database and analytics servers to your inventory, you can start collecting metadata and performance metrics. For more information, see Database and analytics data collection.

Data collected by the Agentless Collector database and analytics data collection module

The Application Discovery Service Agentless Collector (Agentless Collector) database and analytics data collection module collects the following metrics from your data environment. For information about setting up data collection, see Using the database and analytics data collection module.

When you use the database and analytics data collection module to collect **Metadata and database capacity**, it captures the following metrics.

- Available memory on your OS servers
- Available storage on your OS servers
- Database version and edition

- Number of CPUs on your OS servers
- Number of schemas
- Number of stored procedures
- Number of tables
- Number of triggers
- Number of views
- Schema structure

After you launch the schema analysis in the AWS DMS console, your data collection module analyzes and displays the following metrics.

- Database support dates
- Number of lines of code
- Schema complexity
- Similarity of schemas

When you use the database and analytics data collection module to collect **Metadata, database capacity, and resource utilization**, it captures the following metrics.

- I/O throughput on your database servers
- Input/output operations per second (IOPS) on your database servers
- Number of CPUs that your OS servers use
- Memory usage on your OS servers
- Storage usage on your OS servers

You can use the database and analytics data collection module to collect metadata, capacity, and utilization metrics from your Oracle and SQL Server databases. At the same time, for PostgreSQL and MySQL databases, the data collection module can collect only metadata.

Viewing your collected data

> **Important**
>
> End of support notice: On May 20, 2026, AWS will end support for AWS Database Migration Service Fleet Advisor. After May 20, 2026, you will no longer be able to access the AWS DMS Fleet Advisor console or AWS DMS Fleet Advisor resources. For more information, see AWS DMS Fleet Advisor end of support.

You can view the data that your Application Discovery Service Agentless Collector (Agentless Collector) collected in the Migration Hub console by following the steps in Viewing servers in the AWS Migration Hub console.

You can also view the collected metrics for database and analytics servers in the AWS DMS console by taking the following steps.

To view the data discovered by the database and analytics data collection module in the AWS DMS console

1. Sign in to the AWS Management Console and open the AWS DMS console at https://console.aws.amazon.com/dms/v2/.
2. Choose **Inventory** under **Discover**. The **Inventory** page opens.
3. Choose **Analyze inventories** to determine database schema properties, such as similarity and complexity.
4. Choose the **Schemas** tab to see the results of analysis.

You can use the AWS DMS console to identify duplicate schemas, determine the migration complexity, and export the inventory information for the future analysis. For more information, see Using inventories for analysis in AWS DMS Fleet Advisor.

Accessing the Agentless Collector

This section describes how to use the Application Discovery Service Agentless Collector (Agentless Collector).

Topics

The Agentless Collector dashboard

On the Application Discovery Service Agentless Collector (Agentless Collector) dashboard page you can see the status of the collector and choose a method of data collection as described in the following topics.

Topics

Collector status

Collector status gives you status information about the collector. The collector name, the status of the collector's connection to AWS, the Migration Hub home Region, and the version.

If you have AWS connection issues, you might need to edit Agentless Collector configuration settings.

To edit the collector configuration settings, choose **Edit collector settings** and follow the instructions described in Editing Agentless Collector settings.

Data collection

Under **Data collection** you can choose a data collection method. Application Discovery Service Agentless Collector (Agentless Collector) currently supports data collection from VMware VMs and from database and analytics servers. Future modules will support collection from additional virtualization platforms, and operating system level collection.

Topics

VMware vCenter data collection

To collect server inventory, profile, and utilization data from your VMware VMs, set up connections to your vCenter servers. To set up the connections, choose **Set up** in the **VMware vCenter** section and follow the instructions described in Using the VMware vCenter Agentless Collector data collection module.

After you set up vCenter data collection, from the dashboard you can perform the following:

- View data collection status
- Start data collection
- Stop data collection

> **Note**
>
> On the dashboard page, after you set up vCenter data collection, the **Set up** button in the **VMware vCenter** section is replaced with data collection status information, a **Stop data collection** button, and a **View and edit** button.

Database and analytics data collection

You can run your database and analytics data collection module in the following two modes.

Metadata and database capacity

The data collection module collects such information as schemas, versions, editions, CPU, memory, and disk capacity from your database and analytics servers. You can use this collected information to compute target recommendations in the AWS DMS console. If your source database is overprovisioned or underprovisioned, then the target recommendations also will be overprovisioned or underprovisioned.

This is the default mode.

Metadata, database capacity, and resource utilization

In addition to metadata and database capacity information, the data collection module collects actual utilization metrics of CPU, memory, and disk capacity for the databases and analytics servers. This mode provides more accurate target recommendations than the default mode

because the recommendations are based on the actual database workloads. In this mode, the data collection module collects performance metrics every minute.

To start collecting metadata and performance metrics from your database and analytics servers

1. On the **Database and analytics collector** page, choose **Data collection** in the navigation pane.
2. From the **Database inventory** list, select the database and analytics servers for which you want to collect metadata and performance metrics.
3. Choose **Run data collection**. The **Data collection type** dialog box opens.
4. Choose how to collect data for analysis.

   If you choose the **Metadata, database capacity, and resource utilization** option, then set the period of data collection. You can collect data during the **Next 7 days** or set the **Custom range** of 1–60 days.
5. Choose **Run data collection**. The **Data collection** page opens.
6. Choose the **Collection health** tab to see the status of data collection.

After completing the data collection, your data collection module uploads collected data to your Amazon S3 bucket. Then, you can view this collected data as described in Viewing your collected data.

Editing Agentless Collector settings

You configured the collector when you first set up Application Discovery Service Agentless Collector (Agentless Collector) as described in Configuring Agentless Collector. The following procedure describes how to edit Agentless Collector configuration settings.

To edit the collector configuration settings

- Choose the **Edit collector settings** button on the Agentless Collector dashboard.

  On the **Edit collector settings** page, perform the following:

  a. For **Collector name**, enter a name to identify the collector. The name can contain spaces but it cannot contain special characters.

  b. Under **Destination AWS account for discovery data**, enter the AWS access key and secret key for the AWS account to specify as the destination account to receive the data

discovered by the collector. For information about the requirements for the IAM user, see Deploying Application Discovery Service Agentless Collector.

i. For **AWS access-key**, enter the access key of the AWS account IAM user that you're specifying as the destination account.

ii. For **AWS secret-key**, enter the secret key of the AWS account IAM user that you're specifying as the destination account.

c. Under **Agentless Collector password**, change the password to use to authenticate access to the Agentless Collector.

i. For **Agentless Collector password**, enter a password to use to authenticate access to the Agentless Collector.

ii. For **Re-enter Agentless Collector password**, for verification enter the password again.

d. Choose **Save configurations**.

Next, you'll see The Agentless Collector dashboard.

Editing VMware vCenter credentials

To collect server inventory, profile, and utilization data from your VMware VMs, set up connections to your vCenter servers. For information about setting up VMware vCenter connections, see Using the VMware vCenter Agentless Collector data collection module.

This section describes how to edit the vCenter credentials.

> **Note**
>
> Before editing vCenter credentials, make sure you can provide vCenter credentials with Read and View permissions set for the System group.

To edit the VMware vCenter credentials

On the Viewing VMware data collection details page, choose **Edit vCenter servers**.

- On the **Edit vCenter** page, perform the following:

  a. Under **vCenter credentials**:

i. For **vCenter URL/IP**, enter the IP address of your VMware vCenter Server host.

ii. For **vCenter Username**, enter the name of a local or domain user that the connector uses to communicate with vCenter. For domain users, use the form *domain**username* or *username@domain*.

iii. For **vCenter Password**, enter the local or domain user password.

b. Choose **Save**.

Manually updating Application Discovery Service Agentless Collector

When you configure Application Discovery Service Agentless Collector (Agentless Collector), you can choose to enable automatic updates as described in Configuring Agentless Collector. If you do not enable automatic updates, you'll need to manually update Agentless Collector.

The following procedure describes how to manually update Agentless Collector.

To manually update Agentless Collector

1. Obtain the latest Agentless Collector Open Virtualization Archive (OVA) file.
2. (Optional) We recommend that you delete the previous Agentless Collector OVA file, before you deploy the latest one.
3. Follow steps in Deploy Agentless Collector.

The previous procedure only updates the Agentless Collector. It is your responsibility to keep the OS up to date.

To update your Amazon EC2 instance

1. Get the IP address of the Agentless Collector from VMware vCenter.
2. Open the collector's VM console and sign in as **`ec2-user`** using the password **`collector`** as shown in the following example.

```
username: ec2-user
password: collector
```

3. Follow the instructions in Update instance software on your AL2 instance in the Amazon Linux 2 User Guide.

Kernel Live Patching

Agentless Collector version 2

The Agentless Collector version 2 virtual machine uses Amazon Linux 2023 as described in Deploy Agentless Collector.

To enable and use Live Patching for Amazon Linux 2023, see Kernel Live Patching on AL2023 in the *Amazon EC2 User Guide*.

Agentless Collector version 1

The Agentless Collector version 1 virtual machine uses Amazon Linux 2 as described in Deploy Agentless Collector.

To enable and use Live Patching for Amazon Linux 2, see Kernel Live Patching on AL2 in the *Amazon EC2 User Guide*.

To upgrade from Agentless Collector version 1 to version 2

1. Install a new Agentless Collector OVA by using the latest image.
2. Set up credentials.
3. Delete the old virtual appliance.

Troubleshooting Agentless Collector

This section contains topics that can help you troubleshoot known issues with Application Discovery Service Agentless Collector (Agentless Collector).

Topics

- Fixing Unable to retrieve manifest or certificate file error
- Addressing self-signed certification problems when configuring WinRM certificates
- Fixing Agentless Collector cannot reach AWS during setup
- Fixing self-signed certification problems when connecting to the proxy host
- Finding unhealthy collectors

- Fixing IP address issues
- Fixing vCenter credentials issues
- Fixing data forwarding issues in the database and analytics data collection module
- Fixing connection issues in the database and analytics data collection module
- Standalone ESX host support
- Contacting AWS Support for Agentless Collector issues

Fixing `Unable to retrieve manifest or certificate file error`

If you receive this error when you try to deploy the OVA from the Amazon S3 URL in the VMware vCenter UI, ensure that your vCenter server meets the following requirements:

- VMware vCenter Server version 8.0 update 1 or later
- VMware vCenter Server 7.0 Update 3q (ISO Build 23788036) or later

Addressing self-signed certification problems when configuring WinRM certificates

If you enable WinRM certificate checks, you might need to import a self-signed certificate authority into the Agentless Collector.

To import a self-signed certificate authority

1. Open the collector's VM web console in VMware vCenter and sign in as `ec2-user` with the password `collector` as shown in the following example.

   ```
   username: ec2-user
   password: collector
   ```

2. Make sure that every self-signed CA certificate that is used to sign WinRM certificates is under the directory `/etc/pki/ca-trust/source/anchors`. For example:

   ```
   /etc/pki/ca-trust/source/anchors/https-winrm-ca-1.pem
   ```

3. To install the new certificates, run the following command.

```
sudo update-ca-trust
```

4. Restart the Agentless Collector by running the following command

```
sudo shutdown -r now
```

5. (Optional) To verify that certificates have been successfully imported, you can run the following command.

```
sudo trust list --filter=ca-anchors | less
```

Fixing Agentless Collector cannot reach AWS during setup

Agentless Collector requires outbound access over TCP port 443 to several AWS domains. When configuring Agentless Collector in the console you can get the following error message.

> **Could Not Reach AWS**
>
> AWS cannot be reached. Please verify network settings.

This error occurs because of a failed attempt by Agentless Collector to establish an HTTPS connection to an AWS domain that the collector needs to communicate with during the setup process. The Agentless Collector configuration fails if a connection can't be established.

To fix the connection to AWS

1. Check with your IT admin to see if your company firewall is blocking outbound traffic on port 443 to any of the AWS domains that require outbound access. Which AWS domains require outbound access depend on if your home Region is US West (Oregon) Region, us-west-2, or some other Region.

   The following domains require outbound access if your AWS account home Region is us-west-2:

   - `arsenal-discovery.us-west-2.amazonaws.com`
   - `migrationhub-config.us-west-2.amazonaws.com`

- `api.ecr-public.us-east-1.amazonaws.com`
- `public.ecr.aws`

The following domains require outbound access if your AWS account home Region is not `us-west-2`:

- `arsenal-discovery.us-west-2.amazonaws.com`
- `arsenal-discovery.`*`your-home-region`*`.amazonaws.com`
- `migrationhub-config.us-west-2.amazonaws.com`
- `api.ecr-public.us-east-1.amazonaws.com`
- `public.ecr.aws`

If your firewall is blocking outbound access to the AWS domains that Agentless Collector needs to communicate with, configure a proxy host in the **Data synchronization** section under **Collector configuration**.

2. If updating the firewall does not resolve the connection issue, use the following steps to ensure that the collector virtual machine has outbound network connectivity to the domains listed in the previous step.

   a. Get the IP address of the Agentless Collector from VMware vCenter.

   b. Open the collector's VM web console and sign in as **`ec2-user`** using the password **`collector`** as shown in the following example.

   ```
   username: ec2-user
   password: collector
   ```

   c. Test the connection to the listed domains by running telnet on ports 443 as shown in the following example.

   ```
   telnet migrationhub-config.us-west-2.amazonaws.com 443
   ```

3. If telnet cannot resolve the domain, try configuring a static DNS server using the instructions for Amazon Linux 2.
4. If the error continues, for further support, see Contacting AWS Support for Agentless Collector issues.

Fixing self-signed certification problems when connecting to the proxy host

If communication with the optionally provided proxy is via HTTPS and the proxy has a self-signed certificate, you might need to provide a certificate.

1. Get the IP address of the Agentless Collector from VMware vCenter.
2. Open the collector's VM web console and sign in as `ec2-user` with the password `collector` as shown in the following example.

   ```
   username: ec2-user
   password: collector
   ```

3. Paste the body of the certificate that is associated with the secure proxy, including both `-----BEGIN CERTIFICATE-----` and `-----END CERTIFICATE-----`, into the following file:

   ```
   /etc/pki/ca-trust/source/anchors/https-proxy-ca.pem
   ```

4. To install the new certificate, run the following commands:

   ```
   sudo update-ca-trust
   ```

5. Restart the Agentless Collector by running the following command:

   ```
   sudo shutdown -r now
   ```

Finding unhealthy collectors

Status information for every collector is found on the Data collectors page of the AWS Migration Hub (Migration Hub) console. You can identify collectors with problems by finding any collectors with a **Status** of **Requires attention**.

The following procedure describes how to access the Agentless Collector console to identify health issues.

To access the Agentless Collector console

1. Using your AWS account, sign in to the AWS Management Console and open the Migration Hub console at https://console.aws.amazon.com/migrationhub/.

2. In the Migration Hub console navigation pane under **Discover**, choose **Data collectors**.
3. From the **Agentless collectors** tab, make a note of the **IP address** for each connector that has a status of **Requires attention**.
4. To open the Agentless Collector console, open a web browser. Then type the following URL in the address bar: **`https://`***`<ip_address>`***`/`**, where *ip_address* is the IP address of an unhealthy collector.
5. Choose **Log in**, and then enter the Agentless Collector password, which was set up when the collector was configured in Configuring Agentless Collector.
6. On the **Agentless Collector** dashboard page, under **Data collection**, choose **View and edit** in the **VMware vCenter** section.
7. Follow the instructions in Editing VMware vCenter credentials to correct the URL and credentials.

After correcting the health issues, the collector will re-establish connectivity with vCenter server, and the collector's status will change to the **Collecting** state. If the issues persist, see Contacting AWS Support for Agentless Collector issues.

The most common causes for unhealthy collectors are IP address and credentials issues. Fixing IP address issues and Fixing vCenter credentials issues can help you resolve these issues and return a collector to a healthy state.

Fixing IP address issues

A collector can go into an unhealthy state if the vCenter endpoint provided during collector setup is malformed, invalid, or if the vCenter server is currently down and not reachable. In this case, you'll receive a **Connection error** message .

The following procedure can help you resolve IP address issues.

To fix collector IP address issues

1. Get the IP address of the Agentless Collector from VMware vCenter.
2. Open the Agentless Collector console by opening a web browser, and then type the following URL in the address bar: **`https://`***`<ip_address>`***`/`**, where *ip_address* is the IP address of the collector from Deploy Agentless Collector.
3. Choose **Log in**, and then enter the Agentless Collector password, which was set up when the collector was configured in Configuring Agentless Collector.

4. On the **Agentless Collector** dashboard page, under **Data collection**, choose **View and edit** in the **VMware vCenter** section.
5. On the **VMware data collection details** page, under **Discovered vCenter servers**, make a note of the IP address in the **vCenter** column.
6. Using a separate command line tool like `ping` or `traceroute`, validate that the associated vCenter server is active and the IP is reachable from the collector VM.

   - If the IP address is incorrect and the vCenter service is active, then update the IP address in the collector console, and choose **Next**.
   - If the IP address is correct but the vCenter server is inactive, activate it.
   - If the IP address is correct and the vCenter server is active, check if it is blocking ingress network connections due to firewall issues. If yes, update your firewall settings to allow incoming connections from the collector VM.

Fixing vCenter credentials issues

Collectors can go into an unhealthy state if the vCenter user credentials provided when configuring a collector are invalid, or do not have vCenter Read and View account privileges.

If you experience issues related to vCenter credentials, check to make sure that you have vCenter Read and View permissions set for the System group.

For information about editing vCenter credentials, see Editing VMware vCenter credentials.

Fixing data forwarding issues in the database and analytics data collection module

The home page of the database and analytics data collection module in Agentless Collector displays the connection status for **Access to DMS** and **Access to S3**. If you see **No access** for **Access to DMS** and **Access to S3**, then configure data forwarding. For more information, see Configuring data forwarding.

If you experience this issue after you configure data forwarding, then check to make sure that your data collection module can access to the internet. Then, make sure that you added the **DMSCollectorPolicy** and **FleetAdvisorS3Policy** policies to your IAM user. For more information, see Deploying Application Discovery Service Agentless Collector.

If your data collection module can't connect to AWS, then provide outbound access to the following domains.

- `dms.your-home-region.amazonaws.com`
- `s3.amazonaws.com`

Fixing connection issues in the database and analytics data collection module

The database and analytics data collection module in Agentless Collector connects to your LDAP servers to discover OS servers in your data environment. Then, the data collection module connects to your OS servers to discover database and analytics servers. From these database servers, the data collection module gathers capacity and performance metrics. If your data collection module can't connect to these servers, then verify that you can connect to your servers.

In the following examples, replace *replaceable* values with your values.

- To verify that you can connect to your LDAP server, install the `ldap-util` package. To do so, run the following command.

  ```
  sudo apt-get install ldap-util
  ```

  Then, run the following command.

  ```
  ldapsearch -x -D "CN=user,CN=Users,DC=example,DC=com" -w "password" -b
   "dc=example,dc=com" -h
  ```

- To verify that you can connect to a Linux OS server, use the following commands.

  ```
  ssh -i C:\Users\user\private_key.pem -p 22 username@my-linux-host.domain.com
  ```

  Run the previous example as an administrator in Windows.

  ```
  ssh username@my-linux-host.domain.com
  ```

  Run the previous example in Linux.

- To verify that you can connect to a Windows OS server, use the following commands.

```
winrs -r:[hostname or ip] -u:username -p:password cmd
```

Run the previous example as an administrator in Windows.

```
sudo apt install -y winrm
winrm --user=username --password=password [http or https]://[hostname or ip]:[port]
 "[cmd.exe or any other CLI command]"
```

Run the previous example in Linux.

- To verify that you can connect to a SQL Server database, use the following commands.

```
sqlcmd -S [hostname or IP] -U username -P 'password'
SELECT GETDATE() AS sysdate
```

- To verify that you can connect to a MySQL database, use the following commands.

```
mysql -u username -p 'password' -h [hostname or IP] -P [port]
SELECT NOW() FROM DUAL
```

- To verify that you can connect to an Oracle database, use the following commands.

```
sqlplus username/password@[hostname or IP]:port/servicename
SELECT SYSDATE FROM DUAL
```

- To verify that you can connect to a PostgreSQL database, use the following commands.

```
psql -U username -h [hostname or IP] -p port -d database
SELECT CURRENT_TIMESTAMP AS sysdate
```

If you can't connect to your database and analytics servers, then make sure that you provide the required permissions. For more information, see Discovering your database servers.

Standalone ESX host support

The Agentless Collector does not support a standalone ESX host. The ESX host must be part of the vCenter Server instance.

Contacting AWS Support for Agentless Collector issues

If you encounter issues with Application Discovery Service Agentless Collector (Agentless Collector) and need help, contact AWS Support You'll be contacted and might be asked to send the collector logs.

To obtain Agentless Collector logs

1. Get the IP address of the Agentless Collector from VMware vCenter.
2. Open the collector's VM web console and sign in as **`ec2-user`** using the password **`collector`** as shown in the following example.

```
username: ec2-user
password: collector
```

3. Use the following command to navigate to the log folder.

```
cd /var/log/aws/collector
```

4. Zip the log files by using the following commands.

```
sudo cp /local/agentless_collector/compose.log .
docker inspect $(docker ps --format {{.Names}}) | sudo tee docker_inspect.log >/
dev/null
sudo tar czf logs_$(date '+%d-%m-%Y_%H.%M.%S').tar.gz --exclude='db.mv*' *
```

5. Copy the log file from the Agentless Collector VM.

```
scp logs*.tar.gz targetuser@targetaddress
```

6. Give the `tar.gz` file to AWS Enterprise Support.

Importing data into Migration Hub

AWS Migration Hub (Migration Hub) import allows you to import details of your on-premises environment directly into Migration Hub without using the Application Discovery Service Agentless Collector (Agentless Collector) or AWS Application Discovery Agent (Discovery Agent), so you can perform migration assessment and planning directly from your imported data. You also can group your devices as applications and track their migration status.

This page describes the steps to complete an import request. First, you use one of the following two options to prepare your on-premises server data.

- Use common third-party tools to generate a file that contains your on-premises server data.
- Download our comma separated value (CSV) import template, and populate it with your on-premises server data.

After you use one of the two previously described methods to create your on-premises data file, you upload the file to Migration Hub by using the Migration Hub console, AWS CLI, or one of the AWS SDKs. For more information about the two options, see the section called “Supported import formats”.

You can submit multiple import requests. Each request is processed sequentially. You can check the status of your import requests at any time, through the console or import APIs.

After an import request is complete, you can view the details of individual imported records. View utilization data, tags, and application mappings directly from within the Migration Hub console. If errors were encountered during the import, you can review the count of successful and failed records, and you can see the error details for each failed record.

Handling errors: A link is provided to download the error log and failed records files as CSV files in a compressed archive. Use these files to resubmit your import request after correcting the errors.

Limits apply to the number of imported records, imported servers, and deleted records you can keep. For more information, see AWS Application Discovery Service Quotas.

Supported import formats

Migration Hub supports the following import formats.

RVTools

Migration Hub supports importing exports of VMware vSphere via RVTools. When saving data from RVTools, first choose the **Export all to csv** option or the **Export all to Excel** option, then ZIP the folder, and import the ZIP file into Migration Hub. The following files are required in the ZIP: vInfo, vNetwork, vCpu, vMemory, vDisk, vPartition, vSource, vTools, vHost, vNic, vSC_VMK.

Migration Hub import template

Migration Hub import allows you to import data from any source. The data provided must be in the supported format for a CSV file, and the data must contain only the supported fields with the supported ranges for those fields.

An asterisk (*) next to an import field name in the following table denotes that it is a required field. Each record of your import file must have at least one or more of those required fields populated to uniquely identify a server or application. Otherwise, a record without any of the required fields will fail to be imported.

A caret (^) next to an import filed name in the following table denotes that it is a readonly if a serverId is provided.

> **Note**
>
> If you're using either VMware.MoRefId or VMWare.VCenterId, to identify a record, you must have both fields in the same record.

Import Field Name	Description	Examples
ExternalId*^	A custom identifier that allows you to mark each record as unique. For example, **ExternalId** can be the inventory ID for the server in your data center.	Inventory Id 1 Server 2 CMBD Id 3

Import Field Name	Description	Examples
SMBiosId^	System management BIOS (SMBIOS) ID.	
IPAddress*^	A comma-delimited list of IP addresses of the server, in quotes.	192.0.0.2 "10.12.31.233, 10.12.32.11"
MACAddress*^	A comma-delimited list of MAC address of the server, in quotes.	00:1B:44:11:3A:B7 "00-15-E9-2B-99-3C, 00-14-22-01-23-45"
HostName*^	The host name of the server. We recommend using the fully qualified domain name (FQDN) for this value.	ip-1-2-3-4 localhost.domain
VMware.MoRefId*^	The managed object reference ID. Must be provided with a VMware.VC enterId.	
VMware.VCenterId*^	Virtual machine unique identifier. Must be provided with a VMware.MoRefId.	
CPU.NumberOfProcessors^	The number of CPUs.	4
CPU.NumberOfCores^	The total number of physical cores.	8

Import Field Name	Description	Examples
CPU.NumberOfLogicalCores^	The total number of threads that can run concurren tly on all CPUs in a server. Some CPUs support multiple threads to run concurrently on a single CPU core. In those cases, this number will be larger than the number of physical (or virtual) cores.	16
OS.Name^	The name of the operating system.	Linux Windows.Hat
OS.Version^	The version of the operating system.	16.04.3 NT 6.2.8
VMware.VMName^	The name of the virtual machine.	Corp1
RAM.TotalSizeInMB^	The total RAM available on the server, in MB.	64 128
RAM.UsedSizeInMB.Avg^	The average amount of used RAM on the server, in MB.	64 128
RAM.UsedSizeInMB.Max^	The maximum amount of used RAM available on the server, in MB.	64 128
CPU.UsagePct.Avg^	The average CPU utilization when the discovery tool was collecting data.	45 23.9

Import Field Name	Description	Examples
CPU.UsagePct.Max^	The maximum CPU utilization when the discovery tool was collecting data.	55.34 24
DiskReadsPerSecondInKB.Avg^	The average number of disk reads per second, in KB.	1159 84506
DiskWritesPerSecondInKB.Avg^	The average number of disk writes per second, in KB.	199 6197
DiskReadsPerSecondInKB.Max^	The maximum number of disk reads per second, in KB.	37892 869962
DiskWritesPerSecondInKB.Max^	The maximum number of disk writes per second, in KB.	18436 1808
DiskReadsOpsPerSecond.Avg^	The average number of disk read operations per second.	45 28
DiskWritesOpsPerSecond.Avg^	The average number of disk write operations per second.	8 3
DiskReadsOpsPerSecond.Max^	The maximum number of disk read operations per second.	1083 176
DiskWritesOpsPerSecond.Max^	The maximum number of disk write operations per second.	535 71
NetworkReadsPerSecondInKB.Avg^	The average number of network read operations per second, in KB.	45 28

Import Field Name	Description	Examples
NetworkWritesPerSecondInKB.Avg^	The average number of network write operations per second, in KB.	8 3
NetworkReadsPerSecondInKB.Max^	The maximum number of network read operations per second, in KB.	1083 176
NetworkWritesPerSecondInKB.Max^	The maximum number of network write operations per second, in KB.	535 71
Applications	A comma-delimited list of applications that include this server, in quotes. This value can include existing applications and/or new applications that are created upon import.	Application1 "Application2, Application3"
ApplicationWave	The migration wave for this server.	
Tags^	A comma-delimited list of tags formatted as name:value. **Important** Do not store sensitive information (like personal data) in tags.	"zone:1, critical:yes" "zone:3, critical:no, zone:1"
ServerId	The server identifier as seen in the Migration Hub server list.	`d-server-01kk9i6ywwaxmp`

You can import data even if you don't have data populated for all the fields defined in the import template, so long as each record has at least one of the required fields within it. Duplicates are managed across multiple import requests by using either an external or internal matching key. If you populate your own matching key, `External ID`, this field is used to uniquely identify and import the records. If no matching key is specified, import uses an internally generated matching key that is derived from some of the columns in the import template. For more information on this matching, see Matching logic for discovered servers and applications.

> **Note**
>
> Migration Hub import does not support any fields outside of those defined in the import template. Any custom fields supplied will be ignored and will not be imported.

Setting up import permissions

Before you can import your data, ensure that your IAM user has the necessary Amazon S3 permissions to upload (`s3:PutObject`) your import file to Amazon S3, and to read the object (`s3:GetObject`). You also must establish programmatic access (for the AWS CLI) or console access, by creating an IAM policy and attaching it to the IAM user that performs imports in your AWS account.

Console Permissions

Use the following procedure to edit the permissions policy for the IAM user that will make import requests in your AWS account using the console.

To edit a user's attached managed policies

1. Sign in to the AWS Management Console and open the IAM console at https://console.aws.amazon.com/iam/.
2. In the navigation pane, choose **Users**.
3. Choose the name of the user whose permissions policy you want to change.
4. Choose the **Permissions** tab and choose **Add permissions**.
5. Choose **Attach existing policies directly**, and then choose **Create policy**.

a. In the **Create policy** page that opens, choose **JSON**, and paste in the following policy. Remember to replace the name of your bucket with the actual name of the bucket that the IAM user will upload the import files into.

JSON

```
{
  "Version":"2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "s3:GetBucketLocation",
        "s3:ListAllMyBuckets"
      ],
      "Resource": "*"
    },
    {
      "Effect": "Allow",
      "Action": ["s3:ListBucket"],
      "Resource": ["arn:aws:s3:::importBucket"]
    },
    {
      "Effect": "Allow",
      "Action": [
        "s3:PutObject",
        "s3:GetObject",
        "s3:DeleteObject"
      ],
      "Resource": ["arn:aws:s3:::importBucket/*"]
    }
  ]
}
```

b. Choose **Review policy**.

c. Give your policy a new **Name** and optional description, before reviewing the summary of the policy.

d. Choose **Create policy**.

6. Return to the **Grant permissions** IAM console page for the user that will make import requests in your AWS account.

7. Refresh the table of policies, and search for the name of the policy you just created.
8. Choose **Next: Review**.
9. Choose **Add permissions**.

Now that you've added the policy to your IAM user, you're ready to start the import process.

AWS CLI Permissions

Use the following procedure to create the managed policies necessary to give an IAM user the permissions to make import data requests using the AWS CLI.

To create and attach the managed policies

1. Use the `aws iam create-policy` AWS CLI command to create an IAM policy with the following permissions. Remember to replace the name of your bucket with the actual name of the bucket that the IAM user will upload the import files into.

   JSON

```
{
  "Version":"2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": ["s3:ListBucket"],
      "Resource": ["arn:aws:s3:::importBucket"]
    },
    {
      "Effect": "Allow",
      "Action": [
        "s3:PutObject",
        "s3:GetObject",
        "s3:DeleteObject"
      ],
      "Resource": ["arn:aws:s3:::importBucket/*"]
    }
  ]
}
```

For more information on using this command, see create-policy in the *AWS CLI Command Reference*.

2. Use the `aws iam create-policy` AWS CLI command to create an additional IAM policy with the following permissions.

   JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "discovery:ListConfigurations",
                "discovery:CreateApplication",
                "discovery:UpdateApplication",
                "discovery:AssociateConfigurationItemsToApplication",
                "discovery:DisassociateConfigurationItemsFromApplication",
                "discovery:GetDiscoverySummary",
                "discovery:StartImportTask",
                "discovery:DescribeImportTasks",
                "discovery:BatchDeleteImportData"
            ],
            "Resource": "*"
        }
    ]
}
```

3. Use the `aws iam attach-user-policy` AWS CLI command to attach the policies you created in the previous two steps to the IAM user that will be performing import requests in your AWS account using the AWS CLI. For more information on using this command, see attach-user-policy in the *AWS CLI Command Reference*.

Now that you've added the policies to your IAM user, you're ready to start the import process.

Remember that when the IAM user uploads objects to the Amazon S3 bucket that you specified, they must leave the default permissions for the objects set so that the user can read the object.

Uploading your import file to Amazon S3

Next, you must upload your CSV formatted import file into Amazon S3 so it can be imported. Before you begin, you should have an Amazon S3 bucket that will house your import file created and/or chosen ahead of time.

Console S3 Upload

To upload your import file to Amazon S3

1. Sign in to the AWS Management Console and open the Amazon S3 console at https://console.aws.amazon.com/s3/.
2. In the **Bucket name** list, choose the name of the bucket that you want to upload your object to.
3. Choose **Upload**.
4. In the **Upload** dialog box, choose **Add files** to choose the file to upload.
5. Choose a file to upload, and then choose **Open.**
6. Choose **Upload**.
7. Once your file has been uploaded, choose the name of your data file object from your bucket dashboard.
8. From the **Overview** tab of the object details page, copy the **Object URL**. You'll need this when you create your import request.
9. Go to the **Import** page in the Migration Hub console as described in Importing data. Then, paste the object URL in the **Amazon S3 Object URL** field.

AWS CLI S3 Upload

To upload your import file to Amazon S3

1. Open a terminal window and navigate to the directory that your import file is saved to.
2. Enter the following command:

```
aws s3 cp ImportFile.csv s3://BucketName/ImportFile.csv
```

3. This returns the following results:

```
upload: .\ImportFile.csv to s3://BucketName/ImportFile.csv
```

4. Copy the full Amazon S3 object path that was returned. You will need this when you create your import request.

Importing data

After you download the import template from the Migration Hub console and populate it with your existing on-premises server data, you're ready to start importing the data into Migration Hub. The following instructions describe two ways to do this, either by using the console or by making API calls through the AWS CLI.

Console Import

Start data import on the **Tools** page of the Migration Hub console.

To start data import

1. In the navigation pane, under **Discover**, choose **Tools**.
2. If you don't already have an import template filled out, you can download the template by choosing **import template** in the **Import** box. Open the downloaded template and populate it with your existing on-premises server data. You can also download the import template from our Amazon S3 bucket at https://s3.us-west-2.amazonaws.com/templates-7cffcf56-bd96-4b1c-b45b-a5b42f282e46/import_template.csv
3. To open the **Import** page, choose **Import** in the **Import** box.
4. Under **Import name**, specify a name for the import.
5. Fill out the **Amazon S3 Object URL** field. To do this step, you'll need to upload your import data file to Amazon S3. For more information, see Uploading your import file to Amazon S3.
6. Choose **Import** in the lower-right area. This will open the **Imports** page where you can see your import and its status listed in the table.

After following the preceding procedure to start your data import, the **Imports** page will show details of each import request including its progress status, completion time, and the number of successful or failed records with the ability to download those records. From this screen, you can also navigate to the **Servers** page under **Discover** to see the actual imported data.

On the **Servers** page, you can see a list of all the servers (devices) that are discovered along with the import name. When you navigate from the **Imports** (import history) page by selecting the name of the import listed in the **Name** column, you are taken to the **Servers** page where a filter is applied based on the selected import's data set. Then, you only see data belonging to that particular import.

The archive is in a .zip format, and contains two files; `errors-file` and `failed-entries-file`. The errors file contains a list of error messages associated with each failed line and associated column name from your data file that failed the import. You can use this file to quickly identify where problems occurred. The failed entries file includes each line and all the provided columns that failed. You can make the changes called out in the errors file in this file and attempt to import the file again with the corrected information.

AWS CLI Import

To start the data import process from the AWS CLI, the AWS CLI must first be installed in your environment. For more information, see Installing the AWS Command Line Interface in the *AWS Command Line Interface User Guide*.

> **Note**
>
> If you don't already have an import template filled out, you can download the import template from our Amazon S3 bucket here: https://s3.us-west-2.amazonaws.com/templates-7cffcf56-bd96-4b1c-b45b-a5b42f282e46/import_template.csv

To start data import

1. Open a terminal window, and type the following command:

```
aws discovery start-import-task --import-url s3://BucketName/ImportFile.csv --
name ImportName
```

2. This will create your import task, and return the following status information:

```
{
    "task": {
        "status": "IMPORT_IN_PROGRESS",
        "applicationImportSuccess": 0,
        "serverImportFailure": 0,
```

```
        "serverImportSuccess": 0,
        "name": "ImportName",
        "importRequestTime": 1547682819.801,
        "applicationImportFailure": 0,
        "clientRequestToken": "EXAMPLE1-abcd-1234-abcd-EXAMPLE1234",
        "importUrl": "s3://BucketName/ImportFile.csv",
        "importTaskId": "import-task-EXAMPLE1229949eabfEXAMPLE03862c0"
    }
}
```

Tracking your Migration Hub import requests

You can track the status of your Migration Hub import requests using the console, AWS CLI, or one of the AWS SDKs.

Console Tracking

From the **Imports** dashboard in the Migration Hub console, you'll find the following elements.

- **Name** – The name of the import request.
- **Import ID** – The unique ID of the import request.
- **Import time** – The date and time that the import request was created.
- **Import status** – The status for the import request. This can be one of the following values:
  - **Importing** – This data file is currently being imported.
  - **Imported** – The entire data file was successfully imported.
  - **Imported with errors** – One or more of the records in the data file failed to import. To resolve your failed records, choose **Download failed records** for your import task and resolve the errors in the failed entries csv file, and do the import again.
  - **Import Failed** – None of the records in the data file where imported. To resolve your failed records, choose **Download failed records** for your import task and resolve the errors in the failed entries csv file, and do the import again.
- **Imported records** – The number of records in a specific data file that were successfully imported.
- **Failed records** – The number records in a specific data file that weren't imported.

CLI Tracking

You can track the status of your import tasks with the `aws discovery describe-import-tasks` AWS CLI command.

1. Open a terminal window, and type the following command:

```
aws discovery describe-import-tasks
```

2. This will return a list of all your import tasks in JSON format, complete with status and other relevant information. Optionally, you can filter results to return a subset of your import tasks.

When tracking your import tasks, you may find that the `serverImportFailure` value returned is greater than zero. When this happens, your import file had one or more entries that couldn't be imported. This can be resolved by downloading your failed records archive, reviewing the files within, and doing another import request with the modified failed-entries.csv file.

After creating your import task, you can perform additional actions to help manage and track your data migration. For example, you can download an archive of failed records for a specific request. For information on using the failed records archive to resolve import issues, see Troubleshooting failed import records.

View and explore discovered data

Both Application Discovery Service Agentless Collector (Agentless Collector) and AWS Discovery Agent (Discovery Agent) provide system performance data based on average and peak utilization. You can use the system performance data that's collected to perform a high-level total cost of ownership (TCO). Discovery Agents collect more detailed data including time series data for system performance information, inbound and outbound network connections, and processes running on the server. You can use this data to understand network dependencies between servers and group the related servers as applications for migration planning.

In this section you'll find instructions on how to view and work with data discovered by Agentless Collector and Discovery Agent from both the console and the AWS CLI.

Topics

View collected data using the Migration Hub console

For both the Application Discovery Service Agentless Collector (Agentless Collector) and AWS Discovery Agent (Discovery Agent), after the data collection process starts, you can use the console to view their collected data about your servers and VMs. Data appears in the console approximately 15 minutes after data collection starts. You can also view this data in CSV format by exporting the collected data by making API calls using the AWS CLI.

To view collected data about discovered servers in the console, follow the steps in Viewing servers in the AWS Migration Hub console. To learn more about using the console to view, sort, and tag servers discovered by your Agentless Collectors or Discovery Agents, see Discovering data with the AWS Migration Hub console.

The Agentless Collector database and analytics data collection module uploads the collected data to the Amazon S3 bucket. You can view the data from this bucket in the AWS DMS console. To view collected data about discovered database and analytics servers, follow the steps in Viewing your collected data.

Matching logic for discovered servers and applications

AWS Application Discovery Service (Application Discovery Service) has built-in matching logic that identifies when servers that it discovers match existing entries. When this logic finds a match, it updates the information for the already-existing discovered server with new values.

This matching logic handles duplicate servers from multiple sources including AWS Migration Hub (Migration Hub) import, Application Discovery Service Agentless Collector (Agentless Collector), AWS Application Discovery Agent (Discovery Agent), and other migration tools. For more information about Migration Hub import, see Migration Hub Import.

When server discovery occurs, each entry is cross-checked with previously imported records to ensure that the imported server does not already exist. If no match is found, a new record is created and a new unique server identifier is assigned. If a match is found, then a new entry is still created, but it's assigned the same unique server identifier as the existing server. When viewing this server in the Migration Hub console, you only find one unique entry for the server.

Server attributes associated with this entry are merged to show attribute values from a previously available record as well as the newly imported record. If there is more than one value for a given server attribute from multiple sources, e.g., two different values within for `Total RAM` associated with a given server discovered using import and also by the Discovery Agent, then the value that was most recently updated is shown in the matched record for the server.

Matching fields

The following fields are used to match servers when discovery tools are used.

- **ExternalId** – This is the primary field used to match servers. If the value in this field is identical to another `ExternalId` in another entry, then Application Discovery Service matches the two entries, regardless of whether the other fields match or not.
- **IPAddress**
- **HostName**
- **MacAddress**
- **VMware.MoRefId** and **VMware.vCenterId** – Both of these values must be identical to the respective fields in another entry for Application Discovery Service to perform a match.

Exploring data in Amazon Athena

Data exploration in Amazon Athena allows you to analyze the data that's collected from all the discovered on-premises servers by Discovery Agent in one place. Once Data exploration in Amazon Athena is enabled from the Migration Hub console (or by using the StartContinousExport API) and the data collection for agents is turned on, data that's collected by agents is automatically get stored in your S3 bucket at regular intervals. For more information, see Exploring data in Amazon Athena.

Data exploration in Amazon Athena allows you to analyze the data that's collected from all the discovered on-premises servers by Discovery Agents in one place. Once data exploration in Amazon Athena is enabled from the Migration Hub console (or by using the StartContinousExport API) and the data collection for agents is turned on, data that's collected by agents is automatically get stored in your S3 bucket at regular intervals.

You can then visit Amazon Athena to run pre-defined queries to analyze the time-series system performance for each server, the type of processes that are running on each server and the network dependencies between different servers. In addition, you can write your own custom queries using Amazon Athena, upload additional existing data sources such as configuration management database (CMDB) exports, and associate the discovered servers with the actual business applications. You can also integrate the Athena database with Amazon Quick to visualize the query outputs and perform additional analysis.

The topics in this section describe the ways that you can work with your data in Athena to assess and plan for migrating your local environment to AWS.

Turning on data exploration in Amazon Athena

Data exploration in Amazon Athena is enabled by turning on Continuous Export using the Migration Hub console or an API call from the AWS CLI. You must turn on data exploration before you can see and start exploring your discovered data in Amazon Athena.

When you turn on Continuous Export the service-linked role `AWSServiceRoleForApplicationDiscoveryServiceContinuousExport` is automatically used by your account. For more information about this service-linked role, see Service-linked role permissions for Application Discovery Service.

The following instructions show how to turn on data exploration in Amazon Athena by using the console and the AWS CLI.

Turn on with the console

Data exploration in Amazon Athena is enabled by Continuous Export implicitly being turned on when you choose "Start data collection", or click the toggle labeled, "Data exploration in Amazon Athena" on the **Data Collectors** page of the Migration Hub console.

To turn on data exploration in Amazon Athena from the console

1. In the navigation pane, choose **Data Collectors**.
2. Choose the **Agents** tab.
3. Choose **Start data collection**, or if you already have data collection turned on, click the **Data exploration in Amazon Athena** toggle.
4. In the dialog box generated from the previous step, click the checkbox agreeing to associated costs and choose **Continue** or **Enable**.

> **Note**
>
> Your agents are now running in "continuous export" mode which will enable you to see and work with your discovered data in Amazon Athena. The first time this is enable it may take up to 30 minutes for your data to appear in Amazon Athena.

Enable with the AWS CLI

Data exploration in Amazon Athena is enabled by Continuous Export explicitly being turned on through an API call from the AWS CLI. To do this, the AWS CLI must first be installed in your environment.

To install the AWS CLI and turn on data exploration in Amazon Athena

1. Install the AWS CLI for your operating system (Linux, macOS, or Windows). See the AWS Command Line Interface User Guide for instructions.
2. Open the Command prompt (Windows) or Terminal (Linux or macOS).
   a. Type `aws configure` and press Enter.
   b. Enter your AWS Access Key Id and AWS Secret Access Key.
   c. Enter `us-west-2` for the Default Region Name.

d. Enter `text` for Default Output Format.

3. Type the following command:

```
aws discovery start-continuous-export
```

> **Note**
>
> Your agents are now running in "continuous export" mode which will enable you to see and work with your discovered data in Amazon Athena. The first time this is enable it may take up to 30 minutes for your data to appear in Amazon Athena.

Exploring data directly in Amazon Athena

After you turn on data exploration in Amazon Athena, you can begin exploring and working with detailed current data that was discovered by your agents by querying the data directly in Athena. You can use the data to generate spreadsheets, run a cost analysis, port the query to a visualization program to diagram network dependencies, and more.

The following instructions explain how to explore your agent data directly in the Athena console. If you don't have any data in Athena or have not enabled data exploration in Amazon Athena, you will be prompted by a dialog box to enable data exploration in Amazon Athena , as explained in Turning on data exploration in Amazon Athena.

To explore agent-discovered data directly in Athena

1. In the AWS Migration Hub console, choose **Servers** in the navigation pane.
2. To open the Amazon Athena console, choose **Explore data in Amazon Athena**.
3. On the **Query Editor** page, in the navigation pane under **Database**, make sure that **application_discovery_service_database** is selected.

> **Note**
>
> Under **Tables** the following tables represent the datasets grouped by the agents.
>
> - **os_info_agent**
> - **network_interface_agent**

- **sys_performance_agent**
- **processes_agent**
- **inbound_connection_agent**
- **outbound_connection_agent**
- **id_mapping_agent**

4. Query the data in the Amazon Athena console by writing and running SQL queries in the Athena Query Editor. For example, you can use the following query to see all of the discovered server IP addresses.

```
SELECT * FROM network_interface_agent;
```

For more example queries, see Using predefined queries in Amazon Athena.

Visualizing Amazon Athena data

To visualize your data, a query can be ported to a visualization program such as Amazon Quick or other open-source visualization tools such as Cytoscape, yEd, or Gelphi. Use these tools to render network diagrams, summary charts, and other graphical representations. When this method is used, you connect to Athena through the visualization program so that it can access your collected data as a source to produce the visualization.

To visualize your Amazon Athena data using Quick

1. Sign in to Amazon Quick.
2. Choose **Connect to another data source or upload a file**.
3. Choose **Athena**. The **New Athena data source** dialog box displays.
4. Enter a name in the **Data source name** field.
5. Choose **Create data source**.
6. Select the **Agents-servers-os** table in the **Choose your table** dialog box and choose **Select**.
7. In the **Finish dataset creation** dialog box, select **Import to SPICE for quicker analytics**, and choose **Visualize**.

   Your visualization is rendered.

Using predefined queries in Amazon Athena

This section contains a set of predefined queries that perform typical use cases, such as TCO analysis and network visualization. You can use these queries as is or modify them to suit your needs.

To use a predefined query

1. In the AWS Migration Hub console, choose **Servers** in the navigation pane.
2. To open the Amazon Athena console, choose **Explore data in Amazon Athena**.
3. On the **Query Editor** page, in the navigation pane under **Database**, make sure that **application_discovery_service_database** is selected.
4. Choose the plus (**+**) sign in the Query Editor to create a tab for a new query.
5. Copy one of the queries from Predefined queries.
6. Paste the query into the query pane of the new query tab you just created.
7. Choose **Run Query**.

Predefined queries

Choose a title to see information about the query.

Obtain IP addresses and hostnames for servers

This view helper function retrieves IP addresses and hostnames for a given server. You can use this view in other queries. For information about how to create a view, see CREATE VIEW in the *Amazon Athena User Guide*.

```
CREATE OR REPLACE VIEW hostname_ip_helper AS
SELECT DISTINCT
  "os"."host_name"
, "nic"."agent_id"
, "nic"."ip_address"
FROM
  os_info_agent os
, network_interface_agent nic
WHERE ("os"."agent_id" = "nic"."agent_id");
```

Identify servers with or without agents

This query can help you perform data validation. If you've deployed agents on a number of servers in your network, you can use this query to understand if there are other servers in your network without agents deployed on them. In this query, we look into the inbound and outbound network traffic, and filter the traffic for private IP addresses only. That is, IP addresses starting with 192, 10, or 172.

```
SELECT DISTINCT "destination_ip" "IP Address" ,
         (CASE
    WHEN (
    (SELECT "count"(*)
    FROM network_interface_agent
    WHERE ("ip_address" = "destination_ip") ) = 0) THEN
        'no'
        WHEN (
        (SELECT "count"(*)
        FROM network_interface_agent
        WHERE ("ip_address" = "destination_ip") ) > 0) THEN
            'yes' END) "agent_running"
    FROM outbound_connection_agent
WHERE ((("destination_ip" LIKE '192.%')
        OR ("destination_ip" LIKE '10.%'))
        OR ("destination_ip" LIKE '172.%'))
UNION
SELECT DISTINCT "source_ip" "IP ADDRESS" ,
         (CASE
    WHEN (
    (SELECT "count"(*)
    FROM network_interface_agent
    WHERE ("ip_address" = "source_ip") ) = 0) THEN
        'no'
        WHEN (
        (SELECT "count"(*)
        FROM network_interface_agent
        WHERE ("ip_address" = "source_ip") ) > 0) THEN
            'yes' END) "agent_running"
    FROM inbound_connection_agent
WHERE ((("source_ip" LIKE '192.%')
        OR ("source_ip" LIKE '10.%'))
        OR ("source_ip" LIKE '172.%'));
```

Analyze system performance data for servers with agents

You can use this query to analyze system performance and utilization pattern data for your on-premises servers that have agents installed on them. The query combines the `system_performance_agent` table with the `os_info_agent` table to identify the hostname for each server. This query returns the time series utilization data (in 15 minute intervals) for all the servers where agents are running.

```
SELECT "OS"."os_name" "OS Name" ,
    "OS"."os_version" "OS Version" ,
    "OS"."host_name" "Host Name" ,
     "SP"."agent_id" ,
     "SP"."total_num_cores" "Number of Cores" ,
     "SP"."total_num_cpus" "Number of CPU" ,
     "SP"."total_cpu_usage_pct" "CPU Percentage" ,
     "SP"."total_disk_size_in_gb" "Total Storage (GB)" ,
     "SP"."total_disk_free_size_in_gb" "Free Storage (GB)" ,
     ("SP"."total_disk_size_in_gb" - "SP"."total_disk_free_size_in_gb") "Used
 Storage" ,
     "SP"."total_ram_in_mb" "Total RAM (MB)" ,
     ("SP"."total_ram_in_mb" - "SP"."free_ram_in_mb") "Used RAM (MB)" ,
     "SP"."free_ram_in_mb" "Free RAM (MB)" ,
     "SP"."total_disk_read_ops_per_sec" "Disk Read IOPS" ,
     "SP"."total_disk_bytes_written_per_sec_in_kbps" "Disk Write IOPS" ,
     "SP"."total_network_bytes_read_per_sec_in_kbps" "Network Reads (kbps)" ,
     "SP"."total_network_bytes_written_per_sec_in_kbps" "Network Write (kbps)"
FROM "sys_performance_agent" "SP" , "OS_INFO_agent" "OS"
WHERE ("SP"."agent_id" = "OS"."agent_id") limit 10;
```

Track outbound communication between servers based on port number and process details

This query gets the details on the outbound traffic for each service, along with the port number and process details.

Before running the query, if you have not already done so, you must create the `iana_service_ports_import` table that contains the IANA port registry database downloaded from IANA. For information about how to create this table, see Creating the IANA port registry import table.

After the `iana_service_ports_import` table is created, create two view helper functions for tracking outbound traffic. For information about how to create a view, see CREATE VIEW in the *Amazon Athena User Guide*.

To create outbound tracking helper functions

1. Open the Athena console at https://console.aws.amazon.com/athena/.
2. Create the `valid_outbound_ips_helper` view, using the following helper function that lists all distinct outbound destination IP addresses.

```
CREATE OR REPLACE VIEW valid_outbound_ips_helper AS
SELECT DISTINCT "destination_ip"
FROM outbound_connection_agent;
```

3. Create the `outbound_query_helper` view, using the following helper function that determines the frequency of communication for outbound traffic.

```
CREATE OR REPLACE VIEW outbound_query_helper AS
SELECT "agent_id" ,
         "source_ip" ,
         "destination_ip" ,
         "destination_port" ,
         "agent_assigned_process_id" ,
         "count"(*) "frequency"
FROM outbound_connection_agent
WHERE (("ip_version" = 'IPv4')
        AND ("destination_ip" IN
    (SELECT *
    FROM valid_outbound_ips_helper )))
GROUP BY  "agent_id", "source_ip", "destination_ip", "destination_port",
 "agent_assigned_process_id";
```

4. After you create the `iana_service_ports_import` table and your two helper functions, you can run the following query to get the details on the outbound traffic for each service, along with the port number and process details.

```
SELECT hip1.host_name "Source Host Name",
         outbound_connections_results0.source_ip "Source IP Address",
         hip2.host_name "Destination Host Name",
         outbound_connections_results0.destination_ip "Destination IP Address",
         outbound_connections_results0.frequency "Connection Frequency",
         outbound_connections_results0.destination_port "Destination Communication
 Port",
         outbound_connections_results0.servicename "Process Service Name",
         outbound_connections_results0.description "Process Service Description"
FROM
```

```
    (SELECT DISTINCT o.source_ip,
         o.destination_ip,
         o.frequency,
         o.destination_port,
         ianap.servicename,
         ianap.description
    FROM outbound_query_helper o, iana_service_ports_import ianap
    WHERE o.destination_port = TRY_CAST(ianap.portnumber AS integer)) AS
 outbound_connections_results0 LEFT OUTER
JOIN hostname_ip_helper hip1
    ON outbound_connections_results0.source_ip = hip1.ip_address LEFT OUTER
JOIN hostname_ip_helper hip2
    ON outbound_connections_results0.destination_ip = hip2.ip_address
```

Track inbound communication between servers based on port number and process details

This query gets information about inbound traffic for each service, along with the port number and process details.

Before running this query, if you have not already done so, you must create the `iana_service_ports_import` table that contains the IANA port registry database downloaded from IANA. For information about how to create this table, see Creating the IANA port registry import table.

After the `iana_service_ports_import` table is created, create two view helper functions for tracking inbound traffic. For information about how to create a view, see CREATE VIEW in the *Amazon Athena User Guide*.

To create import tracking helper functions

1. Open the Athena console at https://console.aws.amazon.com/athena/.
2. Create the `valid_inbound_ips_helper` view, using the following helper function that lists all distinct inbound source IP addresses.

```
CREATE OR REPLACE VIEW valid_inbound_ips_helper AS
SELECT DISTINCT "source_ip"
FROM inbound_connection_agent;
```

3. Create the `inbound_query_helper` view, using the following helper function that determines the frequency of communication for inbound traffic.

```
CREATE OR REPLACE VIEW inbound_query_helper AS
SELECT "agent_id" ,
         "source_ip" ,
         "destination_ip" ,
         "destination_port" ,
         "agent_assigned_process_id" ,
         "count"(*) "frequency"
FROM inbound_connection_agent
WHERE (("ip_version" = 'IPv4')
        AND ("source_ip" IN
    (SELECT *
    FROM valid_inbound_ips_helper )))
GROUP BY  "agent_id", "source_ip", "destination_ip", "destination_port",
 "agent_assigned_process_id";
```

4. After you create the `iana_service_ports_import` table and your two helper functions, you can run the following query to get the details on the inbound traffic for each service, along with the port number and process details.

```
SELECT hip1.host_name "Source Host Name",
         inbound_connections_results0.source_ip "Source IP Address",
         hip2.host_name "Destination Host Name",
         inbound_connections_results0.destination_ip "Destination IP Address",
         inbound_connections_results0.frequency "Connection Frequency",
         inbound_connections_results0.destination_port "Destination Communication
 Port",
         inbound_connections_results0.servicename "Process Service Name",
         inbound_connections_results0.description "Process Service Description"
FROM
    (SELECT DISTINCT i.source_ip,
         i.destination_ip,
         i.frequency,
         i.destination_port,
         ianap.servicename,
         ianap.description
    FROM inbound_query_helper i, iana_service_ports_import ianap
    WHERE i.destination_port = TRY_CAST(ianap.portnumber AS integer)) AS
 inbound_connections_results0 LEFT OUTER
JOIN hostname_ip_helper hip1
    ON inbound_connections_results0.source_ip = hip1.ip_address LEFT OUTER
JOIN hostname_ip_helper hip2
```

```
    ON inbound_connections_results0.destination_ip = hip2.ip_address
```

Identify running software from port number

This query identifies the running software based on port numbers.

Before running this query, if you have not already done so, you must create the `iana_service_ports_import` table that contains the IANA port registry database downloaded from IANA. For information about how to create this table, see Creating the IANA port registry import table.

Run the following query to identify the running software based on port numbers.

```
SELECT o.host_name "Host Name",
       ianap.servicename "Service",
       ianap.description "Description",
       con.destination_port,
       con.cnt_dest_port "Destination Port Count"
FROM   (SELECT agent_id,
               destination_ip,
               destination_port,
               Count(destination_port) cnt_dest_port
        FROM   inbound_connection_agent
        GROUP  BY agent_id,
                  destination_ip,
                  destination_port) con,
       (SELECT agent_id,
               host_name,
               Max("timestamp")
        FROM   os_info_agent
        GROUP  BY agent_id,
                  host_name) o,
       iana_service_ports_import ianap
WHERE  ianap.transportprotocol = 'tcp'
       AND con.destination_ip NOT LIKE '172%'
       AND con.destination_port = ianap.portnumber
       AND con.agent_id = o.agent_id
ORDER BY cnt_dest_port DESC;
```

Creating the IANA port registry import table

Some of the predefined queries require a table named `iana_service_ports_import` that contains information downloaded from Internet Assigned Numbers Authority (IANA).

To create the iana_service_ports_import table

1. Download the IANA port registry database **CSV** file from Service Name and Transport Protocol Port Number Registry on *iana.org*.
2. Upload the file to Amazon S3. For more information, see How Do I Upload Files and Folders to an S3 Bucket?.
3. Create a new table in Athena named `iana_service_ports_import`. For instructions, see Create a Table in the *Amazon Athena User Guide*. In the following example, you need to replace `my_bucket_name` with the name of the S3 bucket that you uploaded the CSV file to in the previous step.

```
CREATE EXTERNAL TABLE IF NOT EXISTS iana_service_ports_import (
         ServiceName STRING,
         PortNumber INT,
         TransportProtocol STRING,
         Description STRING,
         Assignee STRING,
         Contact STRING,
         RegistrationDate STRING,
         ModificationDate STRING,
         Reference STRING,
         ServiceCode STRING,
         UnauthorizedUseReported STRING,
         AssignmentNotes STRING
)
ROW FORMAT SERDE 'org.apache.hadoop.hive.serde2.lazy.LazySimpleSerDe'
WITH SERDEPROPERTIES (
  'serialization.format' = ',',
  'quoteChar' = '"',
  'field.delim' = ','
) LOCATION 's3://my_bucket_name/'
TBLPROPERTIES ('has_encrypted_data'='false',"skip.header.line.count"="1");
```

Discovering data with the AWS Migration Hub console

AWS Application Discovery Service (Application Discovery Service) is integrated with AWS Migration Hub (Migration Hub) and customers can view and manage their data collectors, servers, and applications within Migration Hub. When you use the Application Discovery Service console, you are redirected to the Migration Hub console. Working with the Migration Hub console requires no extra steps or setup on your part.

In this section, you can find how to manage and monitor Application Discovery Service Agentless Collector (Agentless Collector) and AWS Application Discovery Agent (Discovery Agent) using the console.

Topics

Viewing data in the AWS Migration Hub console dashboard

To view the main dashboard, choose **Dashboard** from the AWS Migration Hub (Migration Hub) console navigation pane. In the Migration Hub main dashboard, you can view high-level statistics about servers, applications, and data collectors such as Application Discovery Service Agentless Collector (Agentless Collector) and AWS Application Discovery Agent (Discovery Agent).

The main dashboard gathers data from the **Discover** and **Migrate** dashboards in a central location. It has four status and information panes and a list of links for quick access. Using the panes, you can see a summary status of your most recently updated applications. You can also get quick access to any of your applications, get an overview of applications in different states, and track the migration progress over time.

To view the main dashboard, choose **Dashboard** from the navigation pane, which is on the left side of the Migration Hub console homepage.

Starting and stopping data collectors in the AWS Migration Hub console

Application Discovery Service Agentless Collector (Agentless Collector) and AWS Application Discovery Agent (Discovery Agent) are the data collection tools that AWS Application Discovery Service (Application Discovery Service) uses to help you discover your existing infrastructure. The following steps explain how to download and deploy these discovery data collection tools, Deploy Agentless Collector and AWS Application Discovery Agent.

These data collection tools store their data in the Application Discovery Service's repository, providing details about each server and the processes running on them. When either of these tools is deployed, you can start, stop, and view the collected data from the AWS Migration Hub (Migration Hub) console.

After the AWS Application Discovery Agent (Discovery Agent) is deployed, you can start or stop the data collection process on the **Data Collectors** page of the AWS Migration Hub (Migration Hub) console.

To start or stop data collection tools

1. Using your AWS account, sign in to the AWS Management Console and open the Migration Hub console at https://console.aws.amazon.com/migrationhub/.
2. In the Migration Hub console navigation pane under **Discover**, choose **Data collectors**.
3. Choose the **Agents** tab.
4. Select the check box of the collection tool you want to start or stop.
5. Choose **Start data collection** or **Stop data collection**.

Sorting data collectors in the AWS Migration Hub console

If you deployed many data collectors, you can sort the displayed list of deployed collector's on the **Data Collectors** page of the console. You sort the list by applying filters in the search bar. You can search and filter on most of the criteria specified in the **Data Collectors** list.

The following table shows the search criteria that you can use for **Agents**, including operators, values, and a definition of the values.

Search Criterion	Operator	Value: Definition
Agent ID	==	Any agent ID selected from the pre-populated list from which a collection tool is installed.
Hostname	== !=	For agents, any host name selected from the pre-popul ated list of hosts where an agent is installed.
Collection status	== !=	Started: Data is being collected and sent to Application Discovery Service Start scheduled: Data collection is scheduled to start. Data will be sent to Application Discovery Service on next ping, and status will change to **Started**. Stopped: Data is not being collected or sent to Applicati on Discovery Service. Stop scheduled: Data collectio n is scheduled to stop. Data will stop being sent to Application Discovery Service on next ping, and status will change to **Stopped**.

Search Criterion	Operator	Value: Definition
Health	== !=	Healthy: Data collection isn't turned on. The tool is functioning normally. Unhealthy: The tool is in an error state. Data isn't being collected or reported. Unknown: No connection established in over an hour. Shutdown: The tool last communicated "shutting down" due to a system, service, or daemon shut down. If a reboot or tool upgrade occurred, status will change to another state at the first reporting cycle. Running: Data collection is turned on. The tool is functioning normally.
IP address	== !=	Any IP address selected from the pre-populated list where a collection tool is installed.

The following table shows the search criteria that you can use for **Agentless collectors**, including operators, values, and a definition of the values.

Search Criterion	Operator	Value: Definition
ID	==	Any agentless collector ID selected from the pre-

Search Criterion	Operator	Value: Definition
		populated list from which a collection tool is installed.
Hostname	== !=	For agentless collectors, any host name selected from the pre-populated list of hosts where an agentless collectors is installed.
Status	== !=	Collecting data: Data collectio n is turned on. The tool is functioning normally. Ready to configure— Data collection isn't turned on. The tool is functioning normally. Requires attention— The tool is in an error state and needs attention. Unknown: No connection established in over an hour. Shut down: The tool last communicated "shutting down" due to a system, service, or daemon shut down. If a reboot or tool upgrade occurred, status will change to another state at the first reporting cycle.
IP address	== !=	Any IP address selected from the pre-populated list where a collection tool is installed.

To sort data collectors by applying search filters

1. Using your AWS account, sign in to the AWS Management Console and open the Migration Hub console at https://console.aws.amazon.com/migrationhub/.
2. In the Migration Hub console navigation pane under **Discover**, choose **Data Collectors**.
3. Choose either the **Agentless collectors** or **Agents** tab.
4. Click inside the search bar and choose a search criterion from the list.
5. Choose an operator from the next list.
6. Choose a value from the last list.

Viewing servers in the AWS Migration Hub console

The **Servers** page provides system configuration and performance data about each server instance known to the data collection tools. You can view server information, sort servers with filters, tag servers with key-value pairs, and export detailed server and system information.

You can get a general view and a detailed view of the servers discovered by the data collection tools.

To view discovered servers

1. Using your AWS account, sign in to the AWS Management Console and open the Migration Hub console at https://console.aws.amazon.com/migrationhub/.
2. In the Migration Hub console navigation pane under **Discover**, choose **Servers**. The discovered servers appear in the servers list.
3. For more detail about a server, choose its server link in the **Server info** column. Doing so displays a screen that describes the server.

The server's detail screen displays system information and performance metrics. You can also find a button to export network dependencies and processes information. To export detailed server information, see Using AWS Migration Hub to export server data.

Sorting servers in the AWS Migration Hub console

To easily find specific servers, apply search filters to sort through all the servers discovered by the collection tools. You can search and filter on numerous criteria.

To sort servers by applying search filters

1. Using your AWS account, sign in to the AWS Management Console and open the Migration Hub console at https://console.aws.amazon.com/migrationhub/.
2. In the Migration Hub console navigation pane under **Discover**, choose **Servers**.
3. Click inside the search bar, and choose a search criterion from the list.
4. Choose an operator from the next list.
5. Type in a case-sensitive value for the search criterion you selected, and press Enter.
6. Multiple filters can be applied by repeating steps 2 - 4.

Tagging servers in the AWS Migration Hub console

To assist migration planning and help stay organized, you can create multiple tags for each server. *Tags* are user-defined key-value pairs that can store any custom data or metadata about servers. You can tag an individual server or multiple servers in a single operation. AWS Application Discovery Service (Application Discovery Service) tags are similar to AWS tags, but the two types of tag cannot be used interchangeably.

You can add or remove multiple tags for one or more servers from the main **Servers** page. On a server's detail page, you can add or remove one or more tags for the selected server. You can do any type of tagging task involving multiple servers or tags in a single operation. You can also remove tags.

To add tags to one or more servers

1. Using your AWS account, sign in to the AWS Management Console and open the Migration Hub console at https://console.aws.amazon.com/migrationhub/.
2. In the Migration Hub console navigation pane under **Discover**, choose **Servers**.
3. In the **Server info** column, choose the server link for the server that you want to add tags for. To add tags to more than one server at a time, click inside the check boxes of multiple servers.
4. Choose **Add tags**, and then choose **Add new tag**.
5. In the dialog box, type a key in the **Key** field, and optionally a value in the **Value** field.

   Add more tags by choosing **Add new tag** and adding more information.
6. Choose **Save**.

To remove tags from one or more servers

1. Using your AWS account, sign in to the AWS Management Console and open the Migration Hub console at https://console.aws.amazon.com/migrationhub/.
2. In the Migration Hub console navigation pane under **Discover**, choose **Servers**.
3. In the **Server info** column, choose the server link for the server that you want to remove tags from. Select the check boxes of multiple servers to remove tags from more than one server at a time.
4. Choose **Remove tags**.
5. Select each tag that you want to remove.
6. Choose **Confirm**.

Using AWS Migration Hub to export server data

This topic explains how to export server data by using the AWS Management Console, the AWS Command Line Interface, or the API.

To use the AWS Management Console to export server data for all servers

1. Sign in to the AWS Management Console and open the Migration Hub console at https://console.aws.amazon.com/migrationhub/.
2. In the left navigation pane under **Discover**, choose **Servers**.
3. Choose **Actions**, and then choose **Export discovery data**.
4. In the **Exports** section at the bottom of the screen, choose **Export server details**. This action generates a .zip file that includes the .csv files that are described in the following table.

File name	Description
{account_id}_Application.csv	Details of each application, including the server count, name, and description.
{account_id}_ApplicationResourceAssociation.csv	The relationship between servers and applications.
{account_id}_ImportTemplate	The summary of each server's application and tags. This file can be modified and

File name	Description
	re-imported to update the application associated with the server.
{account_id}_NetworkInterface.csv	Details of each network interface including the associated server, address, and switch.
{account_id}_Server.csv	Details of each server, including operating system, host name, and hypervisor.
{account_id}_SystemPerformance.csv	Details of each server, including CPU, memory and storage configuration, and performance.
{account_id}_Tags.csv	Details of each tag associated with a server.
{account_id}_VMwareInfo.csv	Details of each VMware configuration, including moRef, vmName, and vCenter.

To use the AWS Management Console to export agent data for a specific server

1. Sign in to the AWS Management Console and open the Migration Hub console at https://console.aws.amazon.com/migrationhub/.
2. In the left navigation pane under **Discover**, choose **Servers**.
3. Place the cursor in the search field under **Servers**. A drop-down list appears. In that list, under **Properties**, choose **Source**, then choose the **=** operator, and then choose **Source = Agent**.
4. In the search results, choose the name of the server for which you want to export data. This action takes you to the details page for that server.
5. Enter a start time and an end time, and then choose **Export**. The exported .zip file includes the .csv files that are described in the following table.

{account_id}_destinationProcessConnection.csv	Details of the inbound connections into the server.

{account_id}_networkInterface.csv	Details of each network interface including address, mask, and name
{account_id}_osInfo.csv	Details of the operating system including CPU type, hypervisor and operating system name.
{account_id}_process.csv	Details of the processes running on the server.
{account_id}_sourceProcessConnection.csv	Details of the outbound connection originating from the server.
{account_id}_systemPerformance.csv	Details of the CPU, memory and storage configuration & performance for the server.

To use the AWS Command Line Interface or the API to export server data

1. Run start-export-task. The corresponding API operation is StartExportTask
2. Run describe-export-tasks. The corresponding API operation is DescribeExportTasks.

Grouping servers in the AWS Migration Hub console

Some of your discovered servers might need to be migrated together to remain functional. In this case, you can logically define and group discovered servers into applications.

As part of the grouping process, you can search, filter, and add tags.

To group servers into a new or existing application

1. Using your AWS account, sign in to the AWS Management Console and open the Migration Hub console at https://console.aws.amazon.com/migrationhub/.
2. In the Migration Hub console navigation pane under **Discover**, choose **Servers**.
3. In the servers list, select each server that you want to group into a new or existing application.

To help choose servers for your group, you can search and filter on any criteria that you specify in the server list. Click inside the search bar and choose an item from the list, choose an operator from the next list, and then type in your criteria.

4. Optional: For each selected server, choose **Add tag**, type a value for **Key**, and then optionally type a value for **Value**.
5. Choose **Group as application** to create your application, or add to an existing one.
6. In the **Group as application** dialog box, choose **Group as a new application** or **Add to an existing application**.

   a. If you chose **Group as a new application**, type a name for **Application name**. Optionally, you can type a description for **Application description**.

   b. If you chose **Add to an existing application**, select the name of the application to add to in the list.
7. Choose **Save**.

Using the Application Discovery Service API to query discovered configuration items

A *configuration item* is an IT asset that was discovered in your data center by an agent or by an import. When you use AWS Application Discovery Service (Application Discovery Service), you use the API to specify filters and query specific configuration items for server, application, process, and connection assets. For information about the API, see Application Discovery Service API Reference.

The tables in the following sections list the available input filters and output sorting options for two Application Discovery Service actions:

- `DescribeConfigurations`
- `ListConfigurations`

The filtering and sorting options are organized by the type of asset to which apply (server, application, process, or connection).

> **Important**
>
> Results returned by `DescribeConfigurations`, `ListConfigurations`, and `StartExportTask` might not contain recent updates. For more information, see the section called "Eventual consistency".

Using the `DescribeConfigurations` action

The `DescribeConfigurations`action retrieves attributes for a list of configuration IDs. All the supplied IDs must be for the same asset type (server, application, process, or connection). Output fields are specific to the asset type selected. For example, the output for a server configuration item includes a list of attributes about the server, such as host name, operating system, and number of network cards. For more information about command syntax, see DescribeConfigurations.

The `DescribeConfigurations`action does not support filtering.

Output fields for `DescribeConfigurations`

The following tables, organized by asset type, list the supported output fields of the `DescribeConfigurations`action. The ones marked as mandatory are always present in the output.

Server assets

Field	Mandatory
server.agentId	
server.applications	
server.applications.hasMore Values	
server.configurationId	x
server.cpuType	
server.hostName	
server.hypervisor	
server.networkInterfaceInfo	
server.networkInterfaceInfo .hasMoreValues	
server.osName	
server.osVersion	
server.tags	
server.tags.hasMoreValues	
server.timeOfCreation	x
server.type	

Field	Mandatory
server.performance.avgCpuUsagePct	
server.performance.avgDiskReadIOPS	
server.performance.avgDiskReadsPerSecondInKB	
server.performance.avgDiskWriteIOPS	
server.performance.avgDiskWritesPerSecondInKB	
server.performance.avgFreeRAMInKB	
server.performance.avgNetworkReadsPerSecondInKB	
server.performance.avgNetworkWritesPerSecondInKB	
server.performance.maxCpuUsagePct	
server.performance.maxDiskReadIOPS	
server.performance.maxDiskReadsPerSecondInKB	
server.performance.maxDiskWriteIOPS	

Field	Mandatory
server.performance.maxDiskWritesPerSecondInKB	
server.performance.maxNetworkReadsPerSecondInKB	
server.performance.maxNetworkWritesPerSecondInKB	
server.performance.minFreeRAMInKB	
server.performance.numCores	
server.performance.numCpus	
server.performance.numDisks	
server.performance.numNetworkCards	
server.performance.totalRAMInKB	

Process assets

Field	Mandatory
process.commandLine	
process.configurationId	x
process.name	
process.path	
process.timeOfCreation	x

Application assets

Field	Mandatory
application.configurationId	x
application.description	
application.lastModifiedTime	x
application.name	x
application.serverCount	x
application.timeOfCreation	x

Using the ListConfigurations action

The ListConfigurationsaction retrieves a list of configuration items according to the criteria that you specify in a filter. For more information about command syntax, see ListConfigurations.

Output fields for ListConfigurations

The following tables, organized by asset type, list the supported output fields of the ListConfigurationsaction. The ones marked as mandatory are always present in the output.

Server assets

Field	Mandatory
server.configurationId	x
server.agentId	
server.hostName	
server.osName	
server.osVersion	

Field	Mandatory
server.timeOfCreation	x
server.type	

Process assets

Field	Mandatory
process.commandLine	
process.configurationId	x
process.name	
process.path	
process.timeOfCreation	x
server.agentId	
server.configurationId	x

Application assets

Field	Mandatory
application.configurationId	x
application.description	
application.name	x
application.serverCount	x
application.timeOfCreation	x
application.lastModifiedTime	x

Connection assets

Field	Mandatory
connection.destinationIp	x
connection.destinationPort	x
connection.ipVersion	x
connection.latestTimestamp	x
connection.occurrence	x
connection.sourceIp	x
connection.transportProtocol	
destinationProcess.configurationId	
destinationProcess.name	
destinationServer.configurationId	
destinationServer.hostName	
sourceProcess.configurationId	
sourceProcess.name	
sourceServer.configurationId	
sourceServer.hostName	

Supported filters for `ListConfigurations`

The following tables, organized by asset type, list the supported filters for the `ListConfigurations`action. Filters and values are in a key/value relationship defined by one of the supported logical conditions. You can sort the output of the indicated filters.

Server assets

Filter	Supported conditions	Supported values	Supported sorting
`server.configurationId`	• EQUALS • NOT_EQUALS • EQ • NE	• Any valid server configuration ID	None
`server.hostName`	• EQUALS • NOT_EQUALS • EQ • NE • CONTAINS • NOT_CONTAINS	• String	• ASC • DESC
`server.osName`	• EQUALS • NOT_EQUALS • EQ • NE • CONTAINS • NOT_CONTAINS	• String	• ASC • DESC
`server.osVersion`	• EQUALS • NOT_EQUALS • EQ • NE • CONTAINS • NOT_CONTAINS	• String	• ASC • DESC

Filter	Supported conditions	Supported values	Supported sorting
server.agentId	• EQUALS • NOT_EQUALS • EQ • NE	• String	None
server.connectorId	• EQUALS • NOT_EQUALS • EQ • NE	• String	None
server.type	• EQUALS • NOT_EQUALS • EQ • NE	String with one of the following values: • EC2 • OTHER • VMWARE_VM • VMWARE_HOST • VMWARE_VM_TEMPLATE	None
server.vmWareInfo.morefId	• EQUALS • NOT_EQUALS • EQ • NE • CONTAINS • NOT_CONTAINS	• String	None

Filter	Supported conditions	Supported values	Supported sorting
server.vmWareInfo.vcenterId	• EQUALS • NOT_EQUALS • EQ • NE • CONTAINS • NOT_CONTAINS	• String	None
server.vmWareInfo.hostId	• EQUALS • NOT_EQUALS • EQ • NE • CONTAINS • NOT_CONTAINS	• String	None
server.networkInterfaceInfo.portGroupId	• EQUALS • NOT_EQUALS • EQ • NE • CONTAINS • NOT_CONTAINS	• String	None
server.networkInterfaceInfo.portGroupName	• EQUALS • NOT_EQUALS • EQ • NE • CONTAINS • NOT_CONTAINS	• String	None

Filter	Supported conditions	Supported values	Supported sorting
server.networkInterfaceInfo.virtualSwitchName	• EQUALS • NOT_EQUALS • EQ • NE • CONTAINS • NOT_CONTAINS	• String	None
server.networkInterfaceInfo.ipAddress	• EQUALS • NOT_EQUALS • EQ • NE • CONTAINS • NOT_CONTAINS	• String	None
server.networkInterfaceInfo.macAddress	• EQUALS • NOT_EQUALS • EQ • NE • CONTAINS • NOT_CONTAINS	• String	None
server.performance.avgCpuUsagePct	• GE • LE • GT • LT	• Percentage	None
server.performance.totalDiskFreeSizeInKB	• GE • LE • GT • LT	• Double	None

Filter	Supported conditions	Supported values	Supported sorting
server.performance.avgFreeRAMInKB	• GE • LE • GT • LT	• Double	None
server.tag.value	• EQUALS • NOT_EQUALS • EQ • NE • CONTAINS • NOT_CONTAINS	• String	None
server.tag.key	• EQUALS • NOT_EQUALS • EQ • NE • CONTAINS • NOT_CONTAINS	• String	None
server.application.name	• EQUALS • NOT_EQUALS • EQ • NE • CONTAINS • NOT_CONTAINS	• String	None

Filter	Supported conditions	Supported values	Supported sorting
`server.application.description`	• EQUALS • NOT_EQUALS • EQ • NE • CONTAINS • NOT_CONTAINS	• String	None
`server.application.configurationId`	• EQUALS • NOT_EQUALS • EQ • NE	• Any valid application configuration ID	None
`server.process.configurationId`	• EQUALS • NOT_EQUALS • EQ • NE	• ProcessId	None
`server.process.name`	• EQUALS • NOT_EQUALS • EQ • NE • CONTAINS • NOT_CONTAINS	• String	None
`server.process.commandLine`	• EQUALS • NOT_EQUALS • EQ • NE • CONTAINS • NOT_CONTAINS	• String	None

Application assets

Filter	Supported conditions	Supported values	Supported sorting
`application.configurationId`	• EQUALS • NOT_EQUALS • EQ • NE	• ApplicationId	None
`application.name`	• EQUALS • NOT_EQUALS • EQ • NE • CONTAINS • NOT_CONTAINS	• String	• ASC • DESC
`application.description`	• EQUALS • NOT_EQUALS • EQ • NE • CONTAINS • NOT_CONTAINS	• String	• ASC • DESC
`application.serverCount`	Filtering not supported.	Filtering not supported.	• ASC • DESC
`application.timeOfCreation`	Filtering not supported.	Filtering not supported.	• ASC • DESC

Filter	Supported conditions	Supported values	Supported sorting
`application.lastModifiedTime`	Filtering not supported.	Filtering not supported.	• ASC • DESC
`server.configurationId`	• EQUALS • NOT_EQUALS • EQ • NE	• ServerId	None

Process assets

Filter	Supported conditions	Supported values	Supported sorting
`process.configurationId`	• EQUALS • NOT_EQUALS • EQ • NE	• ProcessId	
`process.name`	• EQUALS • NOT_EQUALS • EQ • NE • CONTAINS • NOT_CONTAINS	• String	• ASC • DESC
`process.commandLine`	• EQUALS • NOT_EQUALS • EQ • NE	• String	• ASC • DESC

Filter	Supported conditions	Supported values	Supported sorting
	• CONTAINS • NOT_CONTAINS		
server.configurationId	• EQUALS • NOT_EQUALS • EQ • NE	• ServerId	
server.hostName	• EQUALS • NOT_EQUALS • EQ • NE • CONTAINS • NOT_CONTAINS	• String	• ASC • DESC
server.osName	• EQUALS • NOT_EQUALS • EQ • NE • CONTAINS • NOT_CONTAINS	• String	• ASC • DESC
server.osVersion	• EQUALS • NOT_EQUALS • EQ • NE • CONTAINS • NOT_CONTAINS	• String	• ASC • DESC

Filter	Supported conditions	Supported values	Supported sorting
server.agentId	• EQUALS • NOT_EQUALS • EQ • NE • CONTAINS • NOT_CONTAINS	• String	

Connection assets

Filter	Supported conditions	Supported values	Supported sorting
connection.sourceIp	• EQUALS • NOT_EQUALS • EQ • NE • CONTAINS • NOT_CONTAINS	• IP	• ASC • DESC
connection.destinationIp	• EQUALS • NOT_EQUALS • EQ • NE • CONTAINS • NOT_CONTAINS	• IP	• ASC • DESC
connection.destinationPort	• EQUALS • NOT_EQUALS • EQ • NE	• Integer	• ASC • DESC

Filter	Supported conditions	Supported values	Supported sorting
sourceServer.configurationId	• EQUALS • NOT_EQUALS • EQ • NE	• ServerId	
sourceServer.hostName	• EQUALS • NOT_EQUALS • EQ • NE • CONTAINS • NOT_CONTAINS	• String	• ASC • DESC
destinationServer.osName	• EQUALS • NOT_EQUALS • EQ • NE • CONTAINS • NOT_CONTAINS	• String	• ASC • DESC
destinationServer.osVersion	• EQUALS • NOT_EQUALS • EQ • NE • CONTAINS • NOT_CONTAINS	• String	• ASC • DESC

Filter	Supported conditions	Supported values	Supported sorting
destinationServer.agentId	• EQUALS • NOT_EQUALS • EQ • NE • CONTAINS • NOT_CONTAINS	• String	
sourceProcess.configurationId	• EQUALS • NOT_EQUALS • EQ • NE	• ProcessId	
sourceProcess.name	• EQUALS • NOT_EQUALS • EQ • NE • CONTAINS • NOT_CONTAINS	• String	• ASC • DESC
sourceProcess.commandLine	• EQUALS • NOT_EQUALS • EQ • NE • CONTAINS • NOT_CONTAINS	• String	• ASC • DESC
destinationProcess.configurationId	• EQUALS • NOT_EQUALS • EQ • NE	• ProcessId	

Filter	Supported conditions	Supported values	Supported sorting
`destinationProcess.name`	• EQUALS • NOT_EQUALS • EQ • NE • CONTAINS • NOT_CONTAINS	• String	• ASC • DESC
`destinationprocess.commandLine`	• EQUALS • NOT_EQUALS • EQ • NE • CONTAINS • NOT_CONTAINS	• String	• ASC • DESC

Eventual consistency in the AWS Application Discovery Service API

The following update operations are eventually consistent. Updates might not be immediately visible to the read operations StartExportTask, DescribeConfigurations, and ListConfigurations.

- AssociateConfigurationItemsToApplication
- CreateTags
- DeleteApplications
- DeleteTags
- DescribeBatchDeleteConfigurationTask
- DescribeImportTasks
- DisassociateConfigurationItemsFromApplication
- UpdateApplication

Suggestions for managing eventual consistency:

- When you invoke the read operations StartExportTask, DescribeConfigurations, or ListConfigurations (or their corresponding AWS CLI commands), use an exponential backoff algorithm to allow enough time for any previous update operation to propagate through the system. To do this, run the read operation repeatedly, starting with a two-second wait time, and increasing gradually up to five minutes of wait time.
- Add wait time between subsequent operations, even if an update operation returns a 200 - OK response. Apply an exponential backoff algorithm starting with a couple of seconds of wait time, and increase gradually up to about five minutes of wait time.

Access AWS Application Discovery Service using an interface endpoint (AWS PrivateLink)

You can use AWS PrivateLink to create a private connection between your VPC and AWS Application Discovery Service. You can access Application Discovery Service as if it were in your VPC, without the use of an internet gateway, NAT device, VPN connection, or Direct Connect connection. Instances in your VPC don't need public IP addresses to access Application Discovery Service.

You establish this private connection by creating an *interface endpoint*, powered by AWS PrivateLink. We create an endpoint network interface in each subnet that you enable for the interface endpoint. These are requester-managed network interfaces that serve as the entry point for traffic destined for Application Discovery Service.

For more information, see Access AWS services through AWS PrivateLink in the *AWS PrivateLink Guide*.

Considerations for Application Discovery Service

Before you set up an interface endpoint for Application Discovery Service, review Access an AWS service using an interface VPC endpoint in the *AWS PrivateLink Guide*.

Application Discovery Service supports two interfaces: One for making calls to all of its API actions, and a second one for the Agentless Collector and AWS Application Discovery Agent to send discovery data.

Create an interface endpoint

You can create an interface endpoint using either the Amazon VPC console or the AWS Command Line Interface (AWS CLI). For more information, see Access an AWS service using an interface VPC endpoint in the *AWS PrivateLink Guide*.

For Application Discovery Service

Create an interface endpoint for Application Discovery Service using the following service name:

```
com.amazonaws.region.discovery
```

If you enable private DNS for the interface endpoint, you can make API requests to Application Discovery Service using its default Regional DNS name. For example, `discovery.us-east-1.amazonaws.com`.

For Agentless Collector and AWS Application Discovery Agent

Create an interface endpoint using the following service name:

```
com.amazonaws.region.arsenal-discovery
```

If you enable private DNS for the interface endpoint, you can make API requests to Application Discovery Arsenal using its default Regional DNS name. For example, `arsenal-discovery.us-east-1.amazonaws.com`.

Create an endpoint policy for your interface endpoint

An endpoint policy is an IAM resource that you can attach to an interface endpoint. The default endpoint policy allows full access to an AWS service through the interface endpoint. To control the access allowed to an AWS service from your VPC, attach a custom endpoint policy to the interface endpoint.

An endpoint policy specifies the following information:

- The principals that can perform actions (AWS accounts, IAM users, and IAM roles).
- The actions that can be performed.

For more information, see Control access to services using endpoint policies in the *AWS PrivateLink Guide*.

Example: VPC endpoint policies

The following is an example of a custom endpoint policy. When you attach this policy to your interface endpoint, it grants access to the listed actions for all principals on all resources.

Example policy for Application Discovery Service

```
{
```

```
   "Statement": [
      {
         "Principal": "*",
         "Effect": "Allow",
         "Action": [
            "discovery:action-1",
            "discovery:action-2",
            "discovery:action-3"
         ],
         "Resource":"*"
      }
   ]
}
```

Example policy for the Agentless Collector and AWS Application Discovery Agent

```
{
   "Statement": [
      {
         "Principal": "*",
         "Effect": "Allow",
         "Action": [
            "arsenal:RegisterOnPremisesAgent"
         ],
         "Resource":"*"
      }
   ]
}
```

Using the VPC endpoint for the Agentless Collector and AWS Application Discovery Agent

The Agentless Collector and AWS Application Discovery Agent don't support configurable endpoints. Instead, use the private DNS feature for the `arsenal-discovery` Amazon VPC endpoint.

- Set up the Direct Connect route table to route private AWS IP addresses to the VPC. For example, destination = 10.0.0.0/8 and target = local. For this setup you need at least routing for the `arsenal-discovery` Amazon VPC endpoint private IP addresses to the VPC.

- Use the `arsenal-discovery` Amazon VPC endpoint private DNS feature because the Agentless Collector doesn't support configurable Arsenal endpoints.
- Set up the `arsenal-discovery` Amazon VPC endpoint in a private subnet with the same VPC to which you are routing the Direct Connect traffic.
- Set up the `arsenal-discovery` Amazon VPC endpoint with a security group that enables inbound traffic from within the VPC (for example, 10.0.0.0/8).
- Set up an Amazon Route 53 inbound resolver to route DNS resolution for the `arsenal-discovery` Amazon VPC endpoint private DNS name, which will resolve to the private IP of the VPC endpoint. If you don't do that, the collector will perform DNS resolution by using the on-premises resolver and will use the public Arsenal endpoint, and traffic will not go through the VPC.
- If you have all public traffic disabled, the auto-update feature will fail. That is because the Agentless Collector retrieves updates by sending requests to the Amazon ECR endpoint. To get the auto-update feature working without sending requests over the public internet, set up a VPC endpoint for the Amazon ECR service and enable the private DNS feature for this endpoint.

Security in AWS Application Discovery Service

Cloud security at AWS is the highest priority. As an AWS customer, you benefit from a data center and network architecture that are built to meet the requirements of the most security-sensitive organizations.

Security is a shared responsibility between AWS and you. The shared responsibility model describes this as security *of* the cloud and security *in* the cloud:

- **Security of the cloud** – AWS is responsible for protecting the infrastructure that runs AWS services in the AWS Cloud. AWS also provides you with services that you can use securely. The effectiveness of our security is regularly tested and verified by third-party auditors as part of the AWS compliance programs.
- **Security in the cloud** – Your responsibility is determined by the AWS service that you use. You are also responsible for other factors including the sensitivity of your data, your organization's requirements, and applicable laws and regulations.

To use the AWS Application Discovery Agent or the Application Discovery Service Agentless Collector you must provide access keys to your AWS account. This information is then stored on your local infrastructure. As part of the shared responsibility model, you are responsible for securing access to your infrastructure.

This documentation will help you understand how to apply the shared responsibility model when using Application Discovery Service. The following topics show you how to configure Application Discovery Service to meet your security and compliance objectives. You'll also learn how to use other AWS services that can help you to monitor and secure your Application Discovery Service resources.

Topics

Identity and Access Management for AWS Application Discovery Service

AWS Identity and Access Management (IAM) is an AWS service that helps an administrator securely control access to AWS resources. IAM administrators control who can be *authenticated* (signed in) and *authorized* (have permissions) to use Application Discovery Service resources. IAM is an AWS service that you can use with no additional charge.

Topics

Audience

How you use AWS Identity and Access Management (IAM) differs based on your role:

- **Service user** - request permissions from your administrator if you cannot access features (see Troubleshooting AWS Application Discovery Service Identity and Access)
- **Service administrator** - determine user access and submit permission requests (see How AWS Application Discovery Service works with IAM)
- **IAM administrator** - write policies to manage access (see AWS Application Discovery Service identity-based policy examples)

Authenticating with identities

Authentication is how you sign in to AWS using your identity credentials. You must be authenticated as the AWS account root user, an IAM user, or by assuming an IAM role.

You can sign in as a federated identity using credentials from an identity source like AWS IAM Identity Center (IAM Identity Center), single sign-on authentication, or Google/Facebook credentials. For more information about signing in, see How to sign in to your AWS account in the *AWS Sign-In User Guide*.

For programmatic access, AWS provides an SDK and CLI to cryptographically sign requests. For more information, see AWS Signature Version 4 for API requests in the *IAM User Guide*.

AWS account root user

When you create an AWS account, you begin with one sign-in identity called the AWS account *root user* that has complete access to all AWS services and resources. We strongly recommend that you don't use the root user for everyday tasks. For tasks that require root user credentials, see Tasks that require root user credentials in the *IAM User Guide*.

IAM users and groups

An *IAM user* is an identity with specific permissions for a single person or application. We recommend using temporary credentials instead of IAM users with long-term credentials. For more information, see Require human users to use federation with an identity provider to access AWS using temporary credentials in the *IAM User Guide*.

An *IAM group* specifies a collection of IAM users and makes permissions easier to manage for large sets of users. For more information, see Use cases for IAM users in the *IAM User Guide*.

IAM roles

An *IAM role* is an identity with specific permissions that provides temporary credentials. You can assume a role by switching from a user to an IAM role (console) or by calling an AWS CLI or AWS API operation. For more information, see Methods to assume a role in the *IAM User Guide*.

IAM roles are useful for federated user access, temporary IAM user permissions, cross-account access, cross-service access, and applications running on Amazon EC2. For more information, see Cross account resource access in IAM in the *IAM User Guide*.

Managing access using policies

You control access in AWS by creating policies and attaching them to AWS identities or resources. A policy defines permissions when associated with an identity or resource. AWS evaluates these policies when a principal makes a request. Most policies are stored in AWS as JSON documents. For

more information about JSON policy documents, see Overview of JSON policies in the *IAM User Guide*.

Using policies, administrators specify who has access to what by defining which **principal** can perform **actions** on what **resources**, and under what **conditions**.

By default, users and roles have no permissions. An IAM administrator creates IAM policies and adds them to roles, which users can then assume. IAM policies define permissions regardless of the method used to perform the operation.

Identity-based policies

Identity-based policies are JSON permissions policy documents that you attach to an identity (user, group, or role). These policies control what actions identities can perform, on which resources, and under what conditions. To learn how to create an identity-based policy, see Define custom IAM permissions with customer managed policies in the *IAM User Guide*.

Identity-based policies can be *inline policies* (embedded directly into a single identity) or *managed policies* (standalone policies attached to multiple identities). To learn how to choose between managed and inline policies, see Choose between managed policies and inline policies in the *IAM User Guide*.

Resource-based policies

Resource-based policies are JSON policy documents that you attach to a resource. Examples include IAM *role trust policies* and Amazon S3 *bucket policies*. In services that support resource-based policies, service administrators can use them to control access to a specific resource. You must specify a principal in a resource-based policy.

Resource-based policies are inline policies that are located in that service. You can't use AWS managed policies from IAM in a resource-based policy.

Access control lists (ACLs)

Access control lists (ACLs) control which principals (account members, users, or roles) have permissions to access a resource. ACLs are similar to resource-based policies, although they do not use the JSON policy document format.

Amazon S3, AWS WAF, and Amazon VPC are examples of services that support ACLs. To learn more about ACLs, see Access control list (ACL) overview in the *Amazon Simple Storage Service Developer Guide*.

Other policy types

AWS supports additional policy types that can set the maximum permissions granted by more common policy types:

- **Permissions boundaries** – Set the maximum permissions that an identity-based policy can grant to an IAM entity. For more information, see Permissions boundaries for IAM entities in the *IAM User Guide*.
- **Service control policies (SCPs)** – Specify the maximum permissions for an organization or organizational unit in AWS Organizations. For more information, see Service control policies in the *AWS Organizations User Guide*.
- **Resource control policies (RCPs)** – Set the maximum available permissions for resources in your accounts. For more information, see Resource control policies (RCPs) in the *AWS Organizations User Guide*.
- **Session policies** – Advanced policies passed as a parameter when creating a temporary session for a role or federated user. For more information, see Session policies in the *IAM User Guide*.

Multiple policy types

When multiple types of policies apply to a request, the resulting permissions are more complicated to understand. To learn how AWS determines whether to allow a request when multiple policy types are involved, see Policy evaluation logic in the *IAM User Guide*.

How AWS Application Discovery Service works with IAM

Before you use IAM to manage access to Application Discovery Service, you should understand what IAM features are available to use with Application Discovery Service. To get a high-level view of how Application Discovery Service and other AWS services work with IAM, see AWS Services That Work with IAM in the *IAM User Guide*.

Topics

- Application Discovery Service identity-based policies
- Application Discovery Service resource-based policies
- Authorization based on Application Discovery Service tags
- Application Discovery Service IAM roles

Application Discovery Service identity-based policies

With IAM identity-based policies, you can specify allowed or denied actions and resources as well as the conditions under which actions are allowed or denied. Application Discovery Service supports specific actions, resources, and condition keys. To learn about all of the elements that you use in a JSON policy, see IAM JSON Policy Elements Reference in the *IAM User Guide*.

Actions

Administrators can use AWS JSON policies to specify who has access to what. That is, which **principal** can perform **actions** on what **resources**, and under what **conditions**.

The `Action` element of a JSON policy describes the actions that you can use to allow or deny access in a policy. Include actions in a policy to grant permissions to perform the associated operation.

Policy actions in Application Discovery Service use the following prefix before the action: `discovery:`. Policy statements must include either an `Action` or `NotAction` element. Application Discovery Service defines its own set of actions that describe tasks that you can perform with this service.

To specify multiple actions in a single statement, separate them with commas as follows:

```
"Action": [
      "discovery:action1",
      "discovery:action2"
```

You can specify multiple actions using wildcards (*). For example, to specify all actions that begin with the word `Describe`, include the following action:

```
"Action": "discovery:Describe*"
```

To see a list of Application Discovery Service actions, see Actions Defined by AWS Application Discovery Service in the *IAM User Guide*.

Resources

Application Discovery Service does not support specifying resource ARNs in a policy. To separate access, create and use separate AWS accounts.

Condition keys

Application Discovery Service does not provide any service-specific condition keys, but it does support using some global condition keys. To see all AWS global condition keys, see AWS Global Condition Context Keys in the *IAM User Guide*.

Examples

To view examples of Application Discovery Service identity-based policies, see AWS Application Discovery Service identity-based policy examples.

Application Discovery Service resource-based policies

Application Discovery Service does not support resource-based policies.

Authorization based on Application Discovery Service tags

Application Discovery Service does not support tagging resources or controlling access based on tags.

Application Discovery Service IAM roles

An IAM role is an entity within your AWS account that has specific permissions.

Using temporary credentials with Application Discovery Service

Application Discovery Service does not support using temporary credentials.

Service-linked roles

Service-linked roles allow AWS services to access resources in other services to complete an action on your behalf. Service-linked roles appear in your IAM account and are owned by the service. An IAM administrator can view but not edit the permissions for service-linked roles.

Application Discovery Service supports service-linked roles. For details about creating or managing Application Discovery Service service-linked roles, see Using service-linked roles for Application Discovery Service.

Service roles

This feature allows a service to assume a service role on your behalf. This role allows the service to access resources in other services to complete an action on your behalf. Service roles appear in your

IAM account and are owned by the account. This means that an IAM administrator can change the permissions for this role. However, doing so might break the functionality of the service.

Application Discovery Service supports service roles.

AWS managed policies for AWS Application Discovery Service

To add permissions to users, groups, and roles, it is easier to use AWS managed policies than to write policies yourself. It takes time and expertise to create IAM customer managed policies that provide your team with only the permissions they need. To get started quickly, you can use our AWS managed policies. These policies cover common use cases and are available in your AWS account. For more information about AWS managed policies, see AWS managed policies in the *IAM User Guide*.

AWS services maintain and update AWS managed policies. You can't change the permissions in AWS managed policies. Services occasionally add additional permissions to an AWS managed policy to support new features. This type of update affects all identities (users, groups, and roles) where the policy is attached. Services are most likely to update an AWS managed policy when a new feature is launched or when new operations become available. Services do not remove permissions from an AWS managed policy, so policy updates won't break your existing permissions.

Additionally, AWS supports managed policies for job functions that span multiple services. For example, the **ReadOnlyAccess** AWS managed policy provides read-only access to all AWS services and resources. When a service launches a new feature, AWS adds read-only permissions for new operations and resources. For a list and descriptions of job function policies, see AWS managed policies for job functions in the *IAM User Guide*.

AWS managed policy: AWSApplicationDiscoveryServiceFullAccess

The `AWSApplicationDiscoveryServiceFullAccess` policy grants an IAM user account access to the Application Discovery Service and Migration Hub APIs.

An IAM user account with this policy attached can configure Application Discovery Service, start and stop agents, start and stop agentless discovery, and query data from the AWS Discovery

Service database. For an example of this policy, see Granting full access to Application Discovery Service.

AWS managed policy: AWSApplicationDiscoveryAgentlessCollectorAccess

The `AWSApplicationDiscoveryAgentlessCollectorAccess` managed policy grants the Application Discovery Service Agentless Collector (Agentless Collector) access to register and communicate with the Application Discovery Service, and communicate with other AWS services.

This policy must be attached to the IAM user whose credentials are used to configure the Agentless Collector.

Permissions details

This policy includes the following permissions.

- `arsenal` – Allows the collector to register with the Application Discovery Service application. This is necessary to be able to send collected data back to AWS.
- `ecr-public` – Allows the collector to make calls to the Amazon Elastic Container Registry Public (Amazon ECR Public) where the latest updates are found for the collector.
- `mgh` – Allows the collector to call AWS Migration Hub to retrieve the home region of the account used to configure the collector. This is necessary to know which region the collected data should be sent to.
- `sts` – Allows the collector to retrieve a service bearer token so that the collector can make calls to Amazon ECR Public to get the latest updates.

JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "arsenal:RegisterOnPremisesAgent"
            ],
            "Resource": "*"
        },
        {
```

```
            "Effect": "Allow",
            "Action": [
                "ecr-public:DescribeImages"
            ],
            "Resource": "arn:aws:ecr-
public::446372222237:repository/6e5498e4-8c31-4f57-9991-13b4b992ff7b"
        },
        {
            "Effect": "Allow",
            "Action": [
                "ecr-public:GetAuthorizationToken"

            ],
            "Resource": "*"
        },
        {
            "Effect": "Allow",
            "Action": [
                "mgh:GetHomeRegion"
            ],
            "Resource": "*"
        },
        {
            "Effect": "Allow",
            "Action": [
                "sts:GetServiceBearerToken"
            ],
            "Resource": "*"
        }
    ]
}
```

AWS managed policy: AWSApplicationDiscoveryAgentAccess

The `AWSApplicationDiscoveryAgentAccess` policy grants the Application Discovery Agent access to register and communicate with Application Discovery Service.

You attach this policy to any user whose credentials are used by Application Discovery Agent.

This policy also grants the user access to Arsenal. Arsenal is an agent service that is managed and hosted by AWS. Arsenal forwards data to Application Discovery Service in the cloud. For an example of this policy, see Granting access to discovery agents.

AWS managed policy: AWSAgentlessDiscoveryService

The `AWSAgentlessDiscoveryService` policy grants the AWS Agentless Discovery Connector that is running in your VMware vCenter Server access to register, communicate with, and share connector health metrics with Application Discovery Service.

You attach this policy to any user whose credentials are used by the connector.

AWS managed policy: ApplicationDiscoveryServiceContinuousExportServiceRolePolicy

If your IAM account has the `AWSApplicationDiscoveryServiceFullAccess` policy attached, `ApplicationDiscoveryServiceContinuousExportServiceRolePolicy` is automatically attached to your account when you turn on data exploration in Amazon Athena.

This policy allows AWS Application Discovery Service to create Amazon Data Firehose streams to transform and deliver data that's collected by AWS Application Discovery Service agents to an Amazon S3 bucket in your AWS account.

In addition, this policy creates an AWS Glue Data Catalog with a new database called *application_discovery_service_database* and table schemas for mapping data that's collected by the agents. For an example of this policy, see Granting permissions for agent data collection.

AWS managed policy: AWSDiscoveryContinuousExportFirehosePolicy

The `AWSDiscoveryContinuousExportFirehosePolicy` policy is required to use data exploration in Amazon Athena. It allows Amazon Data Firehose to write data that's collected from Application Discovery Service to Amazon S3. For information about using this policy, see Creating the AWSApplicationDiscoveryServiceFirehose role. For an example of this policy, see Granting permissions for data exploration.

Creating the AWSApplicationDiscoveryServiceFirehose role

An administrator attaches managed policies to your IAM user account. When using the `AWSDiscoveryContinuousExportFirehosePolicy` policy, the administrator must first create a role named **AWSApplicationDiscoveryServiceFirehose** with Firehose as a trusted entity and then attach the `AWSDiscoveryContinuousExportFirehosePolicy` policy to the role, as shown in the following procedure.

To create the AWSApplicationDiscoveryServiceFirehose IAM role

1. In the IAM console, choose **Roles** on the navigation pane.
2. Choose **Create Role**.
3. Choose **Kinesis**.
4. Choose **Kinesis Firehose** as your use case.
5. Choose **Next: Permissions**.
6. Under **Filter Policies** search for **AWSDiscoveryContinuousExportFirehosePolicy**.
7. Select the box beside **AWSDiscoveryContinuousExportFirehosePolicy**, and then choose **Next: Review**.
8. Enter **AWSApplicationDiscoveryServiceFirehose** as the role name, and then choose **Create role**.

Application Discovery Service updates to AWS managed policies

View details about updates to AWS managed policies for Application Discovery Service since this service began tracking these changes. For automatic alerts about changes to this page, subscribe to the RSS feed on the Document History for AWS Application Discovery Service page.

Change	Description	Date
AWSApplicationDiscoveryAgentlessCollectorAccess – New policy made available with the Agentless Collector launch	Application Discovery Service added the new managed policy `AWSApplicationDiscoveryAgentlessCollectorAccess` that grants the Agentless Collector access to register and communicate with the Application Discovery Service, and communicate with other AWS services.	August 16, 2022

Change	Description	Date
Application Discovery Service started tracking changes	Application Discovery Service started tracking changes for its AWS managed policies.	March 1, 2021

AWS Application Discovery Service identity-based policy examples

By default, IAM users and roles don't have permission to create or modify Application Discovery Service resources. They also can't perform tasks using the AWS Management Console, AWS CLI, or AWS API. An IAM administrator must create IAM policies that grant users and roles permission to perform specific API operations on the specified resources they need. The administrator must then attach those policies to the IAM users or groups that require those permissions.

To learn how to create an IAM identity-based policy using these example JSON policy documents, see Creating Policies on the JSON Tab in the *IAM User Guide*.

Topics

- Policy best practices
- Granting full access to Application Discovery Service
- Granting access to discovery agents
- Granting permissions for agent data collection
- Granting permissions for data exploration
- Granting permissions to use the Migration Hub console network diagram

Policy best practices

Identity-based policies determine whether someone can create, access, or delete Application Discovery Service resources in your account. These actions can incur costs for your AWS account. When you create or edit identity-based policies, follow these guidelines and recommendations:

- **Get started with AWS managed policies and move toward least-privilege permissions** – To get started granting permissions to your users and workloads, use the *AWS managed policies* that grant permissions for many common use cases. They are available in your AWS account. We recommend that you reduce permissions further by defining AWS customer managed policies

that are specific to your use cases. For more information, see AWS managed policies or AWS managed policies for job functions in the *IAM User Guide*.

- **Apply least-privilege permissions** – When you set permissions with IAM policies, grant only the permissions required to perform a task. You do this by defining the actions that can be taken on specific resources under specific conditions, also known as *least-privilege permissions*. For more information about using IAM to apply permissions, see Policies and permissions in IAM in the *IAM User Guide*.
- **Use conditions in IAM policies to further restrict access** – You can add a condition to your policies to limit access to actions and resources. For example, you can write a policy condition to specify that all requests must be sent using SSL. You can also use conditions to grant access to service actions if they are used through a specific AWS service, such as CloudFormation. For more information, see IAM JSON policy elements: Condition in the *IAM User Guide*.
- **Use IAM Access Analyzer to validate your IAM policies to ensure secure and functional permissions** – IAM Access Analyzer validates new and existing policies so that the policies adhere to the IAM policy language (JSON) and IAM best practices. IAM Access Analyzer provides more than 100 policy checks and actionable recommendations to help you author secure and functional policies. For more information, see Validate policies with IAM Access Analyzer in the *IAM User Guide*.
- **Require multi-factor authentication (MFA)** – If you have a scenario that requires IAM users or a root user in your AWS account, turn on MFA for additional security. To require MFA when API operations are called, add MFA conditions to your policies. For more information, see Secure API access with MFA in the *IAM User Guide*.

For more information about best practices in IAM, see Security best practices in IAM in the *IAM User Guide*.

Granting full access to Application Discovery Service

The AWSApplicationDiscoveryServiceFullAccess managed policy grants the IAM user account access to the Application Discovery Service and Migration Hub APIs.

An IAM user with this policy attached to their account can configure Application Discovery Service, start and stop agents, start and stop agentless discovery, and query data from the AWS Discovery Service database. For more information about this policy, see AWS managed policies for AWS Application Discovery Service.

Example AWSApplicationDiscoveryServiceFullAccess policy

JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Action": [
                "mgh:*",
                "discovery:*"
            ],
            "Effect": "Allow",
            "Resource": "*"
        },
        {
            "Action": [
                "iam:GetRole"
            ],
            "Effect": "Allow",
            "Resource": "*"
        }
    ]
}
```

Granting access to discovery agents

The AWSApplicationDiscoveryAgentAccess managed policy grants the Application Discovery Agent access to register and communicate with Application Discovery Service. For more information about this policy, see AWS managed policies for AWS Application Discovery Service.

Attach this policy to any user whose credentials are used by Application Discovery Agent.

This policy also grants the user access to Arsenal. Arsenal is an agent service that is managed and hosted by AWS. Arsenal forwards data to Application Discovery Service in the cloud.

Example AWSApplicationDiscoveryAgentAccess Policy

JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "arsenal:RegisterOnPremisesAgent"
            ],
            "Resource": "*"
        }
    ]
}
```

Granting permissions for agent data collection

The ApplicationDiscoveryServiceContinuousExportServiceRolePolicy managed policy allows AWS Application Discovery Service to create Amazon Data Firehose streams to transform and deliver data that's collected by Application Discovery Service agents to an Amazon S3 bucket in your AWS account.

In addition, this policy creates an AWS Glue Data Catalog with a new database called `application_discovery_service_database` and table schemas for mapping data that's collected by the agents.

For information about using this policy, see AWS managed policies for AWS Application Discovery Service.

Example ApplicationDiscoveryServiceContinuousExportServiceRolePolicy

JSON

```
{
   "Version":"2012-10-17",
    "Statement": [
        {
            "Action": [
```

```
                "glue:CreateDatabase",
                "glue:UpdateDatabase",
                "glue:CreateTable",
                "glue:UpdateTable",
                "firehose:CreateDeliveryStream",
                "firehose:DescribeDeliveryStream",
                "logs:CreateLogGroup"
            ],
            "Effect": "Allow",
            "Resource": "*"
        },
        {
            "Action": [
                "firehose:DeleteDeliveryStream",
                "firehose:PutRecord",
                "firehose:PutRecordBatch",
                "firehose:UpdateDestination"
            ],
            "Effect": "Allow",
            "Resource": "arn:aws:firehose:*:*:deliverystream/aws-application-
discovery-service*"
        },
        {
            "Action": [
                "s3:CreateBucket",
                "s3:ListBucket",
                "s3:PutBucketLogging",
                "s3:PutEncryptionConfiguration"
            ],
            "Effect": "Allow",
            "Resource": "arn:aws:s3:::aws-application-discovery-service*"
        },
        {
            "Action": [
                "s3:GetObject"
            ],
            "Effect": "Allow",
            "Resource": "arn:aws:s3:::aws-application-discovery-service*/*"
        },
        {
            "Action": [
                "logs:CreateLogStream",
                "logs:PutRetentionPolicy"
            ],
```

```
            "Effect": "Allow",
            "Resource": "arn:aws:logs:*:*:log-group:/aws/application-discovery-
service/firehose*"
        },
        {
            "Action": [
                "iam:PassRole"
            ],
            "Effect": "Allow",
            "Resource": "arn:aws:iam::*:role/
AWSApplicationDiscoveryServiceFirehose",
            "Condition": {
                "StringLike": {
                    "iam:PassedToService": "firehose.amazonaws.com"
                }
            }
        },
        {
            "Action": [
                "iam:PassRole"
            ],
            "Effect": "Allow",
            "Resource": "arn:aws:iam::*:role/service-role/
AWSApplicationDiscoveryServiceFirehose",
            "Condition": {
                "StringLike": {
                    "iam:PassedToService": "firehose.amazonaws.com"
                }
            }
        }
    ]
}
```

Granting permissions for data exploration

The AWSDiscoveryContinuousExportFirehosePolicy policy is required to use data exploration in Amazon Athena. It allows Amazon Data Firehose to write data that's collected from Application Discovery Service to Amazon S3. For information about using this policy, see Creating the AWSApplicationDiscoveryServiceFirehose role.

Example AWSDiscoveryContinuousExportFirehosePolicy

JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "glue:GetTableVersions"
            ],
            "Resource": "*"
        },
        {
            "Effect": "Allow",
            "Action": [
                "s3:AbortMultipartUpload",
                "s3:GetBucketLocation",
                "s3:GetObject",
                "s3:ListBucket",
                "s3:ListBucketMultipartUploads",
                "s3:PutObject"
            ],
            "Resource": [
                "arn:aws:s3:::aws-application-discovery-service-*",
                "arn:aws:s3:::aws-application-discovery-service-*/*"
            ]
        },
        {
            "Effect": "Allow",
            "Action": [
                "logs:PutLogEvents"
            ],
            "Resource": [
                "arn:aws:logs:*:*:log-group:/aws/application-discovery-service/
firehose:log-stream:*"
            ]
        }
    ]
}
```

Granting permissions to use the Migration Hub console network diagram

To grant access to the AWS Migration Hub console network diagram when creating an identity-based policy that allows or denies access to Application Discovery Service or Migration Hub, you might need to add the `discovery:GetNetworkConnectionGraph` action to the policy.

You must use the `discovery:GetNetworkConnectionGraph` action in new policies or update older policies if the following are both true for the policy:

- The policy allows or denies access to Application Discovery Service or the Migration Hub.
- The policy grants access permissions using one more specific discovery actions like `discovery:`*`action-name`* rather than `discovery:*`.

The following example shows how to use the `discovery:GetNetworkConnectionGraph` action in an IAM policy.

Example

JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": ["discovery:GetNetworkConnectionGraph"],
            "Resource": "*"
        }
    ]
}
```

For information about the Migration Hub network diagram, see Viewing network connections in Migration Hub.

Using service-linked roles for Application Discovery Service

AWS Application Discovery Service uses AWS Identity and Access Management (IAM) service-linked roles. A service-linked role is a unique type of IAM role that is linked directly to Application

Discovery Service. Service-linked roles are predefined by Application Discovery Service and include all the permissions that the service requires to call other AWS services on your behalf.

A service-linked role makes setting up Application Discovery Service easier because you don't have to manually add the necessary permissions. Application Discovery Service defines the permissions of its service-linked roles, and unless defined otherwise, only Application Discovery Service can assume its roles. The defined permissions include the trust policy and the permissions policy, and that permissions policy cannot be attached to any other IAM entity.

You can delete a service-linked role only after first deleting their related resources. This protects your Application Discovery Service resources because you can't inadvertently remove permission to access the resources.

Topics

- Service-linked role permissions for Application Discovery Service
- Creating a service-linked role for Application Discovery Service
- Deleting a service-linked role for Application Discovery Service

For information about other services that support service-linked roles, see AWS Services That Work with IAM and look for the services that have **Yes** in the **Service-Linked Role** column. Choose a **Yes** with a link to view the service-linked role documentation for that service.

Service-linked role permissions for Application Discovery Service

Application Discovery Service uses the service-linked role named **AWSServiceRoleForApplicationDiscoveryServiceContinuousExport** – Enables access to AWS Services and Resources used or managed by AWS Application Discovery Service.

The AWSServiceRoleForApplicationDiscoveryServiceContinuousExport service-linked role trusts the following services to assume the role:

- `continuousexport.discovery.amazonaws.com`

The role permissions policy allows Application Discovery Service to complete the following actions:

glue

```
CreateDatabase
```

- UpdateDatabase
- CreateTable
- UpdateTable

firehose

- CreateDeliveryStream
- DeleteDeliveryStream
- DescribeDeliveryStream
- PutRecord
- PutRecordBatch
- UpdateDestination

s3

- CreateBucket
- ListBucket
- GetObject

logs

- CreateLogGroup
- CreateLogStream
- PutRetentionPolicy

iam

- PassRole

This is the full policy showing which resources the above actions apply to:

JSON

```
{
    "Version":"2012-10-17",
```

```
    "Statement": [
        {
            "Action": [
                "glue:CreateDatabase",
                "glue:UpdateDatabase",
                "glue:CreateTable",
                "glue:UpdateTable",
                "firehose:CreateDeliveryStream",
                "firehose:DescribeDeliveryStream",
                "logs:CreateLogGroup"
            ],
            "Effect": "Allow",
            "Resource": "*"
        },
        {
            "Action": [
                "firehose:DeleteDeliveryStream",
                "firehose:PutRecord",
                "firehose:PutRecordBatch",
                "firehose:UpdateDestination"
            ],
            "Effect": "Allow",
            "Resource": "arn:aws:firehose:*:*:deliverystream/aws-application-
discovery-service*"
        },
        {
            "Action": [
                "s3:CreateBucket",
                "s3:ListBucket",
                "s3:PutBucketLogging",
                "s3:PutEncryptionConfiguration"
            ],
            "Effect": "Allow",
            "Resource": "arn:aws:s3:::aws-application-discovery-service*"
        },
        {
            "Action": [
                "s3:GetObject"
            ],
            "Effect": "Allow",
            "Resource": "arn:aws:s3:::aws-application-discovery-service*/*"
        },
        {
            "Action": [
```

```
                "logs:CreateLogStream",
                "logs:PutRetentionPolicy"
            ],
            "Effect": "Allow",
            "Resource": "arn:aws:logs:*:*:log-group:/aws/application-discovery-
service/firehose*"
        },
        {
            "Action": [
                "iam:PassRole"
            ],
            "Effect": "Allow",
            "Resource": "arn:aws:iam::*:role/
AWSApplicationDiscoveryServiceFirehose",
            "Condition": {
                "StringLike": {
                    "iam:PassedToService": "firehose.amazonaws.com"
                }
            }
        },
        {
            "Action": [
                "iam:PassRole"
            ],
            "Effect": "Allow",
            "Resource": "arn:aws:iam::*:role/service-role/
AWSApplicationDiscoveryServiceFirehose",
            "Condition": {
                "StringLike": {
                    "iam:PassedToService": "firehose.amazonaws.com"
                }
            }
        }
    ]
}
```

You must configure permissions to allow an IAM entity (such as a user, group, or role) to create, edit, or delete a service-linked role. For more information, see Service-Linked Role Permissions in the *IAM User Guide*.

Creating a service-linked role for Application Discovery Service

You don't need to manually create a service-linked role. The AWSServiceRoleForApplicationDiscoveryServiceContinuousExport service-linked role is automatically created when Continuous Export is implicitly turned on by a) confirming options in the dialog box presented from the Data Collectors page after you choose “Start data collection”, or click the slider labeled, “Data exploration in Athena”, or b) when you call the StartContinuousExport API using the AWS CLI.

> **Important**
>
> This service-linked role can appear in your account if you completed an action in another service that uses the features supported by this role. To learn more, see A New Role Appeared in My IAM Account.

Creating the service-linked role from the Migration Hub console

You can use the Migration Hub console to create the AWSServiceRoleForApplicationDiscoveryServiceContinuousExport service-linked role.

To create the service-linked role (console)

1. In the navigation pane, choose **Data Collectors**.
2. Choose the **Agents** tab.
3. Toggle the **Data exploration in Athena** slider to the On position.
4. In the dialog box generated from the previous step, click the checkbox agreeing to associated costs and choose **Continue** or **Enable**.

Creating the service-linked role from the AWS CLI

You can use Application Discovery Service commands from the AWS Command Line Interface to create the AWSServiceRoleForApplicationDiscoveryServiceContinuousExport service-linked role.

This service-linked role is automatically created when you start Continuous Export from the AWS CLI (the AWS CLI must first be installed in your environment).

To create the service-linked role (CLI) by starting Continuous Export from the AWS CLI

1. Install the AWS CLI for your operating system (Linux, macOS, or Windows). See the AWS Command Line Interface User Guide for instructions.
2. Open the Command prompt (Windows) or Terminal (Linux or macOS).
   a. Type `aws configure` and press Enter.
   b. Enter your AWS Access Key Id and AWS Secret Access Key.
   c. Enter `us-west-2` for the Default Region Name.
   d. Enter `text` for Default Output Format.
3. Type the following command:

```
aws discovery start-continuous-export
```

You can also use the IAM console to create a service-linked role with the **Discovery Service - Continuous Export** use case. In the IAM CLI or the IAM API, create a service-linked role with the `continuousexport.discovery.amazonaws.com` service name. For more information, see Creating a Service-Linked Role in the *IAM User Guide*. If you delete this service-linked role, you can use this same process to create the role again.

Deleting a service-linked role for Application Discovery Service

If you no longer need to use a feature or service that requires a service-linked role, we recommend that you delete that role. That way you don't have an unused entity that is not actively monitored or maintained. However, you must clean up your service-linked role before you can manually delete it.

Cleaning up the service-linked role

Before you can use IAM to delete a service-linked role, you must first delete any resources used by the role.

> **Note**
>
> If Application Discovery Service is using the role when you try to delete the resources, then the deletion might fail. If that happens, wait for a few minutes and try the operation again.

To delete Application Discovery Service resources used by the AWSServiceRoleForApplicationDiscoveryServiceContinuousExport service-linked role from the Migration Hub Console

1. In the navigation pane, choose **Data Collectors**.
2. Choose the **Agents** tab.
3. Toggle the **Data exploration in Athena** slider to the Off position.

To delete Application Discovery Service resources used by the AWSServiceRoleForApplicationDiscoveryServiceContinuousExport service-linked role from the AWS CLI

1. Install the AWS CLI for your operating system (Linux, macOS, or Windows). See the AWS Command Line Interface User Guide for instructions.
2. Open the Command prompt (Windows) or Terminal (Linux or macOS).
   a. Type `aws configure` and press Enter.
   b. Enter your AWS Access Key Id and AWS Secret Access Key.
   c. Enter `us-west-2` for the Default Region Name.
   d. Enter `text` for Default Output Format.
3. Type the following command:

```
aws discovery stop-continuous-export --export-id <export ID>
```

   - If you don't know the export-ID of the continuous export you want to stop, enter the following command to see the continuous export's ID:

```
aws discovery describe-continuous-exports
```

4. Enter the follow command to ensure that Continuous Export has stopped by verifying its return status is "INACTIVE":

```
aws discovery describe-continuous-export
```

Manually delete the service-linked role

You can delete the AWSServiceRoleForApplicationDiscoveryServiceContinuousExport service-linked role by using the IAM console, the IAM CLI, or the IAM API. If you no longer need to use the Discovery Service - Continuous Export features that require this service-linked role, we recommend that you delete that role. That way you don't have an unused entity that is not actively monitored or maintained. For more information, see Deleting a Service-Linked Role in the *IAM User Guide*.

> **Note**
>
> You must first clean up your service-linked role before you can delete it. See Cleaning up the service-linked role.

Troubleshooting AWS Application Discovery Service Identity and Access

Use the following information to help you diagnose and fix common issues that you might encounter when working with Application Discovery Service and IAM.

Topics

- I Am Not Authorized to Perform iam:PassRole

I Am Not Authorized to Perform iam:PassRole

If you receive an error that you're not authorized to perform the `iam:PassRole` action, your policies must be updated to allow you to pass a role to Application Discovery Service.

Some AWS services allow you to pass an existing role to that service instead of creating a new service role or service-linked role. To do this, you must have permissions to pass the role to the service.

The following example error occurs when an IAM user named `marymajor` tries to use the console to perform an action in Application Discovery Service. However, the action requires the service to have permissions that are granted by a service role. Mary does not have permissions to pass the role to the service.

```
User: arn:aws:iam::123456789012:user/marymajor is not authorized to perform:
 iam:PassRole
```

In this case, Mary's policies must be updated to allow her to perform the `iam:PassRole` action.

If you need help, contact your AWS administrator. Your administrator is the person who provided you with your sign-in credentials.

Logging Application Discovery Service API calls with AWS CloudTrail

AWS Application Discovery Service is integrated with AWS CloudTrail, a service that provides a record of actions taken by a user, role, or an AWS service in Application Discovery Service. You can use CloudTrail to log, continuously monitor, and retain account activity for troubleshooting and auditing purposes. CloudTrail provides an event history of your AWS account activity, including actions taken through the AWS Management Console, AWS SDKs, and command line tools.

CloudTrail captures all API calls for Application Discovery Service as events. The calls captured include calls from the Application Discovery Service console and code calls to the Application Discovery Service API operations.

If you create a trail, you can enable continuous delivery of CloudTrail events to an Amazon S3 bucket, including events for Application Discovery Service. If you don't configure a trail, you can still view the most recent events in the CloudTrail console in **Event history**. Using the information collected by CloudTrail, you can determine the request that was made to Application Discovery Service, the IP address from which the request was made, who made the request, when it was made, and additional details.

To learn more about CloudTrail, see the AWS CloudTrail User Guide.

Application Discovery Service information in CloudTrail

CloudTrail is enabled on your AWS account when you create the account. When activity occurs in Application Discovery Service, that activity is recorded in a CloudTrail event along with other AWS service events in **Event history**. You can view, search, and download recent events in your AWS account. For more information, see Viewing Events with CloudTrail Event History.

For an ongoing record of events in your AWS account, including events for Application Discovery Service, create a trail. A *trail* enables CloudTrail to deliver log files to an Amazon S3 bucket. By

default, when you create a trail in the console, the trail applies to all AWS Regions. The trail logs events from all Regions in the AWS partition and delivers the log files to the Amazon S3 bucket that you specify. Additionally, you can configure other AWS services to further analyze and act upon the event data that's collected in CloudTrail logs. For more information, see the following:

- Overview for Creating a Trail
- CloudTrail Supported Services and Integrations
- Configuring Amazon SNS Notifications for CloudTrail
- Receiving CloudTrail Log Files from Multiple Regions and Receiving CloudTrail Log Files from Multiple Accounts

All Application Discovery Service actions are logged by CloudTrail and are documented in the Application Discovery Service API Reference. For example, calls to the `CreateTags`, `DescribeTags`, and `GetDiscoverySummary` actions generate entries in the CloudTrail log files.

Every event or log entry contains information about who generated the request. The identity information helps you determine the following:

- Whether the request was made with root or AWS Identity and Access Management (IAM) user credentials.
- Whether the request was made with temporary security credentials for a role or federated user.
- Whether the request was made by another AWS service.

For more information, see the CloudTrail userIdentity Element.

Understanding Application Discovery Service log file entries

A trail is a configuration that enables delivery of events as log files to an Amazon S3 bucket that you specify. CloudTrail log files contain one or more log entries. An event represents a single request from any source and includes information about the requested action, the date and time of the action, request parameters, and so on. CloudTrail log files aren't an ordered stack trace of the public API calls, so they don't appear in any specific order.

The following example shows a CloudTrail log entry that demonstrates the `DescribeTags` action.

```
{
    "eventVersion": "1.05",
    "userIdentity": {
```

```
        "type": "AssumedRole",
        "principalId": "AROAJBHMC4H6EKEXAMPLE:sample-user",
        "arn": "arn:aws:sts::444455556666:assumed-role/ReadOnly/sample-user",
        "accountId": "123456789012",
        "accessKeyId": "AKIAIOSFODNN7EXAMPLE",
        "sessionContext": {
            "sessionIssuer": {
                "type": "Role",
                "principalId": "AIDAJQABLZS4A3QDU576Q",
                "arn": "arn:aws:iam::444455556666:role/ReadOnly",
                "accountId": "444455556666",
                "userName": "sampleAdmin"
            },
            "webIdFederationData": {},
            "attributes": {
                "mfaAuthenticated": "false",
                "creationDate": "2020-05-05T15:19:03Z"
            }
        }
    },
    "eventTime": "2020-05-05T17:02:40Z",
    "eventSource": "discovery.amazonaws.com",
    "eventName": "DescribeTags",
    "awsRegion": "us-west-2",
    "sourceIPAddress": "20.22.33.44",
    "userAgent": "Coral/Netty4",
    "requestParameters": {
        "maxResults": 0,
        "filters": [
            {
                "values": [
                    "d-server-0315rfdjreyqsq"
                ],
                "name": "configurationId"
            }
        ]
    },
    "responseElements": null,
    "requestID": "mgh-console-eb1cf315-e2b4-4696-93e5-b3a3b9346b4b",
    "eventID": "7b32b778-91c9-4c75-9cb0-6c852791b2eb",
    "eventType": "AwsApiCall",
    "recipientAccountId": "111122223333"
}
```

AWS Application Discovery Service ARN formats

An Amazon Resource Name (ARN) is a string that uniquely identifies an AWS resource. AWS requires an ARN when you want to specify a resource unambiguously across all of AWS. AWS Application Discovery Service defines the following ARNs.

- **Discovery Agent**: `arn:aws:discovery:`*`region`*`:`*`account`*`:agent/discovery-agent/`*`agentId`*
- **Agentless Collector**: `arn:aws:discovery:`*`region`*`:`*`account`*`:agent/agentless-collector/`*`agentId`*
- **Migration Evaluator Collector**: `arn:aws:discovery:`*`region`*`:`*`account`*`:agent/migration-evaluator-collector/`*`agentId`*
- **Discovery Connector**: `arn:aws:discovery:`*`region`*`:`*`account`*`:agent/discovery-connector/`*`agentId`*

AWS Application Discovery Service Quotas

The Service Quotas console provides information about AWS Application Discovery Service quotas. You can use the Service Quotas console to view the default service quotas or to request quota increases for adjustable quotas.

Currently, the only quota that can be increased is **imported servers per account**.

Application Discovery Service has the following default quotas:

- 1,000 applications per account.

  If you reach this quota, and want to import new applications, you can delete existing applications with the `DeleteApplications` API action. For more information, see DeleteApplications in the *Application Discovery Service API Reference*.
- Each import file can have a maximum file size of 10 MB.
- 25,000 imported server records per account.
- 25,000 deletions of import records per day.
- 10,000 imported servers per account (you can request to increase this quota).
- 1,000 active agents, which are collecting and sending data to Application Discovery Service.
- 10,000 inactive agents, which are responsive but not collecting data.
- 400 servers per application.
- 30 tags per server.

Troubleshooting AWS Application Discovery Service

In this section, you can find information about how to fix common issues with AWS Application Discovery Service.

Topics

Stop data collection by data exploration

To stop data exploration, you can either switch off the toggle switch in the Migration Hub console under Discover > Data Collectors > Agents tab, or invoke the `StopContinuousExport` API. It can take up to 30 minutes to stop the data collection, and during this stage, the toggle switch on the console and the `DescribeContinuousExport` API invocation will show the data exploration state as "Stop In Progress".

> **Note**
>
> If after refreshing the console page, the toggle does not switch off and an error message is thrown or the `DescribeContinuousExport` API returns "Stop_Failed" state, you can try again by switching the toggle switch off or calling the `StopContinuousExport` API. If the "data exploration" still shows error and fails to successfully stop, please reach out to AWS support.

Alternatively, you can manually stop data collection as described in the following steps.

Option 1: Stop Agent Data collection

If you have already completed your discovery using ADS agents and no longer want to collect additional data in the ADS database repository:

1. From the Migration Hub console choose Discover > Data Collectors > Agents tab.

2. Select all existing running agents and choose **Stop Data Collection**.

   This will ensure that no new data is being collected by the agents in both the ADS data repository and your S3 bucket. Your existing data remains accessible.

Option 2: Delete data exploration's Amazon Kinesis Data Streams

If you want to continue collecting data by agents in ADS data repository, but don't want to collect data in your Amazon S3 bucket using data exploration, you can manually delete the Amazon Data Firehose streams created by data exploration:

1. Log in to Amazon Kinesis from the AWS console and choose **Data Firehose** from the navigation pane.
2. Delete the following streams created by the data exploration feature:

   - `aws-application-discovery-service-id_mapping_agent`
   - `aws-application-discovery-service-inbound_connection_agent`
   - `aws-application-discovery-service-network_interface_agent`
   - `aws-application-discovery-service-os_info_agent`
   - `aws-application-discovery-service-outbound_connection_agent`
   - `aws-application-discovery-service-processes_agent`
   - `aws-application-discovery-service-sys_performance_agent`

Remove the data collected by data exploration

To remove data that's collected by data exploration

1. Remove the discovery agent data stored in Amazon S3.

   Data that's collected by AWS Application Discovery Service (ADS) is stored in an S3 bucket named `aws-application-discover-discovery-service-`*`uniqueid`*.

> **Note**
>
> Deleting the Amazon S3 bucket or any of the objects in it while data exploration in Amazon Athena is enabled causes an error. It continues to send new discovery agent data to S3. The deleted data will no longer be accessible in Athena as well.

2. Remove AWS Glue Data Catalog.

   When data exploration in Amazon Athena is turned on, it creates an Amazon S3 bucket in your account to store the data that's collected by ADS agents at regular time intervals. In addition, it also creates an AWS Glue Data Catalog to allow you to query the data stored in a Amazon S3 bucket from Amazon Athena. When you turn off data exploration in Amazon Athena, no new data is stored in your Amazon S3 bucket, but data that was collected previously will persist. If you no longer need this data and want to return your account to the state before data exploration in Amazon Athena was turned on.

   a. Visit Amazon S3 from the AWS console and manually delete the bucket with the name "aws-application-discover-discovery-service-uniqueid"

   b. You can manually remove the data exploration AWS Glue Data Catalog by deleting the *application-discovery-service-database* database and all of these tables:

      - `os_info_agent`
      - `network_interface_agent`
      - `sys_performance_agent`
      - `processes_agent`
      - `inbound_connection_agent`
      - `outbound_connection_agent`
      - `id_mapping_agent`

Removing your data from AWS Application Discovery Service

To have all your data removed from Application Discovery Service, contact AWS Support and request full data deletion.

Fix common issues with data exploration in Amazon Athena

In this section, you can find information about how to fix common issues with data exploration in Amazon Athena.

Topics

Data exploration in Amazon Athena fails to initiate because service-linked roles and required AWS resources can't be created

When you turn on data exploration in Amazon Athena, it creates the service-linked role, `AWSServiceRoleForApplicationDiscoveryServiceContinuousExport`, in your account that allows it to create the required AWS resources for making the agent collected data accessible in Amazon Athena including an Amazon S3 bucket, Amazon Kinesis streams, and AWS Glue Data Catalog. If your account does not have the right permissions for data exploration in Amazon Athena to create this role, it will fail to initialize. Refer to AWS managed policies for AWS Application Discovery Service.

New Agent data doesn't show up in Amazon Athena

If new data does not flow into Athena, it has been more than 30 minutes since an agent started, and data exploration status is Active, check the solutions listed below:

- AWS Discovery Agents

  Ensure that your agent's **Collection** status is marked as **Started** and the **Health** status is marked as **Running**.

- Kinesis Role

  Ensure that you have the `AWSApplicationDiscoveryServiceFirehose` role in your account.

- Firehose Status

Ensure that the following Firehose delivery streams are working correctly:

- `aws-application-discovery-service/os_info_agent`
- `aws-application-discovery-service-network_interface_agent`
- `aws-application-discovery-service-sys_performance_agent`
- `aws-application-discovery-service-processes_agent`
- `aws-application-discovery-service-inbound_connection_agent`
- `aws-application-discovery-service-outbound_connection_agent`
- `aws-application-discovery-service-id_mapping_agent`

- AWS Glue Data Catalog

  Ensure that the `application-discovery-service-database` database is in AWS Glue. Make sure that the following tables are present in AWS Glue:

  - `os_info_agent`
  - `network_interface_agent`
  - `sys_performance_agent`
  - `processes_agent`
  - `inbound_connection_agent`
  - `outbound_connection_agent`
  - `id_mapping_agent`

- Amazon S3 Bucket

  Ensure that you have an Amazon S3 bucket named `aws-application-discovery-service-`*`uniqueid`* in your account. If objects in the bucket have been moved or deleted, they will not show up properly in Athena.

- Your on-premises servers

  Ensure that your servers are running so that your agents can collect and send data to AWS Application Discovery Service.

You have insufficient permissions to access Amazon S3, Amazon Data Firehose, or AWS Glue

If you are using AWS Organizations, and initialization for data exploration in Amazon Athena fails, it can be because you don't have permissions to access Amazon S3, Amazon Data Firehose, Athena or AWS Glue.

You will need an IAM user with administrator permissions to grant you access to these services. An administrator can use their account to grant this access. See AWS managed policies for AWS Application Discovery Service.

To ensure that data exploration in Amazon Athena works correctly, do not modify or delete the AWS resources created by data exploration in Amazon Athena including the Amazon S3 bucket, Amazon Data Firehose Streams, and AWS Glue Data Catalog. If you accidentally delete or modify these resources, please stop and start Data Exploration and it will automatically create these resources again. If you delete the Amazon S3 bucket created by data exploration, you may lose the data that was collected in the bucket.

Troubleshooting failed import records

Migration Hub import allows you to import details of your on-premises environment directly into Migration Hub without using the Discovery Connector or Discovery Agent. This gives you the option to perform migration assessment and planning directly from your imported data. You can also group your devices as applications and track their migration status.

When importing data, it's possible that you'll encounter errors. Typically, these errors occur for one of the following reasons:

- **An import-related quota was reached** – There is a quota associated with import tasks. If you make an import task request that would exceeds the quotas, then the request will fail and return an error. For more information, see AWS Application Discovery Service Quotas.
- **An extra comma (,) was inserted into the import file** – Commas in .CSV files are used to differentiate one field from the next. Having a comma appear within a field is unsupported, because it will always split a field. This can cause a cascade of formatting errors. Be sure that commas are only used between fields, and are not otherwise used in your import files.
- **A field has a value outside of its supported range** – Some fields, like `CPU.NumberOfCores` must have a range of values they support. If you have more or less than this supported range, then the record will fail to be imported.

If any errors occur with your import request, you can resolve them by downloading your failed records for your import task, and resolve the errors in the failed entries CSV file, and do the import again.

Console

To download your failed records archive

1. Sign into the AWS Management Console, and open the Migration Hub console at https://console.aws.amazon.com/migrationhub.
2. From the left-side navigation, under **Discover**, choose **Tools**.
3. From **Discovery Tools**, choose **view imports**.
4. From the **Imports** dashboard, choose the radio button associated an import request with some number of **Failed records**.
5. Choose **Download failed records** from above the table on the dashboard. This will open your browser's download dialog box for downloading the archive file.

AWS CLI

To download your failed records archive

1. Open a terminal window, and type the following command, where *`ImportName`* `is the name of the import task with the failed entries that you want to correct.:`

   ```
   aws discovery describe-import-tasks - -name ImportName
   ```

2. From the output, copy the entire contents of the value returned for `errorsAndFailedEntriesZip`, without the surrounding quotes.
3. Open a web browser, and paste in the contents into the URL text box and press ENTER. This will download the failed records archive, compressed in a .zip format.

Now that you've downloaded your failed records archive, you can extract the two files within and correct the errors. Note that if your errors are tied to service-based limits, you'll either need to request a limit increase, or delete enough of the associated resources to get your account under the limit. The archive has the following files:

- **errors-file.csv** – This file is your error log, and it tracks the line, column name, `ExternalId`, and a descriptive error message for each failed record of each failed entry.
- **failed-entries-file.csv** – This file contains only the failed entries from your original import file.

To correct the non-limit-based errors you've encountered, use the `errors-file.csv` to correct the issues in the `failed-entries-file.csv` file, and then import that file. For instructions on importing files, see Importing data.

Document History for AWS Application Discovery Service

Latest User Guide documentation update: May 16, 2023

The following table describes important changes to the *Application Discovery Service User Guide* after January 18, 2019. For notifications about documentation updates, you can subscribe to the RSS feed.

Change	Description	Date
Maintenance mode	AWS Application Discovery Service is no longer open to new customers. Alternatively, use AWS Transform which provides similar capabilities. For more information, see AWS Application Discovery Service availability change.	November 7, 2025
Transition from Discovery Connector to Agentless Collector	We recommend that customers who are currently using Discovery Connector transition to the new Agentless Collector. Starting November 17, 2025, AWS Application Discovery Service will stop accepting new data from Discovery Connector s. For more information, see Discovery Connector.	November 12, 2024
Released the Agentless Collector Network Data Collection module	The Network Data Collectio n module makes it possible for you to discover dependenc ies among servers in your on-premises data center. For more information, see	November 8, 2024

	Using the Agentless Collector Network Data Collection module.	
Support for agentless collection for dependency mapping	For more information, see Using the VMware vCenter Agentless Collector data collection module.	October 24, 2024
Released Agentless Collector version 2 based on Amazon Linux 2023	For more information, see Prerequisites for Agentless Collector.	September 26, 2024
Updated Agentless Collector prerequisites	For more information, see Prerequisites for Agentless Collector.	September 9, 2024
Eventual consistency in the API	For more information, see Eventual consistency in the AWS Application Discovery Service API .	June 20, 2024
Agentless Collector updates	We added `sts.amazonaws.com` to the lists of domains that require outbound access. For more information, see Configure firewall for outbound access to AWS domains .	June 20, 2024
To separate access, create and use separate AWS accounts.	For more information, see Actions, resources, and condition keys for AWS Application Discovery Service .	April 5, 2024

Introducing the Agentless Collector database and analytics data collection module	The database and analytics data collection module is the new module of Application Discovery Service Agentless Collector (Agentless Collector). You can use this data collection module to connect to your environment and collect metadata and performance metrics from your on-premises database and analytics servers. For more information, see Database and analytics data collection module.	May 16, 2023
Introducing Application Discovery Service Agentless Collector	Application Discovery Service Agentless Collector (Agentless Collector) is the new AWS Application Discovery Service on-premises application that collects information through agentless methods about your on-premises environment to help you effectively plan your migration to the AWS Cloud. For more information, see Agentless Collector.	August 16, 2022

IAM update	The AWS Identity and Access Management (IAM) `discovery:GetNetworkConnectionGraph` action is now available for granting access to the AWS Migration Hub console network diagram when creating an identity-based policy. For more information, see Granting permissions to use the network diagram.	May 24, 2022
Introducing the home Region	The Migration Hub home Region provides a single repository of discovery and migration planning information for your entire portfolio, and a single view of migrations into multiple AWS Regions.	November 20, 2019
Introducing the Migration Hub import feature	Migration Hub import allows you to import information about your on-premises servers and applications into Migration Hub, including server specifications and utilization data. You can also use this data to track the status of application migrations. For more information, see Migration Hub Import.	January 18, 2019

The following table describes documentation releases for the *Application Discovery Service User Guide* before January 18, 2019:

Change	Description	Date
New Feature	Updated docs to support data exploration in Amazon Athena and added Troubleshooting chapter.	August 09, 2018
Major revision	Rewrites to usage & output details; entire document restructured.	May 25, 2018
Discovery Agent 2.0	A new and improved Applicati on Discovery agent was released.	October 19, 2017
Console	The AWS Management Console was added.	December 19, 2016
Agentless discovery	This release describes how to set up and configure agentless discovery.	July 28, 2016
New details for Microsoft Windows Server and command issue fixes	This update adds details about Microsoft Windows Server. It also documents fixes to various command issues.	May 20, 2016
Initial publication	This is the first release of the *Application Discovery Service User Guide.*	May 12, 2016

AWS Glossary

For the latest AWS terminology, see the AWS glossary in the *AWS Glossary Reference*.