

Cloudera Runtime 7.3.1

Migrating Navigator to Atlas for CDP Private Cloud

Date published: 2020-07-28

Date modified: 2024-12-10

CLouDERA

<https://docs.cloudera.com/>

Legal Notice

© Cloudera Inc. 2025. All rights reserved.

The documentation is and contains Cloudera proprietary information protected by copyright and other intellectual property rights. No license under copyright or any other intellectual property right is granted herein.

Unless otherwise noted, scripts and sample code are licensed under the Apache License, Version 2.0.

Copyright information for Cloudera software may be found within the documentation accompanying each component in a particular release.

Cloudera software includes software from various open source or other third party projects, and may be released under the Apache Software License 2.0 (“ASLv2”), the Affero General Public License version 3 (AGPLv3), or other license terms. Other software included may be released under the terms of alternative open source licenses. Please review the license and notice files accompanying the software for additional licensing information.

Please visit the Cloudera software product page for more information on Cloudera software. For more information on Cloudera support services, please visit either the Support or Sales page. Feel free to contact us directly to discuss your specific needs.

Cloudera reserves the right to change any products at any time, and without notice. Cloudera assumes no responsibility nor liability arising from the use of products, except as expressly agreed to in writing by Cloudera.

Cloudera, Cloudera Altus, HUE, Impala, Cloudera Impala, and other Cloudera marks are registered or unregistered trademarks in the United States and other countries. All other trademarks are the property of their respective owners.

Disclaimer: EXCEPT AS EXPRESSLY PROVIDED IN A WRITTEN AGREEMENT WITH CLOUDERA, CLOUDERA DOES NOT MAKE NOR GIVE ANY REPRESENTATION, WARRANTY, NOR COVENANT OF ANY KIND, WHETHER EXPRESS OR IMPLIED, IN CONNECTION WITH CLOUDERA TECHNOLOGY OR RELATED SUPPORT PROVIDED IN CONNECTION THEREWITH. CLOUDERA DOES NOT WARRANT THAT CLOUDERA PRODUCTS NOR SOFTWARE WILL OPERATE UNINTERRUPTED NOR THAT IT WILL BE FREE FROM DEFECTS NOR ERRORS, THAT IT WILL PROTECT YOUR DATA FROM LOSS, CORRUPTION NOR UNAVAILABILITY, NOR THAT IT WILL MEET ALL OF CUSTOMER’S BUSINESS REQUIREMENTS. WITHOUT LIMITING THE FOREGOING, AND TO THE MAXIMUM EXTENT PERMITTED BY APPLICABLE LAW, CLOUDERA EXPRESSLY DISCLAIMS ANY AND ALL IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO IMPLIED WARRANTIES OF MERCHANTABILITY, QUALITY, NON-INFRINGEMENT, TITLE, AND FITNESS FOR A PARTICULAR PURPOSE AND ANY REPRESENTATION, WARRANTY, OR COVENANT BASED ON COURSE OF DEALING OR USAGE IN TRADE.

Contents

Migrating Navigator content to Atlas.....	4
Migrated features and limitations.....	4
Migration assumptions and prerequisites.....	5
Estimating migration time and resources required.....	6
Understanding migrated Navigator entites.....	7
How the migration process works.....	8
Installing Atlas using the Cloudera Manager upgrade wizard.....	10
Migrating Navigator data using customized scripts.....	11
Run the extraction.....	12
Run the transformation.....	13
Run the import.....	13
Validate the migration.....	13
Move Atlas out of migration mode.....	14
Mapping Navigator metadata to Atlas.....	14
Migrating Navigator audit information into Atlas.....	15
Migrate audit reports and processes to Ranger.....	16
What is new in Atlas for Navigator Users.....	17

Migrating Navigator content to Atlas

While upgrading from CDH to Cloudera Base on Private Cloud, you can migrate the metadata from Navigator to Apache Atlas, for a scalable and robust infrastructure that supports enhanced metadata searchability by collecting metadata directly from your cluster.

Cloudera Runtime 7.0.0 includes Apache Atlas to collect technical metadata from cluster services. Atlas replaces Cloudera Navigator Data Management for these clusters. Cloudera has incorporated many features from Navigator into Apache Atlas to make sure that the rich metadata collected in Navigator can be represented in Atlas. Atlas provides scalable and robust infrastructure that supports metadata searches and lineage across enterprise production clusters.



Important: Governance through Apache Atlas is just one element of a secure production cluster: Cloudera supports Atlas when it runs on a cluster where Kerberos is enabled to authenticate users. When upgrading from CDH to Cloudera Base on Private Cloud 7.1.7 and running Apache Atlas, the new cluster must have Kerberos enabled.

You may choose not to migrate Navigator content to Atlas: this document describes how to think about archiving your Navigator audit information and metadata.

Whether you choose to migrate Navigator contents to Atlas or not, you must know how to use Atlas to accomplish the tasks you are accustomed to performing in Navigator.

Related Information

[Migrated features and limitations](#)

[Transitioning Navigator content to Atlas \(CDH 5 to CDP Private Cloud Base\)](#)

[Transitioning Navigator content to Atlas \(CDH 6 to CDP Private Cloud Base\)](#)

[Migration assumptions and prerequisites](#)

[About migrating from Cloudera Navigator to Atlas using Cloudera Manager 6](#)

[About migrating from Cloudera Navigator to Atlas using Cloudera Manager 7](#)

Migrated features and limitations

Some of the migrated features from Navigator to Apache Atlas are described in this page. Some of the features that are not migrated are also listed.

Business metadata is migrated into Atlas, including:

- Tags
- Custom properties (definitions and entity assignments)
- Managed metadata properties (definitions and entity assignments)
- Original and updated entity names and descriptions

Technical metadata from the following sources are migrated into Atlas:

- Hive
- Impala
- Spark
- Referenced HDFS / S3

Features that are not migrated

- Audits. In Cloudera, Ranger collects audit information for successful and failed access to objects under its control. This document includes information on how to migrate your audit information to Ranger and how to archive your existing Navigator audit information.

- Entity Metadata. The following metadata entities in Navigator are not migrated to Atlas:
 - Unreferenced S3 and HDFS entities. Files in HDFS and S3 that are not included in lineage from Hive, Spark, or Impala entities are not migrated.
 - Metadata for Sqoop, Pig, Map-Reduce v1 and v2, Oozie, and YARN.
- Saved Search. Available in both Navigator and Atlas but the implementation is different. The data pertaining to the saved searches are not migrated to Atlas by the nav2atlas script. The users know how to save a search in Atlas, which is similar to the Navigator saved search.
- Policies. Navigator policies are not migrated to Atlas.
- Configuration settings. Configuration properties determining Navigator behavior that you have set in Cloudera Manager are not migrated to the new environment. The properties that may apply in Atlas, such as authentication credentials, must be reset in the new environment.

Limitations of migrating from Navigator to Atlas

Regular expressions in Managed Properties

- Navigator has Managed Properties which are mapped to Business Metadata in Atlas but the regular expressions added for Managed Properties are not mapped to Business Metadata.

Post upgrade deleting Business Metadata associated to Entity is not possible in Atlas

- In Navigator, the Managed Property can be deleted when it is attached to an entity but post upgrade from CDH to CDP, it is not possible, as Atlas treats Business Metadata as a Typedef and deletion of type in Atlas requires no references to any type/entity for the type to be deleted.

Migration assumptions and prerequisites

Before you upgrade your CDH cluster to Cloudera Base on Private Cloud or migrate content from Navigator to Apache Atlas, ensure that you have collected all the credentials and set expectations for the time required for completing the migration process.

The prerequisites in this section help you to prepare in advance to migration.

In addition to the prerequisites outlined for the Cloudera Manager and Cloudera upgrades, you shall need the following for the Navigator to Atlas migration:

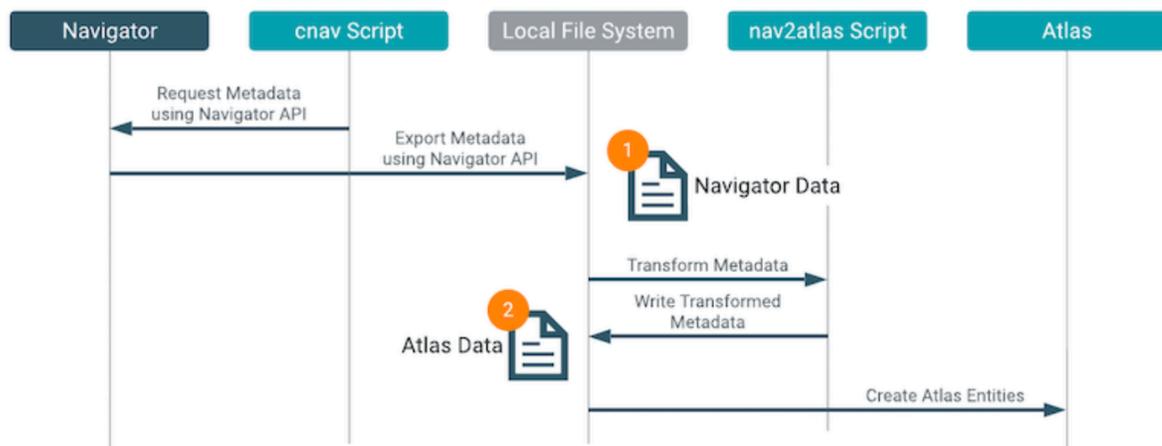
- Deleted entities in Navigator. Check the Navigator Administration page to make sure that a successful purge has run recently. If not, consider running a purge before the migration. See [Managing Metadata Storage with Purge](#).
- Role to host assignments. Before you begin upgrading to CDP, make a plan for where you will install the Atlas server. In addition, Atlas depends on the HBase, Kafka, and Solr services; your plan should include host assignments for installing the components of these services. See [Runtime Cluster Hosts and Role Assignments](#).
- Resources for Atlas service. Atlas requires 16 GB of Java heap (Atlas Max Heapsize property) and 4 Solr shards (Initial Solr Shards for Atlas Collections property). Make sure the host you choose for Atlas has enough resources for all the services' requirements.



Caution: You must note about the default values for Initial Solr Shards for Atlas Collections in your Cloudera Manager UI. Before you commence the Atlas initialization process, based on your performance requirements, you must decide the actual (correct) values for Initial Solr Shards for Atlas Collections in your Cloudera Manager instance. Cloudera recommends to set 4 Solr shards (Initial Solr Shards for Atlas Collections property). Also, you must not update or modify these values after the Atlas initialization has commenced. Additionally, note that once the Atlas initialization process is completed, modifying the value of Initial Solr Shards for Atlas Collections or Initial Solr Replication Factor for Collections will not have any effect on the collections for Atlas in Solr.

- Resources for Solr service. During migration, Solr running to serve as Atlas' index requires 12 GB of Java heap (Java Heap Size of Solr Server in Bytes property). You can reset this to make sure the host you chose for Atlas has enough resources for all the services' requirements.

- Navigator credentials. The migration requires the username and password for a Navigator user with administrator privileges.
- Local disk space needed for intermediate processing. The first two phases of the Navigator-to-Atlas migration produce intermediate files in the /tmp folder in the local file system where Atlas is installed. See [Estimating migration time and resources required](#) on page 6.
- Local disk space for migration staging files. The first two phases of the Navigator-to-Atlas migration produce staging files on the local disk where Atlas is installed. See [Estimating migration time and resources required](#) on page 6.



- Time estimates for migration phases. Each phase of the migration runs independently from the upgrade. You can trigger them to run when convenient. See [Estimating migration time and resources required](#) on page 6.

Estimating migration time and resources required

You must plan your migration process by checking on the time and resources that are required.

While the cluster is starting up, you can plan for and start the migration process.

1. Inspect Navigator installation to determine the number of Navigator entities that will be migrated. See [How many Navigator entities are migrated?](#)
2. Estimate the time and disk space required for each phase of the migration.

The following migration rates are approximate and depend on the resources available on the Atlas host and other unknown factors. Note that the number of entities actually imported may be considerably less than the number of entities extracted. The migration process discards HDFS entities that are not referenced by processes that are migrated such as Hive, Impala, and Spark).

Migration Phase	Migration Rate	Disk Space	Output File Size	Trial Data Points
Extraction	4 minutes / 1 million entities	100 MB / 1 million entities, less as volumes increase	65 MB / 1 million entities	10 million entities takes about 30 minutes; 256 million takes about 18 hours.
Transformation	1.5 minutes / 1 million entities	100 to 150 MB / 1 million entities, higher end of range with larger volumes	150 MB / 1 million entities	10 million entities takes about 20 minutes; 256 million takes about 6 hours.
Import	35 minutes / 1 million migrated entities	N/A	N/A	10 million entities takes about 4 hours; 256 million takes about 6 days.

Understanding migrated Navigator entities

When preparing to migrate content from Navigator to Atlas, it helps to know how many Navigator entities will be extracted. Use Navigator's search facets for additional details.

To determine the number of Navigator entities extracted for the extraction and transformation phases of the migration perform the following steps:

1. Log into Navigator.
2. In the Cluster Group facet in the left panel, select the cluster you are migrating from.

The main panel displays the count of entities in that cluster. Use this value for estimating the extraction and transformation phase durations.

The screenshot shows the Cloudera Navigator Search interface. The top navigation bar includes 'Search', 'Audits', 'Analytics', 'Policies', 'Administration', and 'admin'. The main search area has a search bar and an 'Actions' dropdown. On the left, there are filter facets: 'SOURCE TYPE', 'TYPE', 'OWNER', 'CLUSTER GROUP', and 'TAGS'. The 'CLUSTER GROUP' facet is expanded, showing 'Cluster 1' selected with a count of 2,024. An orange arrow points from this count to the '2,024 results' badge at the top of the main results area. The main results area shows a table of entities for 'Cluster 1', including 'spark', 'yarn', and 'mapreduce'.

Entity Name	Source Type	Last Modified	Created	Owner	Group	Permissions	Source	HDF...	View in Hue
spark	HDFS Directory	Apr 19, 20...	Apr 19, 20...	View in Hue
yarn	HDFS Directory	Apr 19, 20...	Apr 19, 20...	View in Hue
mapreduce	HDFS Directory	Apr 19, 20...	Apr 19, 20...	View in Hue

Not all Navigator entities are imported into Atlas. To estimate the subset of entities included in the import phase:

1. Log into Navigator.
2. In the Cluster Group facet in the left panel, select the cluster you are migrating from.
3. In the Source Type facet in the left panel, select "Hive", "Impala", and "Spark".

The main panel displays the count of entities from these sources in this cluster.

4. Double the number from the search results to account for the physical files that correspond to the tables and jobs. The HDFS entities referenced by the Hive, Impala, and Spark entities are included in the migration.

The migration brings over all business metadata definitions and associations with migrated entities. To determine the number of Navigator-managed properties to migrate perform the following steps:

1. Log into Navigator.
2. In the left Search panel, find the Tags facet.

This facet lists all the tags defined in Navigator. Navigator tags are imported into Atlas as labels.

3. Go to Administration > Managed Properties.

The Navigator namespaces are imported as Atlas business metadata collections. Each managed property is imported as a business metadata attribute.

How the migration process works

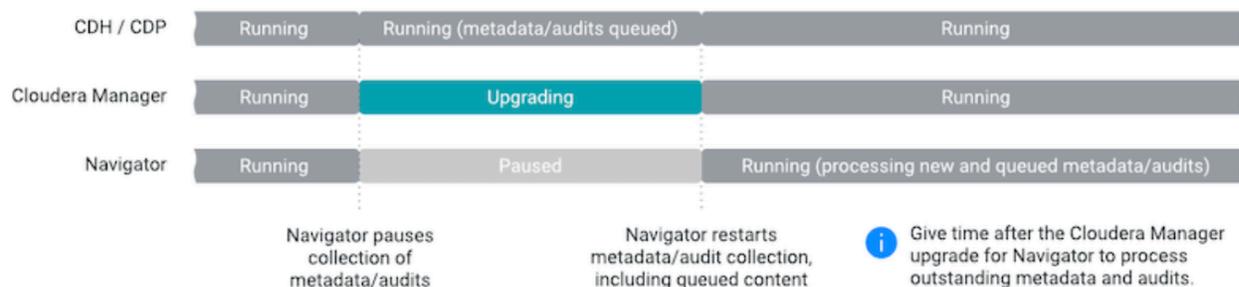
Before migrating from Navigator to Apache Atlas, review the migration paths. You must extract, transform, and import the content from Navigator to Apache Atlas. After the migration is completed, services start producing metadata for Atlas and audit information for Apache Ranger.

There are two main paths that describe a Navigator-to-Atlas migration scenario:

- Upgrading Cloudera Manager to CDP 7.0.0 and upgrading all your CDH clusters to Cloudera Runtime. In this case, you can stop Cloudera Navigator after migrating its content to Atlas.
- Upgrading Cloudera Manager to CDP 7.0.0 but managing some or all of your existing CDH clusters as CDH 5.x or 6.x. In this case, the CDP cluster running Cloudera Navigator continues to extract metadata and audit information from existing CDH clusters and runs Atlas and Ranger to support metadata and audit extraction from new or potential new Cloudera Runtime clusters.

In both the scenarios, you shall complete the upgrade of Cloudera Manager first. While Cloudera Manager is being upgraded, Navigator pauses collection of metadata and audit information from cluster activities. After the upgrade is complete, Navigator processes the queued metadata and audit information.

In the timeline diagrams that follow, the blue color indicates steps and because you trigger the steps manually, you can control their timing.



The migration of Navigator content to Atlas occurs during the upgrade from CDH to CDP. The migration involves three phases:

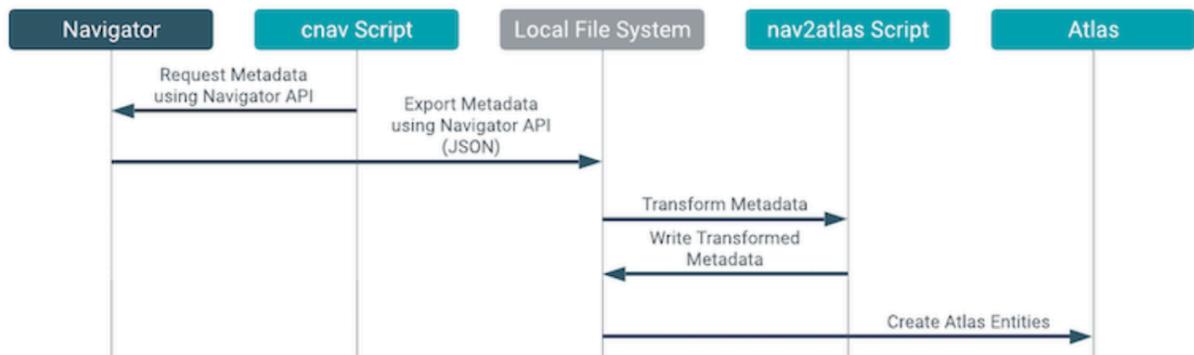
- Extracting metadata from Navigator

The Atlas installation includes a script (`cnav.sh`) that calls Navigator APIs to extract all technical and business metadata from Navigator. The process takes about 4 minutes per one million Navigator entities. The script compresses the result and writes it to the local file system on the host where the Atlas server is installed. Plan for about 100 MB for every one million Navigator entities; lower requirements for larger numbers of entities.

- Transforming the Navigator metadata into a form that Atlas can consume. Including time and resources. The Atlas installation includes a script (`nav2atlas.sh`) that converts the extracted content and again compresses it and writes it to the local file system. This process takes about 1.5 minutes per million Navigator entities. The script compresses the results and writes it to the local file system on the host where the Atlas server is installed. Plan for about 100 to 150 MB for every million Navigator entities; higher end of the range for larger numbers of entities.
- Importing the transformed metadata into Atlas.

After the Cloudera upgrade completes, Atlas starts in "migration mode," where it waits to find the transformed data file and does not collect metadata from cluster services. When the transformation is complete, Atlas begins

importing the content, creating equivalent Atlas entities for each Navigator entity. This process takes about 35 minutes for a million Navigator entities, counting only the entities that are migrated into Atlas.

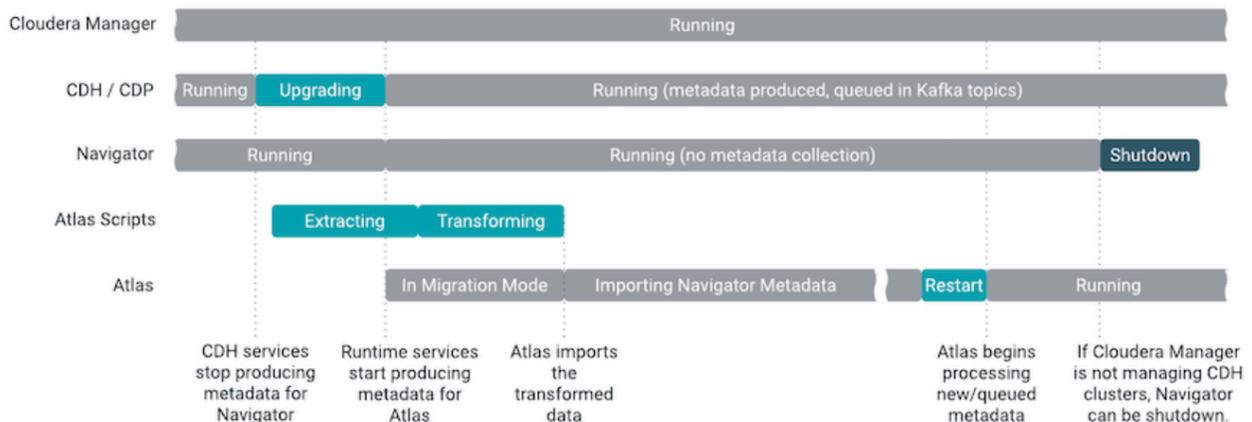


To make sure you do not miss metadata for cluster operations, provide time after the Cloudera Manager upgrade and before the CDH upgrade for Navigator, to process all the metadata produced by CDH service operations. See [Navigator Extraction Timing](#) for more information.

You can start extracting metadata from Navigator as soon as the CDP parcel is deployed on the cluster. After CDP is started, Navigator no longer collects metadata or audit information from the services on that cluster; instead services produce metadata for Atlas and audit information for Ranger.



Important: After the CDH upgrade process is completed, Atlas starts in migration mode and does not process metadata. When the migration process completes, you must manually update the Atlas configuration in Cloudera Manager for Atlas to begin processing metadata.



Migration from Navigator to Atlas can be run only in non-HA mode

Migration import works only with a single Atlas instance.

If Atlas has been set up in HA mode before migration, you must remove the additional instances of Atlas, so that Atlas service has only one instance.

Later, start Atlas in the migration mode and complete the migration. Perform the necessary checks to verify if the data has been imported correctly.

Restart Atlas in non-migration mode.

- If you have Atlas setup in HA mode, retain only one instance and remove the others.
- Ensure that the ZIP files generated as an output from the Nav2Atlas conversion are placed at the same location where the Atlas node is present.

Installing Atlas using the Cloudera Manager upgrade wizard

You can use the wizard in Cloudera Manager to upgrade from CDH to Cloudera Base on Private Cloud, install Apache Atlas and its dependencies, and to generate customized commands to initiate the migration phases.

This page describes the Cloudera Manager wizard steps that help you to set up Atlas and its service dependencies.



Note: Before you perform the upgrade process, note the following pertaining to Atlas operations.

When a cluster is upgraded from CDH 5 or CDH 6 to CDP 7.x, by default the `atlas.metadata.namespace` property is set to `cm`.

If a different namespace property needs to be set, for example: `cluster1`, the same needs to be set while running the `nav2atlas` script as `-clusterName` parameter and also in the `atlas-application` properties in the upgrade wizard.

Post-upgrade, note that the different value `"cluster1"` is not automatically updated in Hooks for services like Hive, HBase, Impala, Spark, and Kafka. You must make sure that before you upgrade your cluster and after the services are installed, you must set the value `"cluster1"` for all the available services. And later complete the upgrade process.

As an example process, follow these steps if namespace other than default `"cm"` needs to be set:

1. Provide the namespace (`"cluster1"`) in the transformation step of the `Nav2Atlas` script.
2. Add `atlas.metadata.namespace =cluster1` in `atlas-application` properties in the upgrade wizard for Atlas.
3. Open another tab of the Cloudera Manager while the upgrade process is in progress and add `atlas.metadata.namespace =cluster1` in advanced configuration Safety Valve of `atlas-application.properties` for all the Hook services (Hive, HiveServer2, Spark, HBase, Impala, and Kafka).
4. Perform all the other steps in the upgrade wizard and complete the upgrade.
5. Remove Atlas from the migration mode.

To return to the main wizard documentation, go to [Upgrading a CDH 6 Cluster](#).

Follow instructions in the upgrade wizard "Install Services" section.

Atlas installation

Refer to [Step 10: Run the Upgrade Cluster Wizard](#).

1. Enable Atlas install. If the CDH cluster being upgraded was running Navigator, the upgrade wizard shows a note recommending that you enable Atlas in the new cluster. Select the Install Atlas option.



2. Install Atlas dependencies.

The wizard steps through the installation for Atlas' dependencies, assuming these services have not already been included in the installation:

- ZooKeeper. Assign one or more hosts for the ZooKeeper role.
- HDFS. Already included in the installation.
- Kafka. Select the optional dependency of HDFS. Atlas requires configuring the Broker service only, not MirrorMaker, Connect, or Gateway.
- HBase. Atlas requires configuring HBase Master and RegionServers only, not REST or Thrift Server. Assign a Master role on at least one host. Assign RegionServers to all hosts.
- Solr. Assign a host for the Solr Server role. Set the Java Heap Size of Solr Server in Bytes property to 12 GB (to support the migration operation).

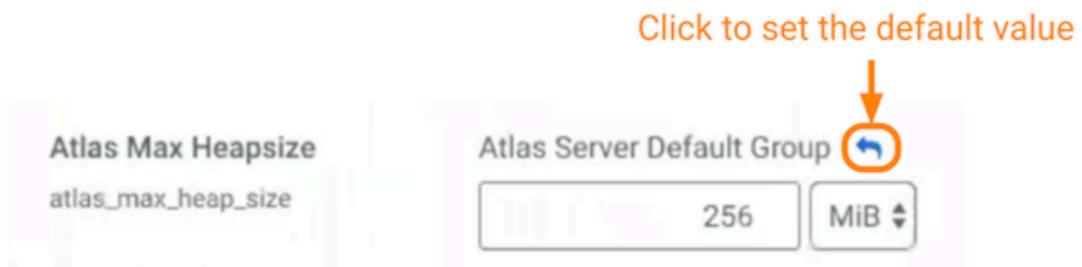
For recommendations on where in the cluster to install the service roles, see [Runtime Cluster Hosts and Role Assignments](#).

3. Click Add Atlas Service. The wizard steps through choosing a host and setting migration details.

- Set the host for the Atlas server roles and click Continue.

Tip: Remember this host because you need to SSH to it later to trigger the content migration from Navigator.

- The Atlas Migrate Navigator Data screen displays. This screen contains migration commands that are customized to your environment. When you fill in the output file paths, the command text changes to incorporate your settings.
- The Atlas Enable Migration Mode screen displays. Review the Atlas advanced configuration Safety Valve content and click Continue. After the migration is complete, you remove these settings manually to start Atlas.
- The Atlas Review Changes screen displays. Review the configurations and make any necessary changes. You must provide a value for the following:
 - Admin Password – choose a password for the preconfigured admin user.
 - Atlas Max Heapsize – set the max heapsize to the default value by clicking the curved blue arrow. If you plan to migrate content from Cloudera Navigator to Atlas, consider setting the heapsize to 16 GB.



4. Click Continue.



Note: To complete the Navigator-to-Atlas migration outside of the Cloudera Runtime upgrade, see [Transitioning Navigator data using customized scripts](#).

The Cloudera Manager upgrade wizard continues with "Other Tasks" and "Inspector Checks" sections. The wizard steps for installing Atlas are complete at this point and you can continue to complete the Cloudera Runtime upgrade.

Migrating Navigator data using customized scripts

You can run the customized scripts generated by the Cloudera Manager wizard to configure the Apache Atlas installation and start the Navigator to Atlas data migration process when you step into the Cloudera upgrade wizard. You can also run the migration scripts independently from the Cloudera upgrade.

The migration has three phases: extraction, transformation, and import. If you haven't already, estimate the time and resource requirements for the migration steps as described in [Assumptions and prerequisites](#).

Run the extraction

You can run the extraction in the background as soon as the CDP runtime parcel is deployed. To customize and run the extraction command.

Procedure

1. Go back to the editor where you saved the extraction commands, from [Copy the extraction command text](#) from the step "Click Add Atlas Service."
2. Open a terminal window or command prompt where you have access to the cluster.
3. Using the provided command, SSH into the Atlas host.
4. Make sure the JAVA_HOME variable is set; if it is not set, run the export command pointing to the location of the JDK.
5. Customize the extraction command to include the Navigator admin user and password.
6. Run the extraction command. When the extraction is complete, a status message is displayed in the command output.

If Navigator is configured with TLS/SSL enabled, the cnav script needs the following credential information:

- Truststore path
- Truststore password
- Keystore path
- Keystore password

To make these parameters available, run the following commands before running the cnav script:

```
export KEYSTORE_PATH=<keystore-path>;
export KEYSTORE_PASSWORD=<keystore-password>;
export TRUSTSTORE_PATH=<truststore-path>;
export TRUSTSTORE_PASSWORD=<truststore-password>
```

For example, the command sequence might look similar to the following (line breaks are introduced for readability):

```
export KEYSTORE_PATH=/opt/cloudera/CMCA/trust-store/acme_truststore.jks;
exportKEYSTORE_PASSWORD=Q01FAeH53dn1HLY74D68KklyMAQVGtOI_cLznArccid48DD
zS0VXY-DWnzpz0Ug10BvikGmoovYaZT2EEEdBGgLPiDCRKHyzFExE30ITRGazjKPtZxAaXOU
zgKMMmQQgJKw-5JW9I6WgLGbHcPkJfBa7vP3z6PFtm6XfyB-o3R6qmcZLWslDIQ18mowuFV
louQIZa;export
TRUSTSTORE_PATH=/var/lib/cloudera-scm-agent/agent-cert/cm-auto-
host_keystore.jks; export TRUSTSTORE_PASSWORD=123420978alngdfdfjliaiu; /
opt/cloudera/cm-agent/service/navigator/cnav.sh -n
https://acme-finance-1.acme-finance:7187 -u admin -p adminpass -
c "Cluster 2" -o
/tmp/nav2atlas/nav2altas_nav_export.zip
```



Note: The -c flag parameter in the nav2atlas script represents the Navigation Metadata Server (NMS) cluster name and NOT the Cloudera Manager cluster name. You must retrieve the value of -c from the Navigator UI to use it while running the nav2atlas script."

Run the transformation

You can run the transformation in the background as soon as the extraction completes. To customize and run the transformation command:

Procedure

1. Go back to the editor where you saved the transformation commands, from [Copy the transformation command text](#) from the step Click Add Atlas Service.
2. (Optional) Open a terminal window or command prompt where you have access to the cluster.
3. If needed, SSH into the Atlas host.
4. If needed, make sure the JAVA_HOME variable is set; if it is not set, run the export command pointing to the location of the JDK.
5. Run the transformation command.
When the transformation is complete, you see a status message in the command output.

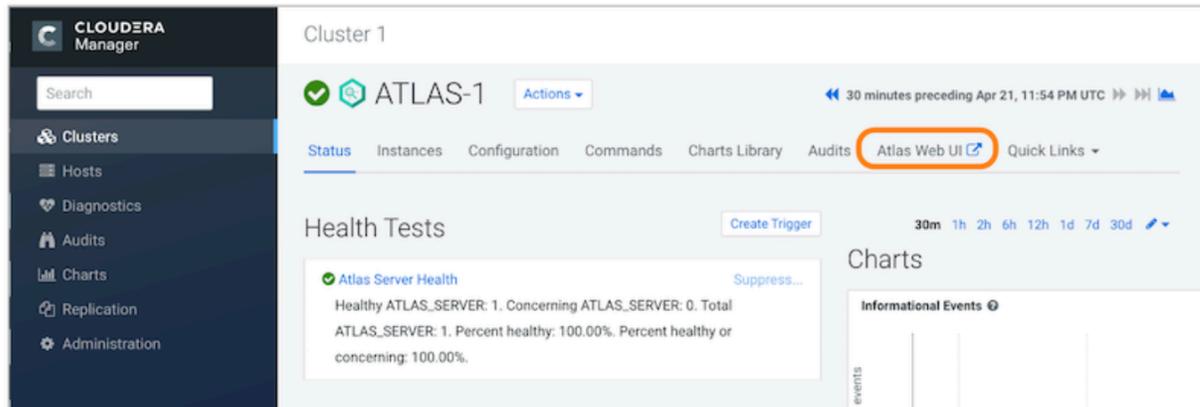
Run the import

When Atlas starts up, it checks the output file location for a completed file from the transformation phase. When Atlas finds the completed file, it begins importing entities automatically.

Procedure

1. Open the Atlas UI.

You can open Atlas web UI from the Atlas service page in Cloudera Manager.



2. Review the migration process in the Statistics page.

The normal Atlas UI does not appear in migration mode; instead you see the Statistics page, which shows a real-time report of the number of entities, classifications, and other metadata that have been created in Atlas.

Validate the migration

To verify that the migration is successful, use the Statistics page in Atlas to compare the metadata in Navigator.

In general, you must verify the following:

- Count of migrated entities. Does the total number of imported entities match what you expect from Navigator? Note that not all Navigator entities are not migrated: HDFS entities are only migrated if they are referenced in Hive, Impala, or Spark operations that are included in the migration cycle.
- Count of managed metadata that became business metadata in Atlas.

- Count of managed metadata assignments. Consider reproducing searches on commonly used business metadata to validate that you see the same results in each system.

Move Atlas out of migration mode

Once the installation is completed, you must move Atlas out of migration mode.

About this task

- Atlas does not collect metadata from services running on the cluster. The metadata remains in Kafka topics and collected later.
- Atlas imports metadata when it finds a final transformation file in the location you specified in [Confirm the output file location](#) from the step "Click Add Atlas Service."

To move Atlas from migration mode into normal operation:

Procedure

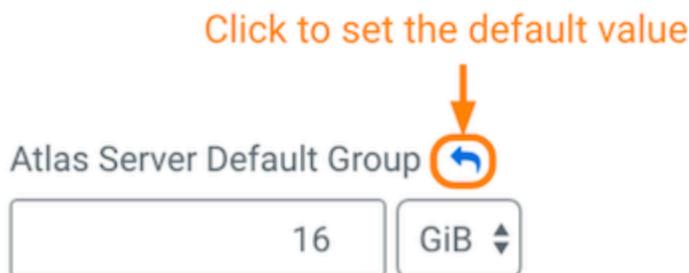
- Open Cloudera Manager to the Atlas service.
- Go to the Configuration tab.
- Filter the list of properties by typing "Safety" in the filter field.
- Remove the migration-specific entries from the Atlas Server Advanced Configuration Snippet (Safety Valve) for conf/atlas-application.properties

Remove the following properties:

```
Atlas.migration.data.filename
Atlas.migration.mode.batch.size
atlas.migration.mode.workers
```

- Reset the Atlas Max Heapsize property back to the default value.

Atlas Max Heapsize
atlas_max_heap_size



- Click Save Changes.
- Restart Atlas. Choose Action.
- Restart.

Mapping Navigator metadata to Atlas

You must validate the Navigator to Apache Atlas migration by reviewing the list of metadata mapping and its types that are migrated from the Navigator.

Use this topic as a checklist to help you validate the migration of metadata from Navigator to Atlas.

User-supplied metadata mapping

The following user-supplied metadata is migrated to Atlas, including definitions and assignments to entities. Enumerations defined in Navigator are created as instances of the enumeration type in Atlas.

Navigator Construct	Atlas Construct
Tag	Label
User-defined properties	User-defined properties
Key	Key
Value	Value
Managed Properties	Business Metadata Attributes
Namespace	Business Metadata name
Namespace description	Business Metadata description
Property Name (original name)	Attribute name
Classes (entity types that property can be assigned to)	Applicable Types (entity types that property can be assigned to)
Display Name	Display Name (not used in the UI)
Property description	Description (not used in the UI)
Multivalued	Enable Multivalues
Type (Text, Number, Boolean, Date, Enumeration)	Type (string, Boolean, byte, short, int, float, double, long, date, enumeration)

Policy mapping

Navigator policies provided the ability to apply metadata to Navigator entities. They do not have an equivalent function in Atlas and are not migrated. You may be able to implement similar functionality through data profilers in the Data Catalog. You can create a report of Navigator policies using the Navigator API, for example, use a GET call to the /policy API:

```
curl -X GET "<navigator host>:<port>/api/v14/policy" -u <username>:<password>
```

Technical metadata mapping

The migration process maps all the needed technical metadata for entities from Navigator to be represented in Atlas. There are some instances where the migration has to convert data from one type to another or generate new content to conform to the Atlas data model. The details of the entity mapping are exhaustively described in the [Atlas technical metadata transition reference](#).

Migrating Navigator audit information into Atlas

Existing Cloudera Navigator audit information is not migrated into the Cloudera cluster.

To migrate reports running against Navigator data to Apache Ranger and other resources you must review the available options.

To manage Navigator audit information in a Cloudera Runtime cluster, consider the following options:

Maintain legacy audit information in Navigator

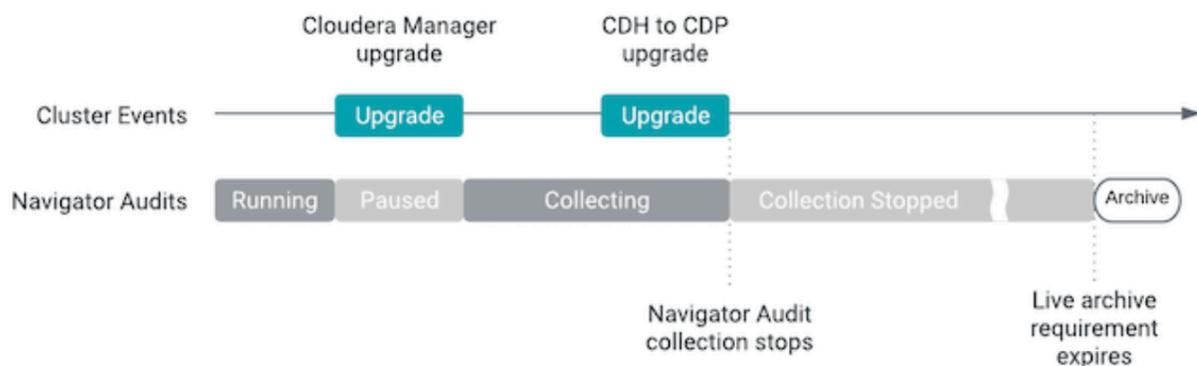
You can continue to run Navigator to access your existing audit information (and/or metadata). If you choose to keep Navigator running, make sure that its users do not add content to the archived system rather than the new Atlas instance. Consider:

- Removing editing privileges from users. If Navigator is configured for LDAP or Active Directory authentication, you can modify users or groups to remove privileges for adding or editing metadata. For details, see [Administering Navigator User Roles](#).
- Marking Navigator entities as stale. If you are managing more than one cluster in Navigator, you can mark entities from the upgraded cluster to indicate to users that the entities are no longer maintained in Navigator. One way to do this is to create a policy that assigns a tag to the entities from the upgraded cluster. For details, see [Using Policies to Automate Metadata Tagging](#).

Archive your Navigator audit information

When Cloudera Manager upgrades to 7.x, it maintains the database of Navigator audit information. After the upgrade, you can access audit information through Navigator; new audit information continues to be collected from CDH services.

When you upgrade a CDH cluster to Cloudera Runtime, the Navigator audit information persists. However, services no longer produce audit information for Navigator. You can continue to run Navigator to be able to access the audit information; at some point—perhaps after your need for immediate access to the audit information expires—you can archive the audit information.



At that point, if Cloudera Manager is not managing another CDH cluster, you can shut down Navigator.

Migrate audit reports and processes to Ranger

In Cloudera, Apache Ranger performs auditing against the data access policies defined for each service.

For example, if a Ranger policy allows only users from the Finance group to access a particular Hive database, Ranger audit information show when those users accessed the database successfully and when other users attempted to access the database and were denied. While the Ranger audit information are a significant subset of the audit information performed by Navigator, the format and content is different enough that Cloudera does not provide a migration path for Navigator audit information into the same repository as Ranger audit information.

When redirecting reports or processes to Ranger, you need to:

- Identify the audited information: does an equivalent exist in Ranger?
- Identify the method of accessing the audit data and map it to Ranger: Did the reporting application use the Navigator API? Did it access archived data or the Navigator audit database? Ranger provides an API to access audit information; audit data is written to HDFS (under `/ranger/audit/<component name>`). 30 days of audit records are indexed in Solr. The audit events are stored in a JSON format. For details, see [Managing Auditing with Ranger](#).

What is new in Atlas for Navigator Users

Reviewing the differences between Navigator and Apache Atlas helps you to know more about the features and functionalities available for metadata management in Apache Atlas.

Customers migrating from CDH with Navigator will be able to take advantage of these Atlas features:

- Control data access in Ranger using Atlas metadata. Atlas classifications can control table, row, or column level access; Ranger can use Atlas classifications as the trigger to instruct services to mask data. See [Configure Atlas Authorization using Ranger](#).
- Reliability and scalability. Atlas uses Kafka to pass metadata and HBase and Solr to store it. See [Apache Atlas architecture](#)
- Additional metadata sources. Atlas is flexible when it comes to adding metadata models for additional sources through REST APIs for metadata capture; it supports NiFi and Kafka metadata in addition to the sources integrated with Navigator. See [Extending Atlas to manage metadata from additional sources](#).

Running Navigator in Cloudera Manager

After upgrading Cloudera Manager, Navigator continues to collect metadata and audit information from CDH cluster services. There are no changes to Navigator functionality; all Navigator data is retained in the Cloudera Manager.



Note: Navigator is recommended to be only in Read-only mode

Post upgrading of the CDH cluster to CDP, Navigator is recommended to be used only in the Read-only mode. No scheduling, updating of properties, and purging is recommended. If any scheduling is created, it must be disabled immediately.

After upgrading a CDH cluster, services that previously sent metadata and audit information to Navigator, such as Hive, Impala, Spark, and HBase, are configured to send metadata to Atlas. Navigator audit collection for those services is disabled. You can still access audit information and metadata through Navigator; however, Navigator will not collect new information from cluster services. When you have exhausted the value of the Navigator audit information and after you have converted Navigator metadata to Atlas content, you can disable Navigator servers.