

Upgrade 1

Upgrading Data Analytics Studio

Date of Publish: 2019-10-04



<https://docs.hortonworks.com>

Contents

Upgrading DAS/DAS-Lite.....	3
Upgrading DAS-Lite to DAS.....	3
Upgrading from a previous DAS release.....	4
Upgrade roadmap.....	4
Pre-upgrade tasks.....	6
Upgrading the DAS cluster components.....	7

Upgrading DAS/DAS-Lite

You can upgrade from DAS-Lite to DAS, or to the DAS 1.4.1 release from the DAS 1.2.0, DAS 1.2.1, or DAS 1.3.0 releases.

Upgrading DAS-Lite to DAS

If you are running the DAS-Lite version, you can upgrade to DAS to get the full feature set of DAS. With DAS, you can get comprehensive usage metrics about your tables and queries by generating reports.

Before you begin

- Download the required repository tarballs from the Hortonworks customer portal, following the instructions provided as part of the product procurement process.
- Make sure that you obtain the DAS Management pack (MPack) of the same release and version numbers as that of DAS-Lite that you are running.

Procedure

1. Log in to the Ambari UI as an admin user.
2. Stop the DAS service.
3. On the Ambari host:
 - a) Download the DAS MPack corresponding to the DAS-Lite release version on your Ambari gateway host.
 - b) Uninstall the DAS-Lite MPack by running the following command:

```
ambari-server uninstall-mpack --mpack=data-analytics-studio-lite-mpack --verbose
```

- c) Install the DAS MPack that you downloaded earlier in this process.

Sample command:

```
ambari-server install-mpack --mpack=<mpack-filename>.tar.gz --verbose
```

Replace the .tar.gz filename with the actual tarball that you have downloaded.

- d) Restart the Ambari server by entering the following command:

```
ambari-server restart
```

4. Carry out the following steps on each of your DAS host, which includes DAS Webapp and the DAS Event Processor hosts:
 - a) SSH in to your DAS host.
 - b) Set up the DAS repository by copying the DAS repository configuration file to the following location:
(On CentOS 7) /etc/yum.repos.d/
(On Debian and Ubuntu) /etc/apt/sources.list.d
 - c) (Only on CentOS): Search and remove the data_analytics_studio packages of the DAS-Lite version by running the following commands:

```
rpm -qa | grep data_analytics_studio
```

```
yum remove <data_analytics_studio_lite>  
<data_analytics_studio_lite_<version>>
```

- d) Install the DAS packages.

(On CentOS 7)

```
yum install hdp3_data_analytics_studio
```

(On Debian and Ubuntu)

```
apt-get install hdp3_data-analytics-studio
```

Upgrading from a previous DAS release

To upgrade to the latest DAS/DAS-Lite release, you must upgrade the DAS cluster components (the DAS Webapp and the DAS Event Processor) on all the hosts on which you have installed DAS/DAS-Lite.

The supported starting points to upgrade to the DAS/DAS-Lite 1.4.1 release are DAS 1.2.0, 1.2.1, and 1.3.0 releases.

You can follow the instructions in this guide to upgrade from either the DAS 1.2.0, DAS 1.2.1, or the DAS 1.3.0 release.

Upgrade roadmap

Review the flowchart and the roadmap for an overview of the upgrade process.

Figure 1: DAS/DAS-Lite upgrade flowchart

1

Pre-upgrade Tasks

Back up the existing DAS configuration

Obtain the latest DAS binaries

Migrate the managed Postgres database

2

Upgrading the DAS cluster components

Remove the existing DAS cluster components

Remove the existing DAS rpm/deb

Uninstall the existing DAS MPack

Install the latest DAS MPack

Configure custom Postgres using the existing database

Configure the new repo URL

Add the DAS service

Table 1: DAS/DAS-Lite upgrade roadmap

No.	Task	Description
Pre-upgrade tasks		
1	Back up the existing DAS configuration.	The upgrade procedure involves deleting the DAS service from the Ambari UI. Therefore, back up all the configuration information before you proceed. For the list of items to back up, see Backing up the configuration .
2	Obtain the latest DAS binaries.	Obtain the latest DAS binaries and create and set up a local repository. See Obtaining the latest DAS binaries .
3	Migrate the managed Postgres database.	If you have configured a managed PostgreSQL database for production use, you can migrate the existing setup either by retaining the existing database and the data, or by creating a new, custom Postgres server, and configuring DAS to use it. See Migrating the managed Postgres database .
Upgrading DAS cluster components		
4	Remove the DAS cluster components.	Delete the DAS cluster components from the Ambari UI. This removes the DAS Webapp and the DAS Event Processor. See Removing the DAS cluster components .
5	Remove the existing DAS rpm/deb from the DAS host.	See Removing DAS rpm/deb .
6	Delete the /usr/das/current directory.	See Deleting the /usr/das/current directory .
7	Uninstall the existing DAS Mpack.	Remove the existing DAS Mpack from the Ambari console. See Uninstalling the existing DAS Mpack .
8	Install the latest DAS Mpack on the Ambari server.	See Installing the latest DAS Mpack .
9	Configure custom Postgres database using the existing, managed database.	See Configuring custom Postgres using an existing database and data .
10	Configure the new Repo URL.	See Configuring the new Repo URL .
11	Add the DAS service.	Use the same configurations while adding the new DAS service through the Ambari UI. See Adding the DAS service .

Pre-upgrade tasks

Before you start the upgrade process, backup your DAS configuration from the Ambari UI and keep other configuration settings and passwords handy.

Backing up the configuration

You must copy the DAS configurations from the Ambari UI manually, or you can use a request URL in a browser to download the configurations.

Before you use the request URL, sign in to the Ambari UI from the same browser.

Format of the request URL:

```
http(s)://{ambari-server}:{port}/api/v1/clusters/
{cluster_name}/configurations/service_config_versions?
service_name=DATA_ANALYTICS_STUDIO&is_current=true
```

To obtain the cluster_name, on the Ambari UI, click **admin > Manage Ambari > Cluster Information** and note the value from the **Cluster Name** field.

Download the configurations in a file. You will need this configuration information when you are adding back the DAS service for the latest release.

The upgrade procedure involves deleting the DAS service from the Ambari UI. This deletes all the configuration that you have done in Ambari. Therefore, make a note of the following information before proceeding to the upgrade procedure:

- Database password (data_analytics_studio_database_password)
- Hive session parameters (das_hive_session_params)
- event_processor_keystore_password
- webapp_keystore_password
- Hosts on which you have installed the existing DAS Event processor and the DAS Webapp

Obtaining the latest DAS binaries

Download the DAS binaries from the Cloudera Customer Portal by following the instructions provided as part of the subscription fulfillment process.

After you download the binaries, create and set up a local repository by following the procedure described in [Create and Set up a Local Repository](#).

Migrating the managed Postgres database

We recommend that you use a custom database in your production environment. If you have configured a managed PostgreSQL database for production use, you can migrate the existing setup either by retaining the existing database and the data, or by creating a new, custom Postgres server, and configuring DAS to use it.

Creating a new Postgres server

We recommend that you create a new Postgres server and configure DAS to use it. But doing so will bootstrap DAS again, and it can cause you to lose data about the queries that are older than the configured cleanup interval for the Hive and Tez events. To retain the older data, you can migrate data from the existing Postgres instance.

Configuring Postgres while keeping existing database and data

You can make the existing, managed Postgres database a custom database while retaining the data. This is a quick and dirty method of using an existing Postgres database. However, note that after migrating the database, the responsibility of maintaining the database need to be taken by your organization.

You need to complete this step during the upgrade process, after you have installed the latest MPack.

Upgrading the DAS cluster components

At this stage, you remove the existing DAS cluster components, rpm/deb, and the existing DAS Mpack, and replace them with the latest DAS binaries.

Removing the DAS cluster components

Procedure

1. Stop the DAS service.
On the Ambari UI, go to **Data Analytics Studio > Summary**, and click **Actions > Stop**.
2. Delete the DAS service.
Go to **Data Analytics Studio > Summary** and click **Actions > Delete Service**.

Removing DAS rpm/deb

Procedure

1. SSH in to the DAS host.
2. Obtain the name of the DAS rpm/deb files by running the following command:
On CentOS:

```
rpm -qa | grep data_analytics_studio
```

Sample output:

```
hdp3_data_analytics_studio_1_2_1_0_31-1.2.1.1.2.1.0-31.noarch
hdp3_data_analytics_studio-1.2.1.1.2.1.0-31.noarch
```

On Debian/Ubuntu:

```
dpkg-query -l | grep data-analytics-studio
```

3. Remove all the DAS rpm/deb files by using the following command:

On CentOS:

```
yum remove <DAS_RPM_name>
```

Example:

```
yum remove hdp3_data_analytics_studio_1_2_1_0_31-1.2.1.1.2.1.0-31.noarch
hdp3_data_analytics_studio-1.2.1.1.2.1.0-31.noarch
```

On Debian/Ubuntu:

```
apt-get remove <das_dep_name>
or
apt-get purge <das_dep_name>
```

Deleting the /usr/das/current directory

Procedure

1. SSH in to the Ambari console.
2. Delete the /usr/das/current/ directory by using the following command:

```
rm -rf /usr/das/current/
```

Uninstalling the existing DAS MPack

Procedure

1. SSH in to the Ambari host.
2. Uninstall the DAS MPack by running the following command:

```
ambari-server uninstall-mpack --mpack-name=data-analytics-studio-mpack
```

To uninstall a DAS-Lite MPack, run the following command:

```
ambari-server uninstall-mpack --mpack-name=data-analytics-studio-lite-mpack
```

Installing the latest DAS MPack

About this task

You must install the latest DAS MPack on the same host on which you had the existing DAS cluster running before the upgrade.

Procedure

1. Install the latest DAS MPack on the Ambari host by running the following command:

```
ambari-server install-mpack --mpack=<mpack-name> --verbose
```

Replace the <mpack-name> with the name of the DAS 1.4.1 MPack. Typically, the format of the MPack is hdp-data-analytics-studio-mpack-X.X.X.tar.gz.

2. Restart the Ambari server by entering the following command:

```
ambari-server restart
```

Configuring custom Postgres using an existing database and data

About this task

To make the existing, managed Postgres database a custom database while retaining the data:

Procedure

1. From the Ambari UI, go to **Data Analytics Studio > CONFIGS > Advanced data_analytics_studio-database**, and uncheck the **Create Data Analytics Studio database** option.
2. In the **Data Analytics Studio database hostname** field, specify the host where you had installed the DAS Webapp (before the 1.4.1 release).
3. Start the Postgres service manually from the same machine where the DAS Webapp was previously installed.

Configuring the new Repo URL

Procedure

1. On the Ambari UI, go to **Admin > Stacks and Versions > Versions**.
2. Click **Show Details**.
3. On the **Version Details** pop-up, click the **Edit Repositories** icon.
4. On the **Repositories** window, specify the new Repo URL in the **dasbn-repo** field.
5. Click **Save**.

Adding the DAS service

About this task

You must specify the same configuration values that you extracted earlier in the process while you are adding the DAS service through the Ambari UI.

Procedure

1. On the Ambari UI, click **Actions > Add Service**.
The **Add Service Wizard** is displayed.
2. On the **Choose Services** page of the Wizard, select **Data Analytics Studio**.
Other required services are automatically selected.
3. When prompted to confirm the addition of dependent services, give a positive confirmation to all.
This adds the other required services.
4. On the **Assign Masters** page, assign the hosts on which you had installed the existing DAS Webapp and the Event Processor.
5. On the **Customize Services** page, expand **Advance_data_analytics_studio-database** and fill in the database details and other required fields that are highlighted.

**Note:**

- You must specify the same database name, username, and password in Ambari while configuring DAS after the upgrade.
- Enter a password in the password field. Do not leave this field empty.
- If you are using a managed Postgres database, then leave the **Create Data Analytics Studio database** checkbox in checked state, and do not specify the database hostname in the **Data Analytics Studio database hostname** field.
- If you are using a custom Postgres database, then uncheck the **Create Data Analytics Studio database** checkbox, and specify the same database hostname in the **Data Analytics Studio database hostname** field on which your database is running.

6. In the **user_authentication** field, specify the authentication method that you want to use. You can specify one of the following:

- **NONE**: Specify NONE if you do not want to enable user authentication. In this case, all the queries will be executed as the hive service user.
- **KNOX_SSO**: Specify KNOX_SSO to enable single sign-on experience to authenticate users as configured in the Knox SSO topology.
- **KNOX_PROXY**: Specify KNOX_PROXY to use DAS behind a Knox proxy. The user is authenticated as configured in the Knox proxy topology. It also enables SPNEGO authentication.
- **SPNEGO**: Specify SPNEGO to authenticate users using a token-based authentication.



Note: The **user_authentication** field replaces the **knox_sso_enabled** checkbox from DAS 1.3.0 onward.

You can configure the user authentication separately, after you have added the DAS service. See [Configuring user authentication](#).

7. In rest of the fields, enter the same configuration values as they were in the existing DAS configuration. You can find the configuration details in the file that you saved while [Backing up the configuration](#).
8. Complete the remaining installation wizard steps and exit the wizard.
9. Ensure that all components required for the DAS service start successfully.
10. Make sure to restart all the affected services in Ambari.

What to do next

After all the services are up after the restart, you can verify whether the upgrade was successful by checking the DAS version on the Ambari UI.