

Administration 1

DataPlane Platform Administration

Date of Publish: 2018-12-14



<http://docs.hortonworks.com>

Contents

Working with clusters in DataPlane.....	3
Add a cluster.....	3
Enable a cluster for use.....	4
Edit cluster details.....	5
View cluster details.....	5
Remove a cluster from DataPlane.....	6
Advanced: Add host entries to the DataPlane environment.....	6
Advanced: Upload SSL certificate to DataPlane.....	7
Managing Users and Groups.....	7
Add a user or group.....	8
Edit a user or group.....	9
Edit LDAP settings.....	10
Managing DP Apps.....	10
Install and Start DP Apps.....	10
Navigating between services.....	11
Recover a DP Instance.....	11
Prerequisites for moving a DP Instance.....	12
Initialize the target DP host.....	12
Configure the target DP instance.....	13
Disable and enable data telemetry.....	13
Roles required for working with DataPlane.....	14
DataPlane Admin role.....	14
Roles required for installation and troubleshooting.....	15
Troubleshooting DataPlane administration.....	15
Endpoint not accessible.....	15
Logging in using the DataPlane local admin role.....	15
Ranger UI does not display deny policy items.....	15
Embedded database is lost due to docker storage and mounting issues.....	16

Working with clusters in DataPlane

From DP Platform, you can add clusters, view cluster details, and enable a cluster to be used with a DP app.

Operation	Description
Add a cluster	To make a cluster available for use in DataPlane.
Enable a cluster	To enable a cluster for use with a DP app service.
View and edit cluster details	To view and modify clusters that have been added to DataPlane.

Add a cluster

You must register clusters with DataPlane instance before you can use those clusters with DP Apps.

About this task



Caution: After you register a cluster in DP Platform, do not change the cluster name in Ambari. A cluster name change in Ambari does not currently propagate to DataPlane, which can result in issues when using clusters with DP Apps.

Before you begin

- You must be logged in using the DataPlane Admin (DP Admin) role with valid LDAP credentials, to perform this task.
 - If your clusters are configured for Knox Gateway and you are using a self-signed SSL certificate with Knox, you must have uploaded your certificate to DataPlane prior to starting this task (or you can disable certificate validation on registration). Refer to “Upload a certificate to DataPlane” for more information.
 - Clusters must have been created using Apache Ambari.
- Clusters that are not managed by Ambari cannot be registered to DataPlane.
- All clusters must meet the requirements identified in DataPlane *Getting Started*.

Procedure

1. In the DP Platform navigation pane, click the  (Clusters) icon.
2. Click **Add Cluster**.
3. (Optional) Check **Ambari and Cluster Services Behind Knox Gateway** only if you are using Knox Gateway to proxy Ambari, cluster services(including DP App agents).
4. (Optional) Check **Validate the SSL certificate and only allow trusted connections** only if your cluster is SSL-enabled.

When checked, DataPlane validates the SSL certificate.

If you plan to validate the SSL certificate but the certificate is self-signed, be sure to upload the certificate to DataPlane. Refer to “Upload a certificate to DataPlane” for more information.

5. Enter the URL of the Ambari host that manages the cluster.
 - If the cluster is behind a Knox Gateway, use the following format:

```
https://knox_host_fqdn:knox_port/gateway_name/dp-proxy/ambari
```

By default, the gateway_name is gateway.

- If the cluster is not using Knox Gateway, use the following format:

```
https://ambari_host_fqdn:ambari_port/
```

You can use HTTP or HTTPS. HTTPS is used if Wire Encryption is enabled for Ambari.

You can also enter the IP address instead of the FQDN, but the FQDN is recommended.

The host name must either be in a DNS server or configured in the etc/hosts file.



Important:

DP Platform host must be able to resolve and reach the Ambari URL, whether you are using the FQDN or the IP address. That is, you should be able to use curl or wget to access the Ambari URL from the DP Platform host. If this requirement is not met, cluster registration fails.

If host names are resolved from /etc/hosts, you should explicitly register the cluster host names on the DataPlane container before the cluster is registered with DataPlane.

6. Click Go.

The cluster name, IP address, and enabled cluster services are displayed.

7. Add the following cluster details:

Cluster Location

Data Center

Tags (optional)

Description (optional)

What to do next

You can now enable the registered cluster for the DP App you want to use it with.

Related Concepts

[Configure Knox Gateway for DataPlane](#)

[Advanced: Upload SSL certificate to DataPlane](#)

Enable a cluster for use

After registering clusters with DataPlane, you must enable the clusters that can be used with DP apps. Each DP app has specific configuration requirements that a cluster must meet before it can be used with the app.

About this task

When you enable a DP app, a check is run to determine if the required engine or agent is installed on any clusters. If the engine is installed but some configuration is still required, the cluster displays on the Services page with the button Enable. If the cluster meets all requirements for the service it is automatically enabled, and the enabled cluster can only be viewed on the Services page by selecting the Show All Clusters action for the service.

Before you begin

The DataPlane Admin role is required to perform this task.

Clusters must be managed by Apache Ambari and registered with DataPlane.

The services you are associating with the clusters must already be enabled in DP Platform.

Procedure

1. Click the  (Services) icon in the DP Platform navigation pane.
The Services page displays. Services listed in the table have been enabled. Services identified by a tile icon are available to be enabled.
2. Click a service.

A list displays of any clusters that have the required service engine or agent installed but have not yet been configured for use with the service.



Tip: If no clusters display for the service, verify that the clusters you expect to see have been registered with DP Platform, and that the proper service engine or agent has been installed on the clusters.

3. Click **Enable** for the cluster you want to use with the service.

A check runs to determine what configuration is required on the cluster for the service you selected. For example, a required service such as Apache Ranger might be installed on the clusters but not enabled in Apache Ambari.

Edit cluster details

You can modify the Cluster Location, Data Center name, Tags, or Description for any cluster registered in DP Platform.

Before you begin

The DataPlane Admin role is required to perform this task.

About this task



Important: After you register a cluster in DP Platform, do not change the cluster name in Ambari. A cluster name change in Ambari currently does not propagate to DataPlane which can result in issues when using clusters with DP apps.

Procedure

1. Click the  (Clusters) icon in the DP Platform navigation pane.
2. Optional: Enter a cluster name in the search field and press **Enter**.
You can only search by cluster name. You can search by partial or full names.
3. In the cluster list, locate the row for the cluster you want to edit.
4. At the end of the row, click the  (Actions) icon and then click **Edit**.
The Edit Cluster page displays.
5. Modify the cluster details.
6. Click **Update**.
The Clusters page displays a list with the updated cluster.

View cluster details

DP Platform provides two levels of detail about each cluster.

About this task

- You can view general information about a cluster such as status, location, HDP or HDF version, number of nodes, and so forth, from the Clusters page.
- You can view information such as the status of DataNodes and NodeManagers, heap size, disk space, and so forth, from the cluster Details page.

Before you begin

The DataPlane Admin role is required to perform this task.

Procedure

1. Click the  (Clusters) icon in the DP Platform navigation pane.
2. In the cluster list, locate the row for the cluster you want to edit and click the cluster name. The Details page displays more information about the selected cluster.
3. Refresh the values on the Details page by clicking the  (Refresh) icon. Updated data is pulled from Ambari.
4. To view or edit cluster information in Ambari, click the  (Actions) icon and then click **Go to Ambari**. A new browser tab opens to the login page for the Ambari host that manages the cluster.

Remove a cluster from DataPlane

If you want to remove a cluster from DataPlane after installing and configuring DataPlane, use the `rm_dp_cluster.sh` command to remove the cluster from DataPlane.

About this task

Perform the following steps:

Procedure

1. Navigate to this folder:
`/usr/dp/current/core/bin`
2. Run the following command.

```
sh rm_dp_cluster.sh <DP_JWT> <DP_HADOOP_JWT> <DP_HOST_NAME> <CLUSTER_NAME>
<DATA_CENTER_NAME>
```

The parameters are as follows:

- `<DP_JWT>` - Value of the `dp_jwt` cookie from a valid user's browser session.
 - `<DP_HADOOP_JWT>` - Value of the `dp-hadoop-jwt` cookie from a valid user's browser session.
 - `<DP_HOST_NAME>` - Hostname or IP address of the DataPlane server.
 - `<CLUSTER_NAME>` - Name of the cluster to be deleted.
 - `<DATA_CENTER_NAME>` - Name of the data center the cluster belongs to.
3. Enter `yes` to continue.

Advanced: Add host entries to the DataPlane environment

If you are using hosts that are not publicly addressable from a DNS server, you must add IP address and host name mapping to the `/etc/hosts` file of each associated *DataPlane container* in your DP instance. DataPlane provides a utility command to help with this to ensure that all hosts to be used with DataPlane are addressable.

About this task

You must use the specific method and format identified in this task to make the mapping information usable by DP Platform.

Perform this task only after DP Platform is installed and running.



Attention:

This task is required only for `/etc/hosts` handling. This setup is not required if Ambari, Knox, or hosts in your cluster are accessible via DNS.

Procedure

1. Log in as the root user to a terminal on the DataPlane host.
2. Type the following command to add the host to the /etc/hosts file of the container:

```
dpdeploy.sh utils add-host <ip> <host>
```

3. Repeat this procedure on each cluster registered with DP Platform.

Advanced: Upload SSL certificate to DataPlane

If you are registering clusters that are using a self-signed SSL certificate (for Ambari or Knox Gateway), you can upload the certificate to DataPlane. This enables DataPlane to validate the certificate during cluster registration. Alternatively, you can disable certificate validation during cluster registration.

Procedure

1. In the Navigation pane, click **Settings**.
2. Click **Upload** and complete the form.
 - Certificate Name: Create a name for the certificate.
 - Certificate File: Browse to the location of the certificate, such as a .pem file.

Admin / Settings

Certificate Settings

<input type="checkbox"/>	Certificate Name	Certificate
No certificates added. Please click 'Upload' to add a certificate		

Upload

Certificate Name*

Certificate File*

 Choose

3. Click **Upload**.

Managing Users and Groups

From the DP Platform UI, you can add and edit users and groups for all apps associated with DataPlane. Users and groups must exist in the corporate LDAP directory before you can add them in DP Platform.

Add a user or group

When you install DataPlane, the DataPlane Admin role is created. This role has access to DP Platform and all DP Apps. You should add additional users with permissions limited to the service or services you want the user to access.

About this task

You can also add groups, which enable you to more easily manage users. You might configure groups for all users who can perform specific tasks, such as creating replications, creating datasets, and so forth.



Tip: User-level assignments override group-level assignments.

Before you begin

User and group accounts must already exist within your corporate LDAP directory.

The DataPlane Admin role is required to perform this task.

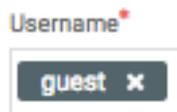
Procedure

1. Click the  (Users) icon in the DP Platform navigation pane.
2. Click **Add User**.
3. In **Username**, enter the name of a user from your corporate LDAP directory, and then click the name when it pops up.



Tip:

You must click the name of the user when it displays and ensure it appears in the Username field on a dark background.



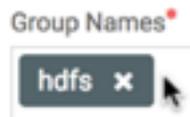
If the name appears on a white background, it means the name is not recognized and the action fails.

4. Select one or more roles to assign to the user.
 - **DataPlane Admin (or DP Admin)**
Can perform all actions in DP Platform, and can access and perform all actions in the UI of enabled services.
 - **Infra Admin**
Can perform all actions in the Data Lifecycle Manage (DLM) service UI, and can manage DLM-enabled clusters in DP Platform.
 - **Data Steward**
Can perform all actions in the Data Steward Studio (DSS) service UI, and can manage DSS-enabled clusters in DP Platform.
5. Click **Save**.
The new user displays in the list on the Users page.
6. Click the **Groups** tab, and then click **Add Group**.
7. In **Group Names**, enter the name of a group from your corporate LDAP directory, and then click the name when it pops up.



Tip:

You must click the name of the group when it displays and ensure it appears in the Group Name field on a dark background.



If the name appears on a white background, it means the name is not recognized and the action fails.

8. Select the roles to assign to the group.

9. Click **Save**.

The new group displays in the list on the Groups page.

Edit a user or group

You can edit any user or group that has already been added to DP Platform.

Before you begin

The DataPlane Admin role is required to perform this task.

About this task

User-level assignments override group-level assignments.

Procedure

1. Click the  (Users) icon in the DP Platform navigation pane.
The Users page displays a list of existing users, their roles, and status.
2. Locate the user you want to edit by browsing the user list or entering a user name in the search field.
You can only search by user name. You can search by partial or full names.
3. Select the user to edit by doing one of the following:
 - Click a user name in the list of users.
 - Click  (Actions icon) and Edit.
The user's information displays in a slide-out panel.
4. Change the user status, or add or delete roles.
5. Click **Save**.
The modifications are shown in the Users list.
6. Click the **Groups** tab.
7. Locate the group you want to edit by browsing the group list or entering a group name in the search field.
You can only search by group name. You can search by partial or full names.
8. Select the group to edit by doing one of the following:
 - Click a group name in the list of groups.
 - Click  (Actions icon) and Edit.
The group's information displays in a slide-out panel.
9. Add or delete roles assigned to the group.
10. Click **Save**.
The modifications are shown in the Groups list.

Edit LDAP settings

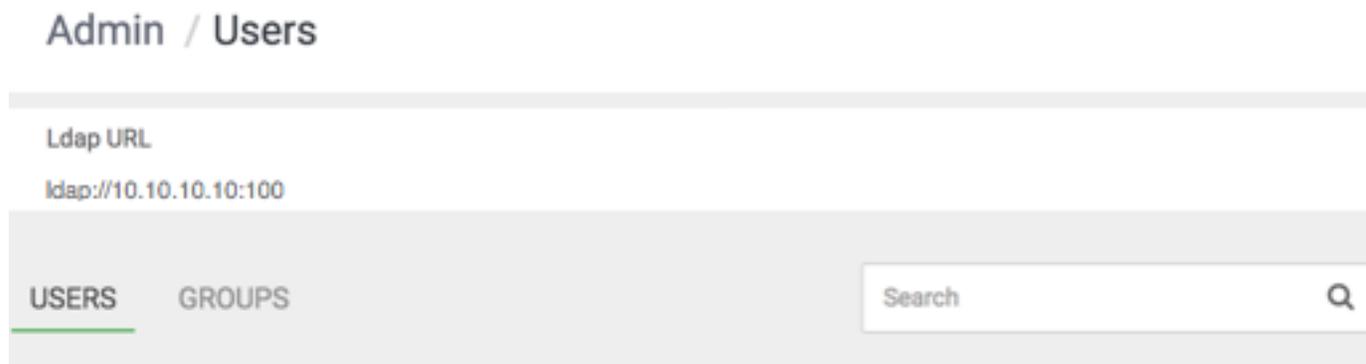
If the corporate LDAP administrator Bind DN and password are changed, such as for security purposes or policy requirements, or the LDAP server URL changes, such as due to hardware issues, the DataPlane Admin can modify those settings in DataPlane so that users can continue to log in to DataPlane.

Before you begin

You must be logged in as the DataPlane Admin to complete this task.

Procedure

1. In the DP Platform UI, click the  (Users) icon.
2. To the right of the LDAP URL, click the  (Actions) icon, and then click **Edit**.



3. In the Edit LDAP page, you can modify the following settings:
 - URL
 - Administrator Bind DN
 - Administrator Password
 - Upload a new certificate
4. Click **Save**.

Managing DP Apps

DataPlane supports running one or more DP apps that you can install and configure for use with your registered clusters.

The service app life cycle includes the following phases:

1. Install the required components for each service in your clusters.
2. Add your clusters to DataPlane.
3. For each service app, enable clusters for use.
4. Access the service app from the service navigation icon or the Services menu from the DataPlane dashboard.

Install and Start DP Apps

To start DP apps, you must enable clusters for the service.

Procedure

1. From the Admin page, click Services  (Services) icon on the left panel.

The Services page displays.

2. On the Services page, you can see the available service apps.
3. Click on a service app to view the list of available clusters.



Note: Make sure you install the required components in the clusters before you enable the cluster for any service. See the service app specific documentation for more details of required components.

4. In the Actions column, select the cluster that you want to enable for the service and click Enable. Thus, you can add clusters for each service.

Navigating between services

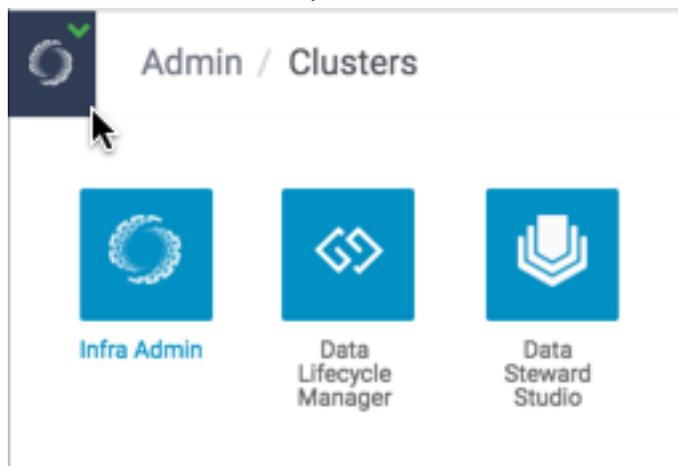
You can access any service for which you have been assigned the proper role. The DataPlane Admin has access to all DataPlane services.

Before you begin

The DataPlane Admin must have assigned you the required role for any service you want to access.

Procedure

1. Click the  (Service Navigation) icon in the upper left corner of any page in DP Platform.
2. Click the tile for the service you want.



If the service you want to access is not displayed, either the service is not enabled or you have not been assigned the role required to access the service. The DataPlane Admin can enable services and assign roles.

Recover a DP Instance

Perform the pre-requisite tasks, initialize the target host, and configure the target DP instance.

Procedure

1. Perform the pre-requisite tasks.
2. Initialize the target DP host.
3. Configure the target DP instance.

Prerequisites for moving a DP Instance

Make sure the following prerequisites are met before you recover the instance.

Requirement	Description
External database	Your DP Instance configured for an external database, not the default embedded PostgreSQL. See Configure an external database for more information.
Local configuration backup	Perform periodic backup of the following file: <code>/usr/dp/current/core/bin/config.env.sh</code>
Local keystore Backup	Perform periodic backup of the following directory, especially after any operation that involves saving/editing your LDAP setup or cloud credentials: <code>/usr/dp/current/core/bin/certs/*</code>
Target DP Host	Prepare a second host (the target) that has the same versions of DP Platform and DP apps software installed on the original host. Do NOT perform initialization of the target host until instructed.

Initialize the target DP host

Procedure

1. Copy the public and private key files from the source host to the target host:

```
/usr/dp/current/core/bin/certs/ssl-cert.pem
/usr/dp/current/core/bin/certs/ssl-key.pem
```

2. On the target host, edit `/usr/dp/current/core/bin/config.env.sh` file and set to the above:

```
DATAPLANE_CERTIFICATE_PUBLIC_KEY_PATH
DATAPLANE_CERTIFICATE_PRIVATE_KEY_PATH
```



Note: Leave `USE_EXTERNAL_DB` set to no and do not configure the target host to use the external database at this time. Otherwise, the external database will be reset, losing your DP information.

3. On the target host, initialize the deployment.

```
dpdeploy init --all
```

4. Stop the DP instance on source host if it is still running so that no new changes are happening to the instance.

```
dpdeploy stop
```

5. Copy the keystore from the source host to the target host:

```
/usr/dp/current/core/bin/certs/dp-keystore.jceks
```

6. Replace `config.env.sh` from `/usr/dp/current/core/bin/` of target host with `/usr/dp/current/core/bin/config.env.sh` of the source host.

7. On the target host, stop the DP instance:

```
dpdeploy stop
```

8. Start the DP instance on the target host and be sure to enter the same Master Password that was used with your source host:

```
dpdeploy start
```

Configure the target DP instance

Update the Knox SSO settings.

Procedure

1. After the containers are started on the target, log in with Super Administrator credentials configured in the source host and navigate to the Admin / Users page.

```
https://<target.hostname>/sign-in
```

2. Click **Edit** to modify the LDAP setup. Enter the LDAP Administrator password and click **Save**.
3. On the target host, execute the following to copy the Knox SSO topology file out of the running container.

```
docker  
cp Knox:/usr/hdp/current/knox-server/conf/topologies/knoxssso.xml .
```

4. Open `knoxssso.xml` and change the `knoxssso.redirect.whitelist.regex` entry as follows to include target host in Knox whitelist.

```
<param>  
  <name>knoxssso.redirect.whitelist.regex</name>  
  <value>^https?:\\\/(\<target.hostname>|localhost|127.0.0.1|  
0:0:0:0:0:0:0:1|::1)(:[0-9])*.*$</value>  
</param>
```

5. Copy the Knox SSO topology file into the running container:

```
docker cp Knoxssso.xml  
knox:/usr/hdp/current/knox-server/conf/topologies/knoxssso.xml
```

6. Log in to the target host as usual and verify everything is working fine.

```
https://<target.hostname>/sign-in
```

Disable and enable data telemetry

As part of the installation process, data collection using cookies and other telemetry mechanisms is turned on by default. This topic describes how to disable tracking and telemetry, and enable or check the status of telemetry.

About this task

Use the `dpdeploy help` command to learn more about `dpdeploy` command options.

Procedure

1. To disable tracking, in a terminal enter the following command:

```
./dpdeploy.sh utils disable-config dps.ga.tracking.enabled
```

You see the following output:

```
UPDATE 1  
Config value for key dps.ga.tracking.enabled was updated successfully with  
value false
```

2. To enable tracking, in a terminal enter the following command:

```
./dpdeploy.sh utils enable-config dps.ga.tracking.enabled
```

You see the following output:

```
UPDATE 1  
Config value for key dps.ga.tracking.enabled was updated successfully with  
value true
```

3. To enable or disable getting tracking status, in a terminal enter the following command:

```
./dpdeploy.sh utils get-config dps.ga.tracking.enabled
```

You see the following output if enabling tracking status:

```
dps.ga.tracking.enabled: Enabled
```

You see the following output if disabling tracking status:

```
dps.ga.tracking.enabled: Disabled
```

Roles required for working with DataPlane

Access to DP apps and functionality within those apps requires a different role or set of roles for each service.

To perform actions in DP Platform and associated services, you must be logged in as DataPlane Admin (DP Admin) user.

DataPlane Admin role

The DataPlane Admin has access to DP Platform and administrative permissions to perform all actions in DP Platform. A DataPlane Admin role is created during installation, so you can initially log in to DP Platform.

The DataPlane Admin has the following capabilities and restrictions:

- Can access DP Platform and perform all actions in DP Platform related to clusters, users, and enabling services.
- Can access all services enabled with DP Platform, and perform the same actions as each administrator role assigned to the enabled services, such as Infra Admin, Data Steward, and so forth.

In addition to this role, there are app-specific admin roles that are needed to manage the enabled services. For more information, see the app-specific documentation.

Roles required for installation and troubleshooting

You need the Ambari Admin or Cluster Admin roles to install DataPlane, add clusters to Ambari, troubleshoot cluster issues, and so forth.

To perform actions in Apache Ambari that impact DataPlane (such as creating clusters, changing configuration settings for services, and so forth) too, you must be an Ambari administrator or a cluster administrator.

See [Apache Ambari Administration](#) for further details about these roles.

Troubleshooting DataPlane administration

You can troubleshoot various issues encountered while performing operations on DataPlane.

Endpoint not accessible

DLM does not support updating any cluster endpoints (HDFS, Hive, Ranger, or DLM Engine).

If an endpoint must be modified, contact Hortonworks Support for assistance.

Logging in using the DataPlane local admin role

The local admin role allows you to perform administrative activities and troubleshoot problems when access through LDAP and Knox is not available. The local admin is also the role you use to log in to DataPlane the first time, before LDAP is configured in DataPlane for SSO.

About this task

When you log in as the local DataPlane Admin, you bypass Knox.

For login, the default username is “admin”. The password you use to log in is set during the installation process.

Procedure

Log in by appending /sign-in to the DataPlane login URL, for example:

```
http://dataplane-host-name/sign-in
```

Ranger UI does not display deny policy items

When a policy with deny conditions is created on Ranger-Admin for a replication relationship, the Policy Details page in Ranger should show the deny policy items also. However, the deny policy items do not display on the Ranger admin Policy Details page.

Procedure

1. Enable deny conditions for policies from **Ambari>Ranger>Configs>Advanced>Custom ranger-admin-site**.
2. Add `ranger.servicedef.enableDenyAndExceptionsInPolicies=true`.
3. Restart target ranger-admin.

Embedded database is lost due to docker storage and mounting issues

DSS configuration data is sometimes lost during a `dpdeploy.sh` restart or stop if you are using the embedded database for production purposes.

About this task

Due to issues with the Docker storage engine, conflicts with other processes can prevent the mounted storage under a docker container from being unmounted correctly. This can result in the loss of data in the embedded database.

To prevent this problem from occurring:

1. Avoid using the embedded database for production purposes.
2. Regularly backup and create a local copy of the embedded Postgres database to allow for graceful recovery.

To perform backup and recovery of the database:

Procedure

1. Back up the database.

```
docker exec -u postgres dp-database pg_dumpall -c > dump_`date +%d-%m-%Y"_"%H_%M_%S`.sql
```

2. Remove the docker container.

```
docker rm -f dp-database
```

3. Remove the docker volume.

```
docker volume rm postgresql-data
```

4. Start the database container.

```
sh docker-database.sh
```

5. Restore the database.

```
cat <Name of the backup file.sql> | docker exec -i dp-database psql -U dp_admin -d dataplane
```