DgSecure – Installation Guide for Linux

Version 7.2

Dataguise, Inc.
39650 Liberty Street, Suite 400

Fremont, CA 94538
877-632-0522

http://www.dataguise.com

Contents

1. Introduction

DgSecure is a complete data security solution that enables enterprises to leverage their data to achieve greater business goals while minimizing the risk of exposure and running afoul of data handling regulations such as PII, PCI, HIPAA and GDPR.

This document is intended to guide users through installation, configuration, and administration of DgSecure. By the end of this document, the user should be clear on basic configuration options and be able to clearly administer the product.

1.1 Organization

This installation guide is structured to reflect different activities workflow. The chapters are:

- Installation
- Configuration
- DgAdmin

1.2 Related Sources

This document is intended to support user deployment of DgSecure on the customer site. Other documents are also available:

- *DgSecure User Guide* – Orients system users and introduces general concepts within DgSecure's various solutions (DBMS, Files, Hadoop).
- *DgSecure REST API Reference* – Describes in detail the REST APIs that are supported by DgSecure.
- *Online Help* – Accessible from any screen in DgSecure, it shares in-depth information on each screen as well as context specific tips. Once the controller, repository database, and IDPs are installed, and configured, DgSecure can be launched using a web browser or from the Start menu. DgSecure is bundled with a default set of software it needs to function.

2. Installation

A DgSecure installation consists installation of a controller, repository database, the monitoring subsystem, and one or more *Intelligent Data processors (IDPs)*. There are two types of deployment scenarios:

1. Single Host Machine Deployment
2. Distributed Deployment

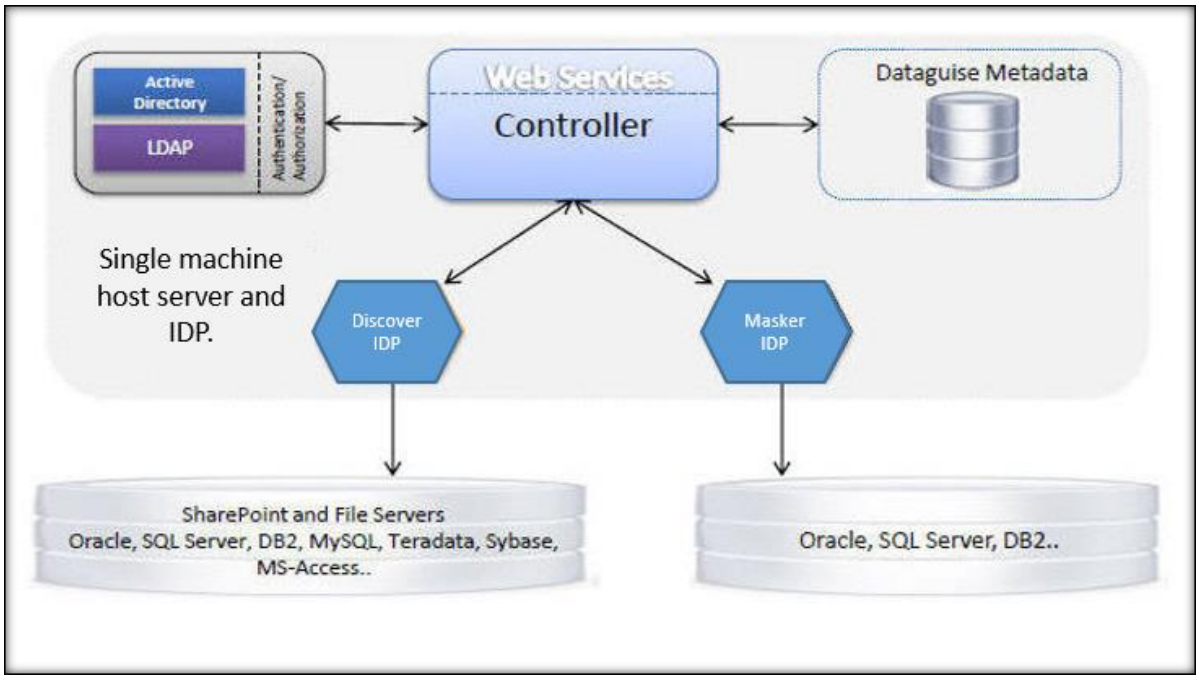
2.1 Deployment Scenarios

The IDPs should be installed on hosts from where they can connect to the data stores that need to be scanned or masked. The exact location depends on how a customer's network and array of machines are set up and utilized.

Below are two types of deployment:

2.1.1 Single Host Machine

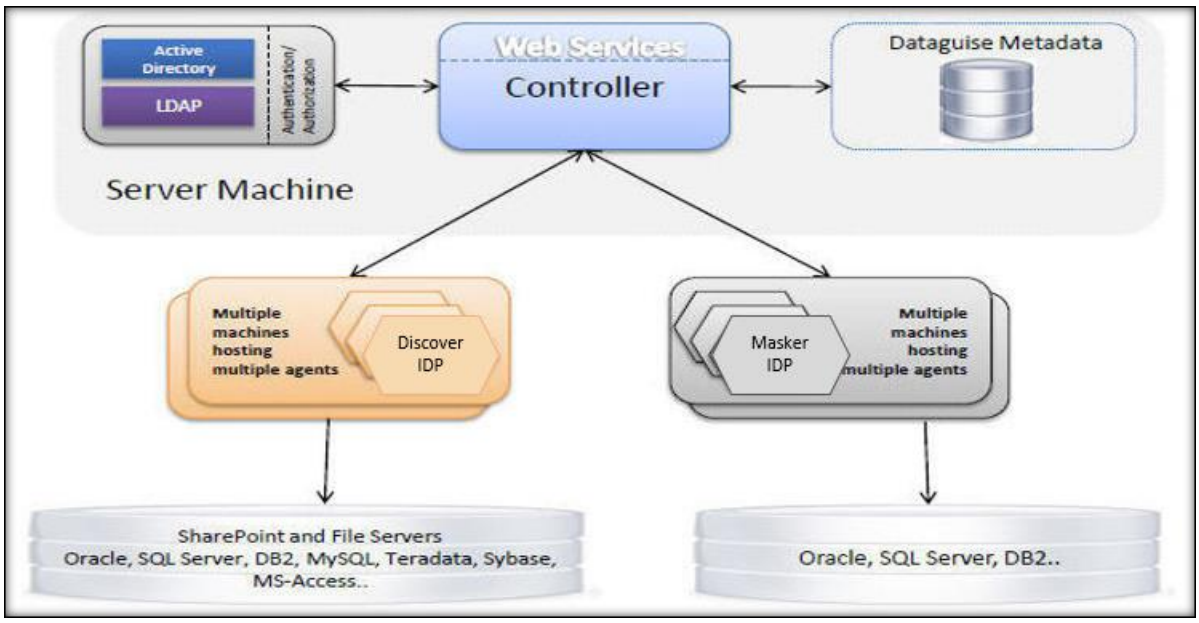
If there is a single, robust server with access to all the target databases and file systems, all of the DgSecure components can be installed on that server. A single host installer is available for all IDPs.



2.1.2 Distributed Deployment

On a segmented or widely-distributed network, multiple IDPs may need to be installed to reach all the target data. This also helps distribute the DgSecure IDP

workload across multiple servers. In this configuration, the controller can be installed on one machine and the IDPs on one or more remote machine(s).



2.2 System Requirements

2.2.1 Controller Host Machine

This software distribution bundles certain additional applications required to run DgSecure. One is Apache Tomcat (Tomcat 9.x) which is installed automatically along with the controller. Another is PostgreSQL server which is also installed automatically unless you choose to use MySQL/Oracle/SQL Server, which must have already been installed.

- Operating System 64 bit: CentOS 6.7 or 6.9 or 7.2, RHEL 6.5, RHEL 7.2, OpenSuSe 42.1
- Java Runtime Environment: JRE v8 (Oracle or Open JRE)
- Authentication Systems: Options include Active Directory Server, LDAP, and DB Authentication
- Metadata Repository: Options include PostgreSQL (Bundled with distribution), MySQL, Oracle, and SQL Server.

Following system configurations are required to install and use DgSecure:

Minimum System Requirements:

	No of VMs	CPUs	Memory (GB)	DG Disk Space (GB)	Operating Systems
DgSecure controller machine	1	8	16	60	64 bit CentOS 6.3 or higher, RHEL 6.3 or higher, 64 bit Windows 2003, 2008/R2, 8, 10

Recommended System Requirements:

	No of VMs	CPUs	Memory (GB)	DG Disk Space (GB)	Operating Systems
DgSecure controller machine	1	8	32	60	64 bit CentOS 6.3 or higher, RHEL 6.3 or higher, 64 bit Windows 2003, 2008/R2, 8, 10

Note: When installing with MySQL as the backend (metadata) repository, ensure the minimum packet size is set to 1GB.

2.2.2 IDP Host Machines

The IDP host machine must have access to the machines on which data you intend to scan, or mask are hosted. For example, if firewalls are set up around sub-nets that the data host machines are part of, appropriate ports should be open for the IDPs to perform the scanning/masking operations. System requirements for the various IDPs are listed below.

All IDPs are compatible with:

Minimum System Requirements:

	No of VMs	CPUs	Memory (GB)	DG Disk Space (GB)	Operating Systems

DgSecure IDP Host Machine	1	4	8	100	64 bit CentOS 6.3 or higher, RHEL 6.3 or higher, 64 bit Windows 2003, 2008/R2, 8, 10

Recommended System Requirements:

	No of VMs	CPUs	Memory (GB)	DG Disk Space (GB)	Operating Systems
DgSecure IDP Host Machine	1	4	16	100	64 bit CentOS 6.3 or higher, RHEL 6.3 or higher, 64 bit Windows 2003, 2008/R2, 8, 10

When using Format Preserving Masking (FPM - available for Hive, HDFS, and Files), RDBMS/RDS / Java 8 is required on the IDP machine.

Additional Requirements for HDFS IDP on Azure Blob Storage:

- Recommendation: the IDP should be deployed on an edge node instead of the head node or a worker node.
- In order to run tasks on storage accounts not linked with a cluster, the access policy on the container should be "container".

Additional Requirements for DgSecure's Monitoring Sub-System:

- Logstash (for Monitoring on S3, HDFS & Hive)
- MySQL 5.5 or above is required to be used as DgSecure's metadata repository in order to use DgSecure Monitor on S3, HDFS & Hive.
- In order to monitor MapR, the Monitoring sub-system must be installed on the MapR cluster node, or on a non-cluster node with the MapR client installed.

Note: IDPs can be installed on the same machine as the controller or on another machine.

2.3 Prerequisites

1. Uninstall any existing DgSecure systems before installing.

2. Ensure that the Metadata repository-Oracle server is already installed on the same machine on which DgSecure is installed, or the correct access is available from that machine.
3. Ensure that you have the required license key and challenge password which will be used during initial setup and configuration.
4. Make sure the machine(s) on which the IDP(s) are installed, have access to the target database(s), and/or file system(s).

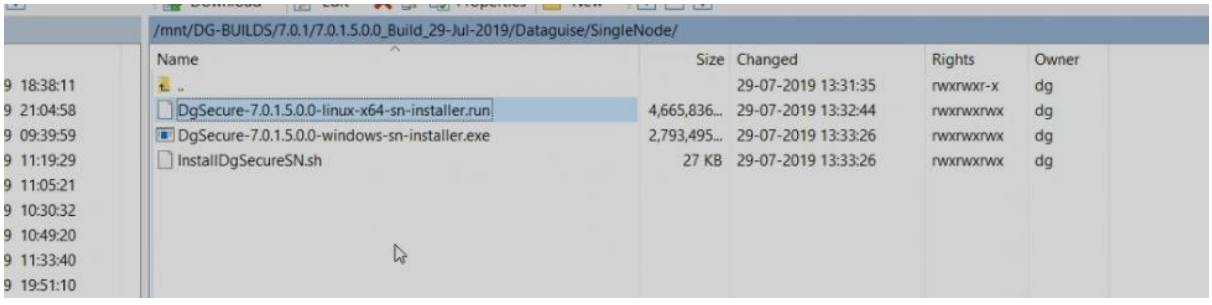
2.4 Single Node Installation

The Single Node procedure has the options to install the following DgSecure components on the single host machine:

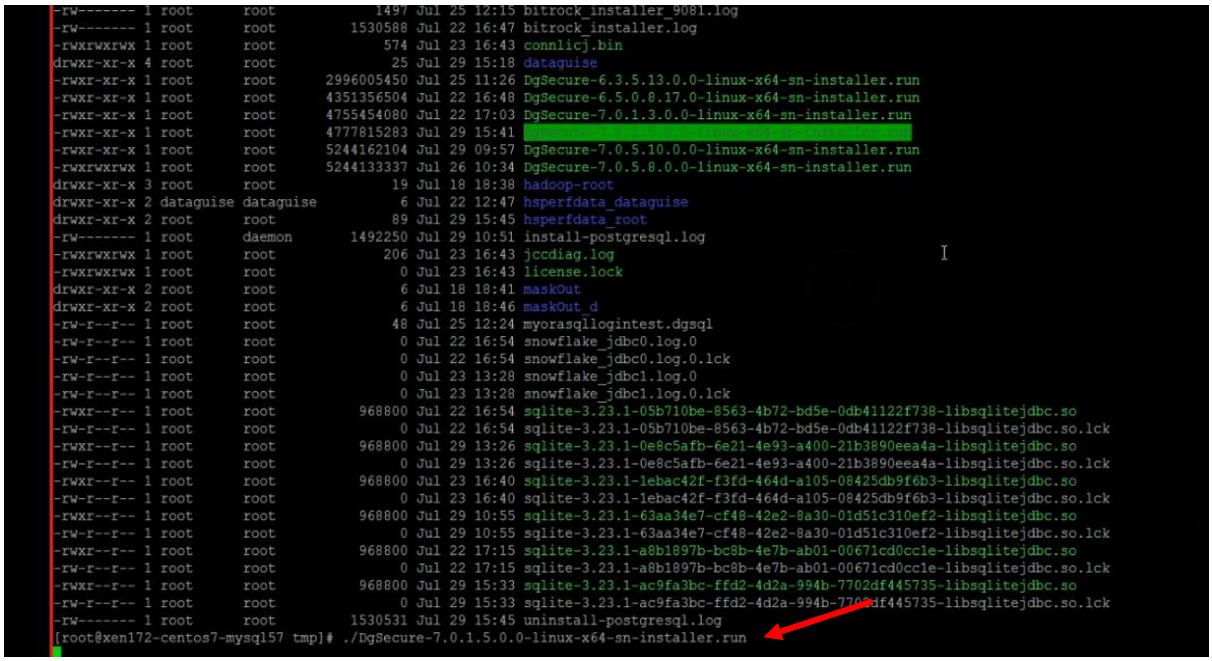
- DgSecure controller
- Postgres Database
- Discover IDP
- Masker IDP
- Files IDP
- Command Line utility
- Hadoop IDP
- Hive IDP
- Hadoop Control IDP
- GDPR IDP

To install DgSecure

i. Copy the DgSecure Single Node Installer for Linux to the machine on which you will install the DgSecure components.

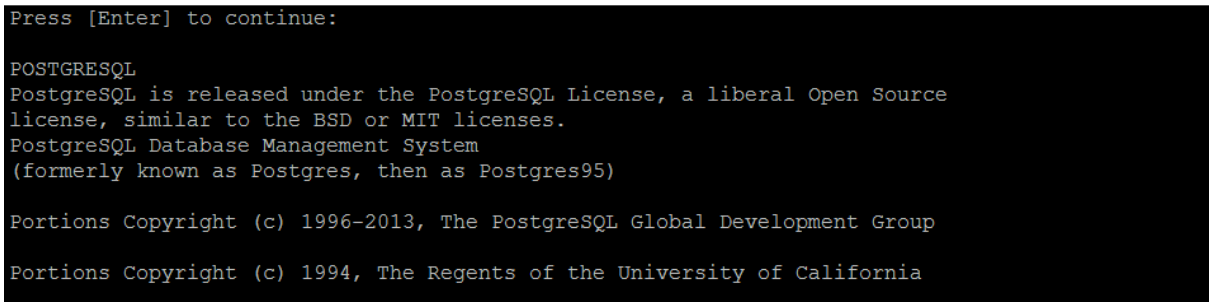


ii. Run the installation file. DgSecure-xxx-linux-x64-sn-installer.run.

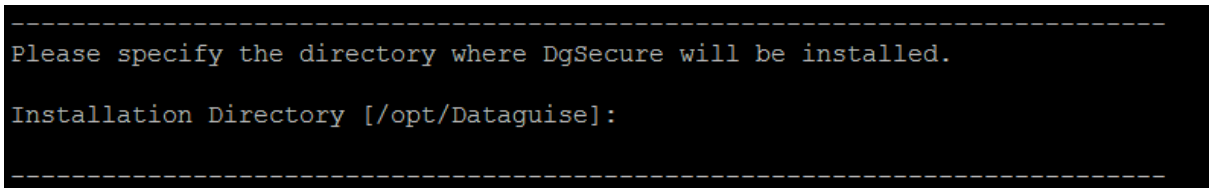


Note: The installer by default, runs as "Root User". To run the non-root installer, user will require 777 permissions on the /tmp directory if installing with PostgreSQL metadata repository (in order to copy the postgres script). Also, once installation is complete, the user will need to manually start the DgSecure IDP services, as they are not installed by the non-root user. These can be started directly from the IDP directories.

iii. Accept the license agreement. Press Enter to continue.



iv. Provide the path to the directory where DgSecure will be installed or press Enter to continue with the default directory.



Note: If user does not wish to give complete access to the temporary directory for certain environments, steps to create a new directory are:

A parameter "sys_temp_dir" is available, which can be used to set temporary directory where installer can write, create and execute files during installation, the syntax is:

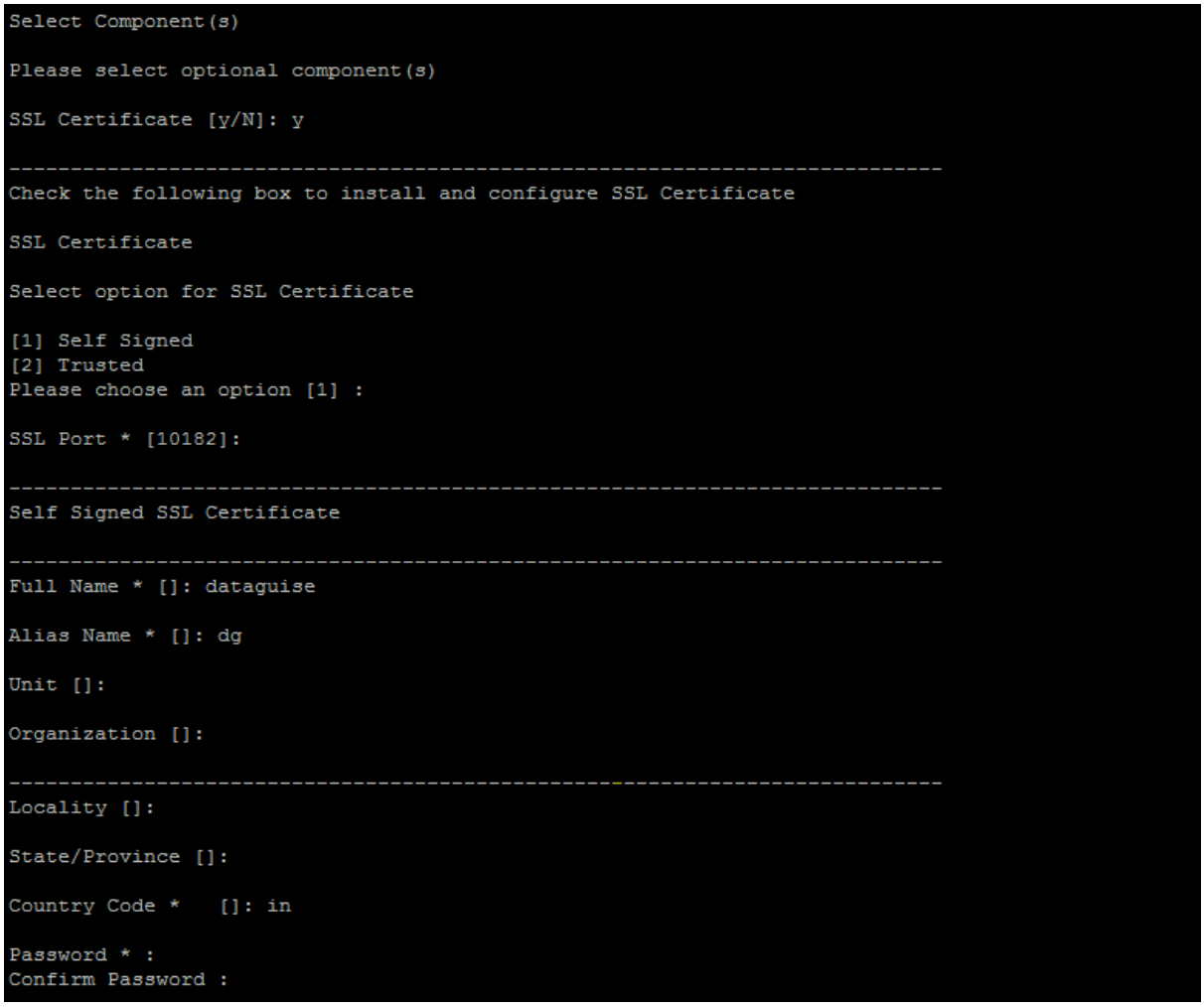
<InstallerName> --sys_temp_dir "<Absolute_path_of_Custom_temp_folder>"

For installation on Linux environment minimum permission requirement is 755 for the folder -"<Absolute_path_of_Custom_temp_folder>".

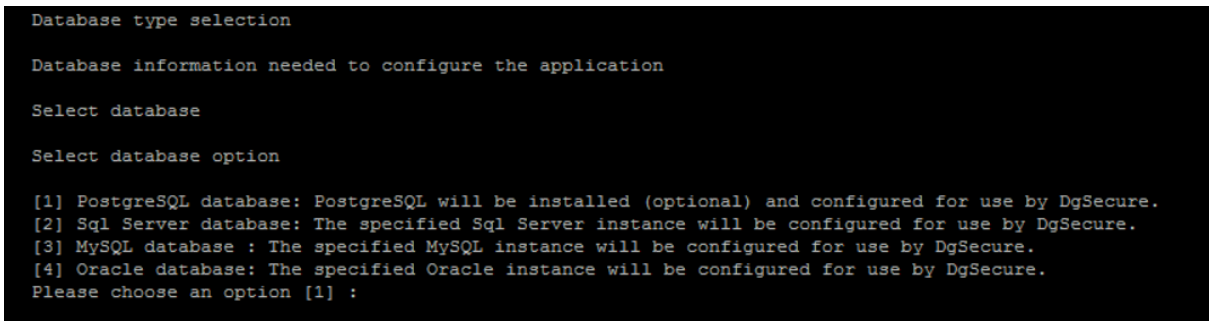
v. Select the IDPs that need to be installed. Enter Y to install the IDP and N if you would like to skip the installation of an IDP. To confirm your selection, enter Y else enter N to change the selection of IDPs.

vi. An SSL certificate enhances security by allowing you to run DgSecure in a secure environment. Skip to step 8, if you wish to proceed without SSL.

vii. Enter Y to provide and chose trusted or self-signed. Provide the SSL port number.



viii. Choose whether you want to install the bundled PostgreSQL database, or use an existing installation of PostgreSQL, MySQL, SQL Server, or Oracle.

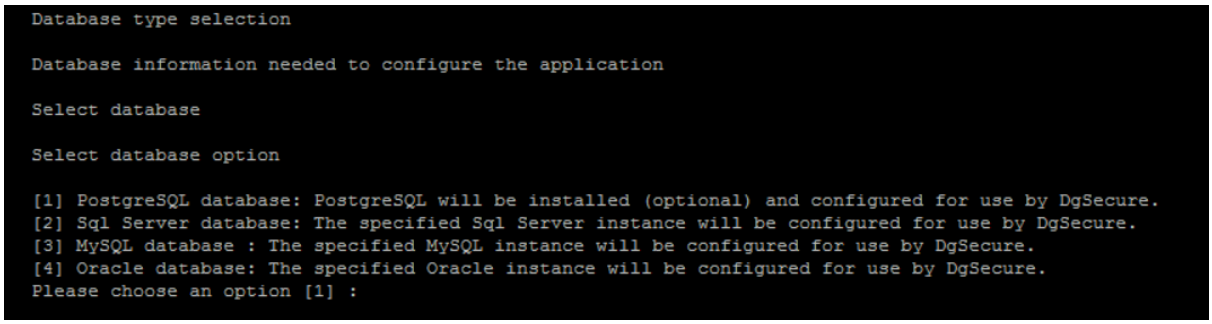


2.4.1 Select the Database Type

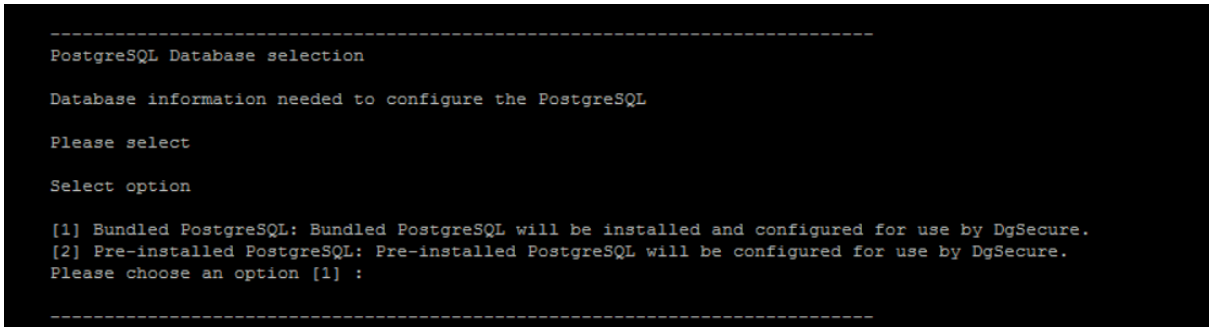
Following are the steps to configure different Database Types.

PostgreSQL

ix. To configure Postgres, enter [1].



x. Select Bundled PostgreSQL or Pre-Installed PostgreSQL.



Note: Incase if bundled PostgreSQL does not exists then create a dummy db named "dataguise".

xi. If user has selected bundled PostgreSQL, then enter the required information.

```
--------------------------------------------------------------------------------
PostgreSQL Configuration

Information needed to configure PostgreSQL

               Port [5432]: 5477

Please provide a password for the database superuser (postgres).

               Password :

               Confirm Password :


               Enable SSL [y/N]:
```

xii. If user has selected Pre-Installed PostgreSQL, then enter the required information (as shown below).

```
PostgreSQL Configuration

Information needed to configure PostgreSQL

Database Name [dg]:

Controller Schema Name [dgcontroller]:

Control Schema Name [dgcontrol]:

HDFS Info Schema Name [dghdfsinfo]:

Dashboard Schema Name [dgstar]:
```

MySQL

ix. To configure MySQL, enter [3].

```
Database type selection

Database information needed to configure the application

Select database

Select database option

[1] PostgreSQL database: PostgreSQL will be installed (optional) and configured for use by DgSecure.
[2] Sql Server database: The specified Sql Server instance will be configured for use by DgSecure.
[3] MySQL database : The specified MySQL instance will be configured for use by DgSecure.
[4] Oracle database: The specified Oracle instance will be configured for use by DgSecure.
Please choose an option [1] : 3
```

x. Provide MySQL Server connection details.

```
--------------------------------------------------------------------------------
MySQL Configuration

MySQL information needed to configure the application

               Port [3306]:

               Host/IP [localhost]: 192.168.0.25

               User Name [root]:

               Password :
```

xi. Select the database option, With Fresh DB/Schemas or With Existing DB/Schemas.

```
  Select database option

  Information needed to configure MySQL

  [1] With Fresh DB/Schemas
  [2] With Existing DB/Schemas
  Please choose an option [1] :

  With Fresh DB/Schemas
```

xii. Provide MySQL Server configuration details.

```
--------------------------------------------------------------------------------
MySQL Configuration

Information needed to configure MySQL

Controller Schema Name [dgcontroller]: A_dgcontroller

Control Schema Name [dgcontrol]: A_dgcontrol

HDFS Info Schema Name [dghdfsinfo]: A_dghdfsinfo

Dashboard Schema Name [dgstar]: A_dgstar
```

SQL Server

ix. To configure SQL Server, enter [2].

```
Select database option

[1] PostgreSQL database: PostgreSQL will be installed (optional) and configured for use by DgSecure.
[2] Sql Server database: The specified Sql Server instance will be configured for use by DgSecure.
[3] MySQL database : The specified MySQL instance will be configured for use by DgSecure.
[4] Oracle database: The specified Oracle instance will be configured for use by DgSecure.
Please choose an option [1] : 2
```

x. If user wants to connect through port, enter [2] and provide the port number.

```
--------------------------------------------------------------------------------
Sql Server Configuration

Sql Server information needed to configure the application

               Server Name [localhost]: 192.168.0.151

Connect Using

               Connect Using

[1] Instance
[2] Port
Please choose an option [1] : 2

                []: 1433

--------------------------------------------------------------------------------
```

xi. Select the database option, With Fresh DB/Schemas or With Existing DB/Schemas.

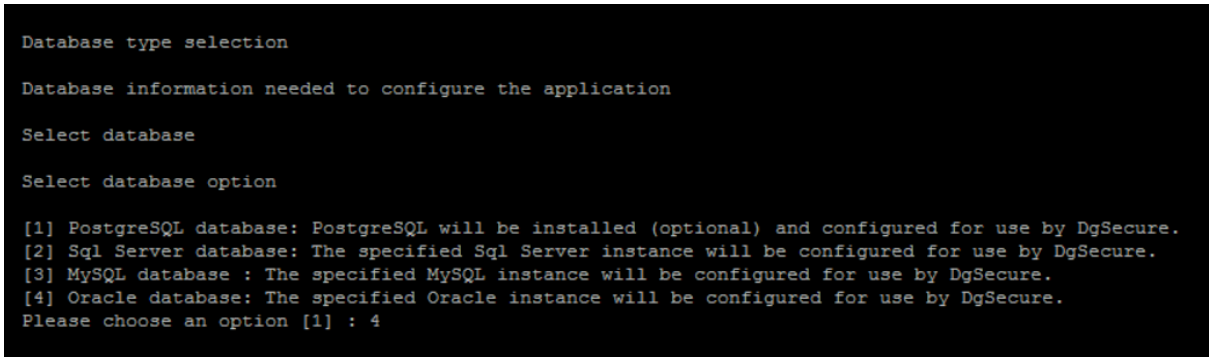
```
--------------------------------------------------------------------------------
Sql Server Configuration

Sql Server information needed to configure the application

               User Name [sa]: sa

               Password :

Select database option

Information needed to configure SQL Server

[1] With Fresh DB/Schemas
[2] With Existing DB/Schemas
Please choose an option [1] :

With Fresh DB/Schemas

--------------------------------------------------------------------------------
```

xii. Provide SQL Server configuration details.

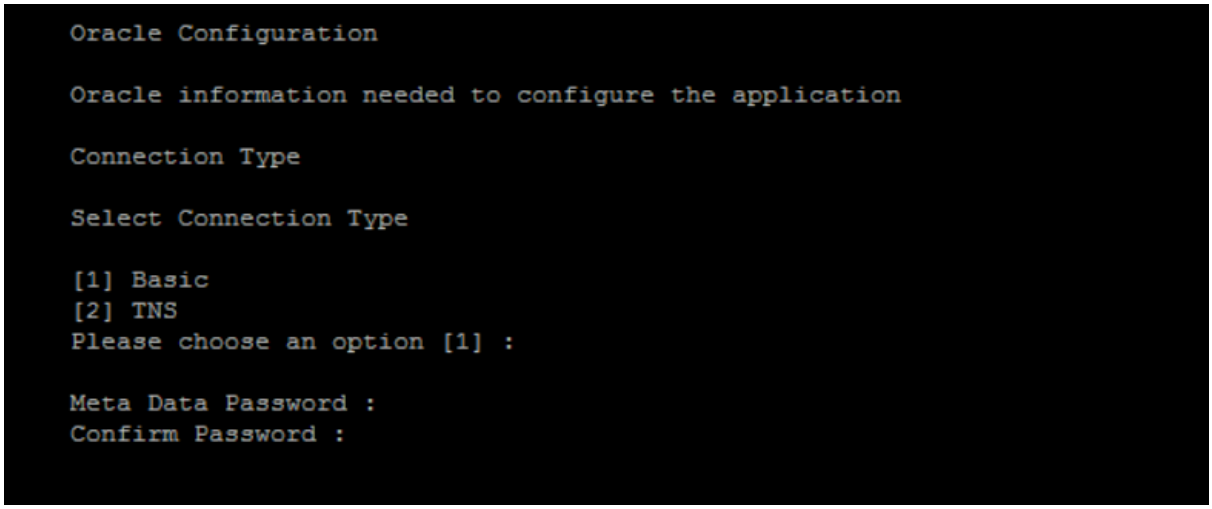
```
--------------------------------------------------------------------------------
SQL Server Configuration

Information needed to configure SQL Server

Database Name [dg]: A_dg

Controller Schema Name [dgcontroller]: A_dgcontroller

Control Schema Name [dgcontrol]: A_dgcontrol

HDFS Info Schema Name [dghdfsinfo]: A_dghdfsinfo

Dashboard Schema Name [dgstar]: A_dgstar
```

Oracle

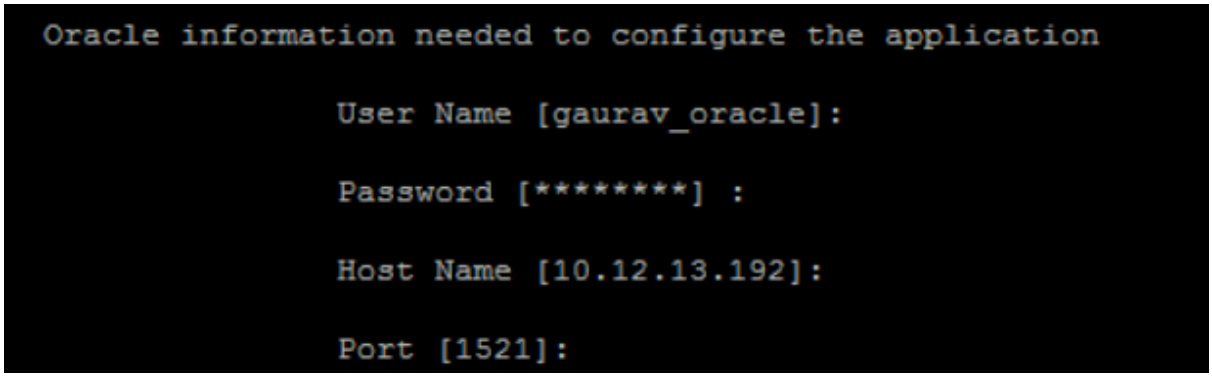
ix. To configure Oracle, enter [4].



x. Select either Basic or TNS.



xi. If you choose Basic, enter the required information which includes: user name, password, host name, and port number.



xii. If you choose TNS, enter the required information which includes: user name, password, TNS name, and TNS admin path.

xiii. Select the database option, With Fresh DB/Schemas or With Existing DB/Schemas.

```
Select database option

Information needed to configure Oracle

[1] With Fresh DB/Schemas
[2] With Existing DB/Schemas
Please choose an option [1] :

With Fresh DB/Schemas
```

xiv. Provide Oracle Configuration details.

```
Oracle Configuration

Information needed to configure Oracle

Controller Schema Name [DGCONTROLLER]: A_DGCONTROLLER

Control Schema Name [DGCONTROL]: A_DGCONTROL

HDFS Info Schema Name [DGHDFSINFO]: A_DGHDFSINFO

Dashboard Schema Name [DGSTAR]: A_DGSTAR
```

xv. Enter a Port number or use the default Tomcat port (10181). If you want to configure SSL certificate, enter SSL port as 10182.

```
Information needed to configure Tomcat

Port [10181]:
```

Note: For non-root user, enter the port number as 1024 or above.

xvi. Enter Controller ID. Press enter to continue installing using the default controller ID or provide a controller ID.

```
Controller Configuration

Information needed to configure controller

Controller ID [cdcdf48e4434e248c1e345e08013a9d2]: dataguise
```

Note: Ensure that the controller ID is same for all the IDPs.

2.4.2 Configure IDPs

The types of DgSecure IDPs with their default port number are as follows:

- Discover (Databases) – 8889
- Masker (Databases) – 8888
- Files – 8082
- Hadoop Control – 8089
- Hive – 9980
- GDPR – 1433

i. Hadoop IDP: Select the type of Distribution System, version and enter the controller ID.

```
HDFS IDP Selection

Select option to deploy HDFS IDP

[1] MapR
[2] Cloudera
[3] Hortonworks
[4] Pivotal
[5] EMR
[6] Spark
Please choose an option [1] : 2

Cloudera

[1] CDH 5.12.1: HDFS IDP compatible with CDH 5.12.1 will be deployed for use by DgSecure.
[2] CDH 5.3.1: HDFS IDP compatible with CDH 5.3.x to 5.11.x will be deployed for use by DgSecure.
[3] CDH 5.1: HDFS IDP compatible with CDH 5.1.x to 5.2.x will be deployed for use by DgSecure.
Please choose an option [1] :

--------------------------------------------------------------------------------
HDFS IDP Configuration

Information needed to configure HDFS IDP

Controller ID [cdcdf48e4434e248c1e345e08013a9d2]: dataguise
```

ii. Monitoring IDP: Select the type of Hadoop distribution for monitoring, version and enter the controller ID.

```
Monitoring IDP Selection

Select Hadoop distribution to deploy Monitoring IDP

[1] MapR
[2] Cloudera
[3] Hortonworks
[4] Pivotal
[5] EMR
Please choose an option [1] : 2

Cloudera

[1] CDH 5.3.1 Yarn: Monitoring IDP compatible with CDH 5.3.1 Yarn will be deployed for use by DgSecure.
[2] CDH 5.1 Yarn: Monitoring IDP compatible with CDH 5.1 Yarn will be deployed for use by DgSecure.
Please choose an option [1] :
```

iii. Discover IDP: Enter the controller ID.

```
Discover IDP Configuration

Information needed to configure Discover IDP

Controller ID [cdcdf48e4434e248c1e345e08013a9d2]: dataguise
```

iv. Masker IDP: Enter the controller ID.

```
--------------------------------------------------------------------------------
Masker IDP Configuration

Information needed to configure Masker IDP

Controller ID [cdcdf48e4434e248c1e345e08013a9d2]: dataguise

--------------------------------------------------------------------------------
```

v. Files IDP: Select the type of Files System and enter the controller ID.

```
--------------------------------------------------------------------------------
Files IDP Selection

Select option to deploy Files IDP

[1] Files System
Please choose an option [1] :

Files System

[1] Local File System: Files IDP compatible with Local File System will be deployed for use by DgSecure.
[2] S3 File System: Files IDP compatible with S3 File System will be deployed for use by DgSecure.
Please choose an option [1] :

--------------------------------------------------------------------------------
Files IDP Configuration

Information needed to configure Files IDP

Controller ID [cdcdf48e4434e248c1e345e08013a9d2]: dataguise
```

vi. Hive IDP: Select the type of Distribution system for Hive and enter the controller ID.

```
Hive IDP Selection

Please select distribution to configure Hive IDP

[1] MapR: Hive IDP for MapR Support
[2] Cloudera: Hive IDP for Cloudera Support
[3] Hortonworks: Hive IDP for Hortonworks Support
[4] Pivotal: Hive IDP for Pivotal Support
[5] HD Insight: Hive IDP for HD Insight Support
Please choose an option [1] : 2

--------------------------------------------------------------------------------
Hive IDP Configuration

Information needed to configure Hive IDP

Controller ID [cdcdf48e4434e248c1e345e08013a9d2]: dataguise
```

vii. GDPR IDP: Select the database type, SQL Server or oracle, enter the server configuration and controller ID.

```
Database information needed to configure the GDPR IDP

Select database

Database type selection

[1] Sql Server database: The specified Sql Server instance will be configured for use by GDPR IDP.
[2] Oracle database: The specified Oracle instance will be configured for use by GDPR IDP.
Please choose an option [1] :

--------------------------------------------------------------------------------
Sql Server Configuration

Sql Server information needed to configure GDPR IDP

               Server Name: [localhost]: 192.168.0.151

Connect Using:

               Connect Using:

[1] Instance
[2] Port
Please choose an option [1] : 2

                []: 1433

--------------------------------------------------------------------------------
Sql Server Configuration

Sql Server information needed to configure GDPR IDP

               User Name: [sa]:

               Password: :

--------------------------------------------------------------------------------
Sql Server Configuration

Sql Server information needed to configure GDPR IDP

               Database Name: [GDPR]: GDPR_TEST

               Schema Name: [dbo]:

--------------------------------------------------------------------------------
GDPR IDP Configuration

Information needed to configure GDPR IDP

Controller ID [cdcdf48e4434e248c1e345e08013a9d2]: dataguise
```

viii. DgWalker IDP: Refer to section 2.5.11.

ix. Cloud IDP: Select the cloud IDP type, i.e., S3, GCS or HD Insight, and enter the controller ID.

```
Information needed to configure HDFS IDP for Cloud

IDP Type

Please select IDP type

[1] S3 HDFS/Local IDP
[2] GCS HDFS IDP
[3] HDInsignt HDFS IDP
Please choose an option [1] : 2

Controller ID [84a5787c203467d9de0ff6b1a200d1c6]: dataguise
```

S3 IDP

x. Enter [1] to select S3 IDP and provide controller ID.

```
IDP Type

Please select IDP type

[1] S3 HDFS/Local IDP
[2] GCS HDFS IDP
[3] HDInsignt HDFS IDP
Please choose an option [1] :

Controller ID [9d802cf06b4f431e525886e0637dd28d]: dataguise
```

xi. Provide the complete location and select the file system.

```
--------------------------------------------------------------------------------
S3 Cloud Configuration for S3 IDP/Files IDP

Information needed to configure S3 IDP

S3 IDP Configuration String [--ClusterTimediffMillisecs "0" --DgMetaDir "/dataguise\$" --S3filesystem "s3" --HadoopConfigPath "/etc/hadoop/conf" --ControllerUrl "http\\://localhost\\:10181
]:

File System

Select File System:

[1] Already provisioned CDH5/EMR/HW cluster / Files IDP
[2] EMR cluster to be provisioned by DgSecure Cloud IDP
Please choose an option [1] : 2
```

xii. Provide cloud configuration.

```
--------------------------------------------------------------------------------
S3 IDP Configuration

Information needed to configure S3 IDP

Meta Name for S3 IDP [dgsecure-test]:

AWS Compute Region for S3 IDP [us-east-1]:

--------------------------------------------------------------------------------
```

GCS IDP

x. Enter [2] to select GCS IDP and provide controller ID.

```
IDP Type

Please select IDP type

[1] S3 HDFS/Local IDP
[2] GCS HDFS IDP
[3] HDInsignt HDFS IDP
Please choose an option [1] : 2

Controller ID [84a5787c203467d9de0ff6b1a200d1c6]: dataguise
```

xi. Provide the complete location.

```
Information needed to configure GCS IDP

GCS IDP Configuration String [--ClusterTimediffMillisecs 0 --DgMetaDir /dataguise\$ --HadoopConfigPath /etc
cs 0 --DgMetaDir /dataguise\$ --HadoopConfigPath /etc/hadoop/conf --ControllerUrl http\\://35.197.75.230 \\
```

xii. Provide Google Project ID, size and Bucket name.

```
Google Project ID []: wide-isotope-147019

Google Zone [us-central1-a]: us-east1-b

Bucket Name [dgsecure-1564570592]: dg-sw
```

HD Insight IDP

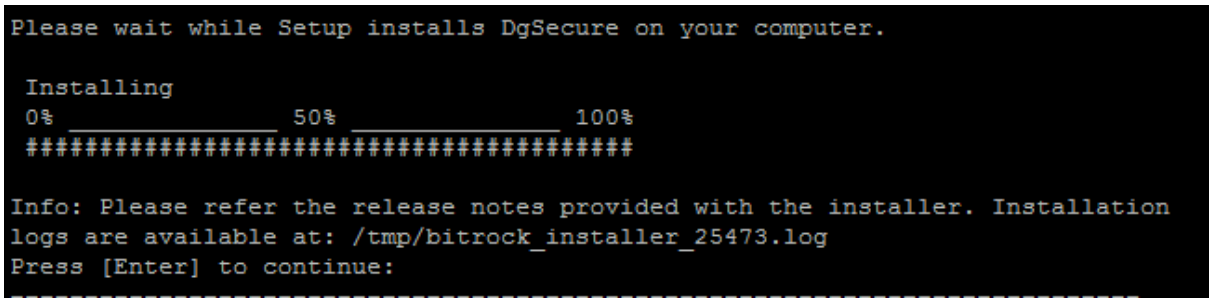
x. Enter [3] to select HDInsight IDP and provide controller ID.

```
--------------------------------------------------------------------------------
HDFS/Local IDP (S3 Only) Configuration for Cloud

Information needed to configure HDFS IDP for Cloud

IDP Type

Please select IDP type

[1] S3 HDFS/Local IDP
[2] GCS HDFS IDP
[3] HDInsignt HDFS IDP
Please choose an option [1] : 3

Controller ID [7e61e04bbf0447796808209144026d3a]: dataguise
```

xi. Provide HDInsight IDP Configuration String.

```
[--ClusterTimediffMillisecs "0" --DgMetaDir "/dataguise\$" --HadoopConfigPath "/etc/hadoop/conf" --ControllerUrl "http\\://localhost\\:10181"]:
```

xii. After the tomcat, controller and IDP setup, installation will begin.



2.5 Distributed Installation

Installation files for all the IDPs are also provided separately. This is to facilitate the installation of IDPs and Controller on a distributed system consisting of multiple machines.

The controller installer (for distributed installations) has the options to install:

- The DgSecure controller
- Postgres Database

Following installer files are provided:

IDP	**Installer File**
Hive	DgSecureHiveIDP-xxx-linux-64-installer.run
Monitoring	DgSecureMonitoringIDP-xxx-linux-64-installer.run
HDFS	DgSecureHDFSIDP-xxx-linux-64-installer.run
Command line	DgSecureCommandlineIDP-xxx-linux-64-installer.run
GDPR	DgSecureGDPRIDP-xxx-linux-64-installer.run
NoSQL	DgSecureNoSQLIDP-xxx-linux-64-installer.run
DgWalker	DgSecureDgWalkerIDP-xxx-linux-64-installer.run
Discover	DgSecureDiscoverIDP-xxx-linux-64-installer.run
Masker	DgSecureMaskerIDP-xxx-linux-64-installer.run
DgMonitor	DgSecureDgMonitotIDP-xxx-linux-64-installer.run
Files	DgSecureFilesIDP-xxx-linux-64-installer.run

Big Query	DgSecureBigQueryIDP-xxx-linux-64-installer.run
Cloud	DgSecureCloudIDP-xxx-linux-64-installer.run

2.5.1 Install HDFS IDP

1. Locate and copy the executable file DgSecureHDFSIDP-xxx-linux-installer.run to the machine on which you will install the IDP.
2. Run DgSecureHDFSIDP-xxx-linux-installer.run.
3. Accept the license agreement.
4. Install the IDP in the default directory or enter a different location.
5. Refer to Section 2.4.2, step 1 for configuration of HDFS IDP
6. Enter the Controller ID.
7. Installation will begin.

2.5.2 Install Monitoring IDP

1. Locate and copy the executable file DgSecureMonitoringIDP-xxx-linux-installer.run to the machine on which you will install the IDP.
2. Run DgSecureMonitoringIDP-xxx-linux-installer.run.
3. Accept the license agreement.
4. Install the IDP in the default directory or enter a different location.
5. Refer to Section 2.4.2, step 2 for configuration of Monitoring IDP
6. Enter the Controller ID.
7. Installation will begin.

2.5.3 Install Discover IDP

1. Locate and copy the executable file DgSecureDiscoverIDP-xxx-linux-installer.run to the machine on which you will install the IDP.
2. Run DgSecureDiscoverIDP-xxx-linux-installer.run.
3. Accept the license agreement.
4. Install the IDP in the default directory or enter a different location.
5. Refer to Section 2.4.2, step 3 for configuration of Discover IDP
6. Installation will begin.

2.5.4 Install Masker IDP

1. Locate and copy the executable file DgSecureMaskerIDP-xxx-linux-installer.run to the machine on which you will install the IDP.
2. Run DgSecureMaskerIDP-xxx-linux-installer.run.
3. Accept the license agreement.
4. Install the IDP in the default directory or enter a different location.
5. Refer to Section 2.4.2, step 4 for configuration of Masker IDP
6. Installation will begin.

2.5.5 Install Files IDP

1. Locate and copy the executable file DgSecureFilesIDP-xxx-linux-installer.run to the machine on which you will install the IDP.
2. Run DgSecureFilesIDP-xxx-linux-installer.run.
3. Accept the license agreement.
4. Install the IDP in the default directory or enter a different location.
5. Refer to Section 2.4.2, step 5 for configuration of Files IDP
6. Enter the Controller ID.
7. Installation will begin.

2.5.6 Install Hive IDP

1. Locate and copy the executable file DgSecureHiveIDP-xxx-linux-installer.run to the machine on which you will install the IDP.
2. Run DgSecureHiveIDP-xxx-linux-installer.run.
3. Accept the license agreement.
4. Install the IDP in the default directory or enter a different location.
5. Refer to Section 2.4.2, step 6 for configuration of Hive IDP
6. Enter the Controller ID.
7. Installation will begin.

2.5.7 Install DSAR IDP

Note: The user's login credentials, database and schema must be created before Privacy agent installation. To run the pre-requisite scripts for Azuresql, Sql server and Oracle access it from location: "**…/x.x.x.x.x_BUILD_DD-M-YYYY/ PrerequisiteScripts/…/…**"

The sequence in which Azuresql scripts need to trigger are:

1. login_database.sql

2. database_grant.sql
3. schema.sql

Similarly, run the pre-requisite scripts for Sqlserver and Oracle. For Oracle, run the **user_grant.sql** script. For Sql server, run **login_database_schema.sql** script.

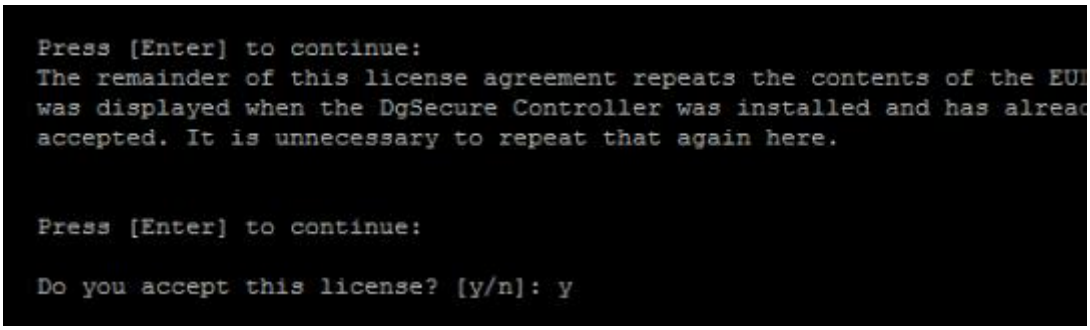
1. Locate and copy the executable file DgSecureDSARIDP-xxx-linux-installer.run to the machine on which you will install the IDP.
2. Run DgSecureDSARIDP-xxx-linux-installer.run.
3. Accept the license agreement.
4. Install the IDP in the default directory or enter a different location.
5. Refer to Section 2.4.2, step 7 for configuration of DSAR IDP
6. Enter the Controller ID.
7. Installation will begin.

2.5.8 Install Cloud IDP

1. Locate and copy the executable file DgSecureCloudIDP-xxx-linux-installer.run to the machine on which you will install the IDP.
2. Run DgSecureCloudIDP-xxx-linux-installer.run.
3. Accept the license agreement.
4. Install the IDP in the default directory or enter a different location.
5. Refer to Section 2.4.2, step ix for configuration of Cloud IDP
6. Enter the Controller ID.
7. Installation will begin.

2.5.9 Install DgMonitor

1. Double click the **DgMonitor-7.2.0.x.x.x-linux-x64-installer.run** file.
2. Accept the license agreement.



3. Provide the location to install the DgMonitor.

```
--------------------------------------------------------------------------------
Please specify the directory where DgMonitor will be installed.

Installation Directory [/opt/Dataguise]: /opt/monitor

--------------------------------------------------------------------------------
```

4. Select the components that need to be installed. Enter N for the components already available at your machine.

```
--------------------------------------------------------------------------------
Before proceeding, select the components you want to install; unselect the
components you do not want to install.

Apache Zookeeper [Y/n] :

Apache Kafka [Y/n] :

Apache Storm [Y/n] :

DgAnomalyDetection [Y/n] :

DgLogstash [y/N] :

DgMonitorSubSystem : Y (Cannot be edited)

Is the selection above correct? [Y/n]:

--------------------------------------------------------------------------------
```

5. Enter 1 to install all source systems except MapR or 2 to install MapR. This installation has been done using option 1.

```
Select DgMonitorSubSystem Support

[1] Support for All (except MapR)
[2] Support for MapR
Please choose an option [1] : 1
```

6. Enter 1 to install all Storm version 0.x or 2 to install Storm version 1.x. This installation has been done using option 1.

```
Support for All (except MapR)

[1] Storm version 0.x: DgMonitorSubSystem support for All (except MapR) and storm version 0.x will be deployed.
[2] Storm version 1.x: DgMonitorSubSystem support for All (except MapR) and storm version 1.x will be deployed.
Please choose an option [1] : 1

--------------------------------------------------------------------------------
```

7. Provide Host details, Eagle port number, Service user ID and create a password for the service.

```
DgMonitorSubSystem Environment Configuration

Service Host * [localhost]:

Eagle Service Port * [9090]:

Service User * [admin]:

Service Password * [********] :
Confirm Password [********] :
```

8. Provide user information

```
DgMonitorSubSystem Environment Configuration

First Name * [Admin]:

Last Name * [Admin]:

Email * [mock-admin@dataguise.com]:
```

9. Provide Apache Storm host name and port number

```
DgMonitorSubSystem Environment Configuration

Nimbus Host * [localhost]:

Nimbus Port * [6627]:
```

10. Provide Apache Zookeeper hostname and client port number

```
DgMonitorSubSystem Service Configuration

Transaction Zookeeper Server * [localhost]:

Zookeeper Client Port * [2181]:
```

11. To configure email notifications, configure the SMTP server

```
DgMonitorSubSystem Email Notification Configuration

SMTP Username * []: sds

SMTP Password * :

SMTP Server * [pod51009.outlook.com]:

SMTP Port * [587]:

SMTP Connection * [tls]:


SMTP Authentication [Y/n]:

The specified port 587 lies in the restricted port ranges. Do you wish to continue the installation at this port? [Y/n]:
```

12. Select the required database configuration for DgMonitor Subsystem.MySQL:

```
DgMonitorSubSystem Configuration MySQL

MySQL Host * [localhost]:

MySQL Port * [3306]:

Storage Username * [eagle]: root

Storage Password * :

Storage Database * [dgmonitor]:
```

- PostgreSQL:

```
DgMonitorSubSystem Configuration PostgreSQL

PostgreSQL Host * [localhost]:

PostgreSQL Port * [5432]:

Storage Username * [postgres]:

Storage Password * :

Storage Database * [dgmonitor]:
```

13. Press enter to confirm the configuration.

```
DgMonitorSubSystem Configuration

You are about to install DgMonitor.


Please review the below information:
   Storage Connection URL:    jdbc:mysql://localhost:3306/dgmonitor
   StorageDriverClass:    com.mysql.jdbc.Driver
```

14. Enter Y to begin installation.

```
Setup is now ready to begin installing DgSecure DgMonitor on your computer.

Do you want to continue? [Y/n]: y
```

15. Exit setup when installation is complete.

```
Please wait while Setup installs DgSecure DgMonitor on your computer.

 Installing
 0%              50%              100%
```

2.5.10 Install NoSQL IDP

1. Locate and copy the executable file DgSecureNoSQLIDP-xxx-linux-installer.run to the machine on which you will install the IDP.
2. Run DgSecureNoSQLIDP-xxx-linux-installer.run.
3. Accept the license agreement.

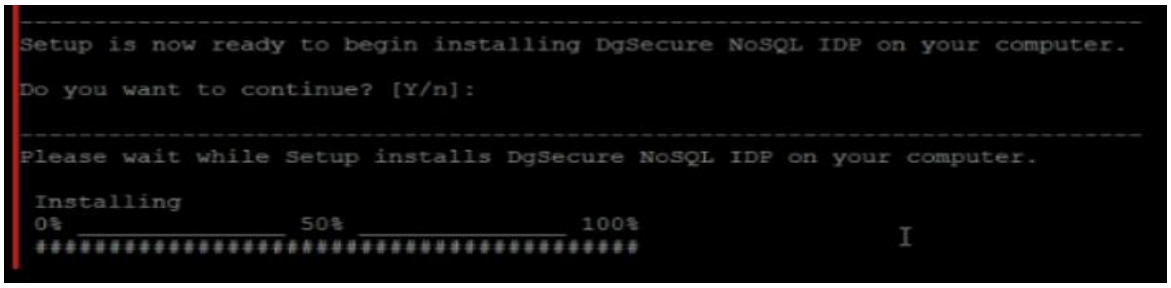
```
Do you accept this license? [y/n]: y
```

4. Install the IDP in the default directory or enter a different location.

```
Please specify the directory where NoSQL IDP will be installed.

Installation Directory [/opt/Dataguise]:
```

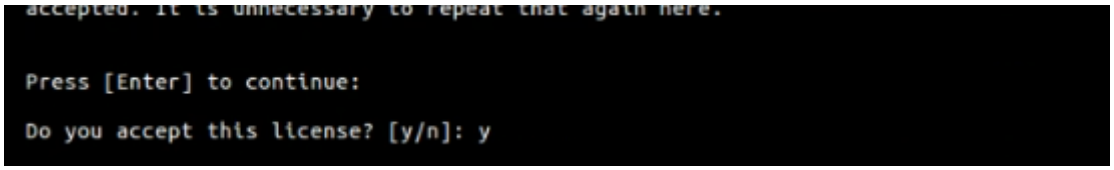
5. To configure Spark Master, enter its URL.

6. Enter Y to begin installation.

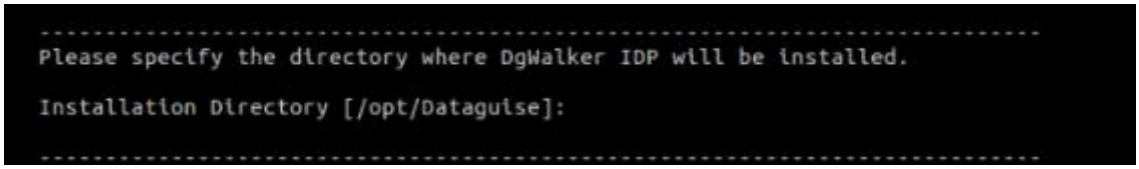


2.5.11 Install DgWalker IDP

1. Locate and copy the executable file DgSecureDgWalkerIDP-xxx-linux-installer.run to the machine on which you will install the IDP.
2. Run DgSecureDgWalkerIDP-xxx-linux-installer.run.
3. Accept the license agreement.



4. Install the IDP in the default directory or enter a different location.



5. Select the Installation type, Default or Advanced.

```
Installation type selection

[1] Default: Installation Type: Default.
[2] Advanced: Installation Type: Advanced.
Please choose an option [1] :


Installation type selection

[1] Default: Installation Type: Default.
[2] Advanced: Installation Type: Advanced.
Please choose an option [1] :
```

Note: If you select [1], then the database selection screen will be skipped and the default database will be set to In-Memory for DgWalker installation.

6. Select the database if you have selected [2] in step 5.

```
--------------------------------------------------------------------------------
Database type selection for DgWalker IDP

Database information needed to configure the application

Select database

Database type selection

[1] PostgreSQL database: The specified PostgreSQL instance will be configured for use by DgWalker IDP.
[2] MySQL database: The specified MySQL instance will be configured for use by DgWalker IDP.
[3] SQL Server database: The specified SQL Server instance will be configured for use by DgWalker IDP.
[4] In-Memory database: The specified in-memory instance will be configured for use by DgWalker IDP.
Please choose an option [4] :

--------------------------------------------------------------------------------
```

Note: Ensure that PostgreSQL, MySQL, or SQL Server is installed on the machine where you install DgWalker IDP.

7. Refer to section 2.4, step vii-PostgreSQL (existing database) or MySQL to configure the database.
8. Select the source system for DgWalker deployment.

```
--------------------------------------------------------------------------------
DgWalker IDP Selection

Select distribution option to deploy DgWalker IDP

[1] Hadoop
[2] Cloud (AWS)
Please choose an option [1] : 2
```

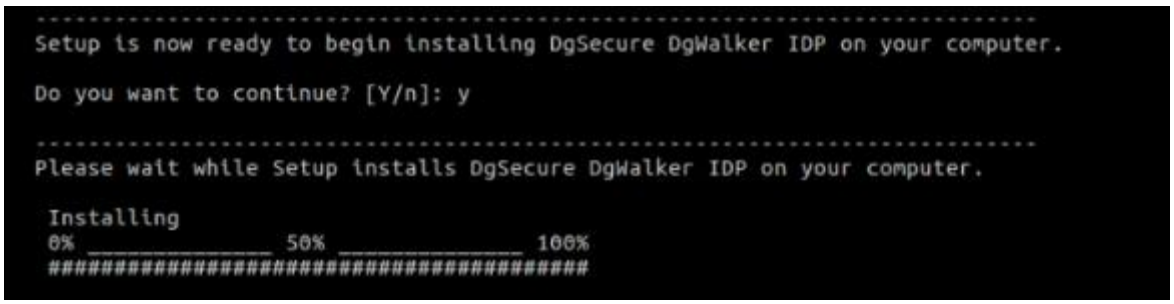
9. Select the distribution system if you have selected [1] in step 8, and provide configuration details.

10. Provide S3 bucket details if you have selected [2] in step 8.

11. Enter the number of threads, this is set to 32 by default.

12. Enter the Controller ID.

13. Enter Y to begin installation.



3. Configuration: DgSecure Admin

3.1 Setup the Admin Console

The DgSecure Admin console provides tools for monitoring DgSecure operations and managing licenses, users, and agents. When you launch it for the first time, you are guided through a one-time process in which you install the product license key and create a DgSecure super user and password.

You will need these items which you received when you purchased DgSecure:

- License file
- Challenge password file

➢ **To launch the DgSecure Admin console**

i. On the machine where you installed the controller, open the Services Console and restart the Apache Tomcat Service.

ii. Copy your Dataguise license (.lic) file to the machine hosting the controller.

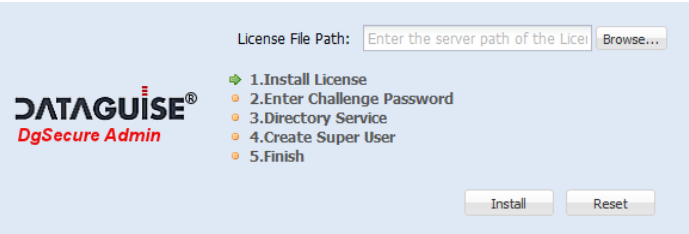
iii. On the same machine, open a browser window and enter this url:

http:\\localhost:<*port*>\dgadmin

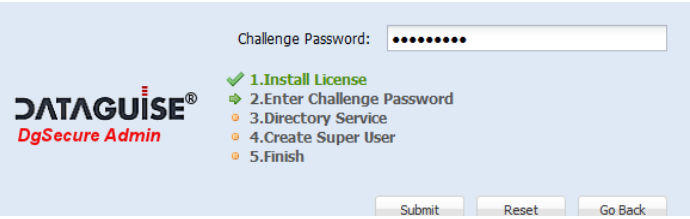
where <*port*> is the port number you used to install DgSecure.

Note: If you selected SSL as an option during installation, you must prefix the host URL with https:\\, and the default port number 10182.

iv. The Install License dialog box is displayed. Enter the path to the license file, or click **Browse** to locate it and click **Install**.



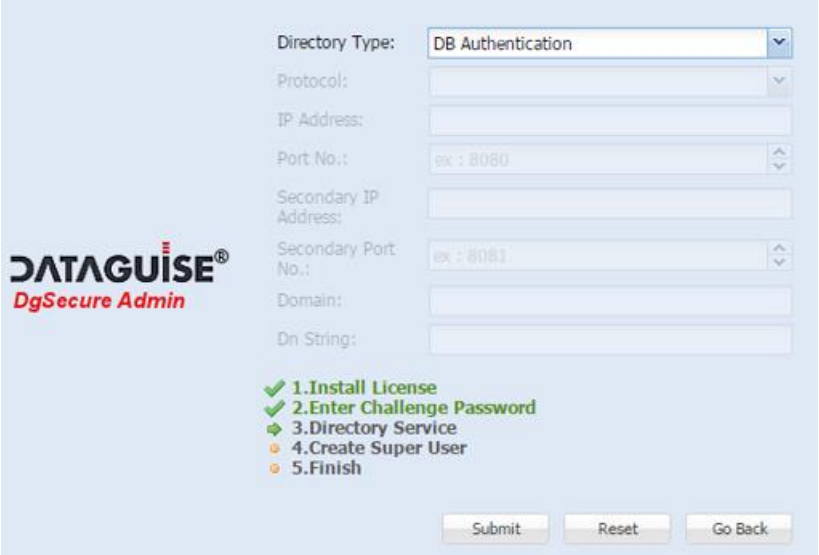
v. The Challenge Password dialog box is displayed. Enter the challenge password you received with the license file, click **Submit**.



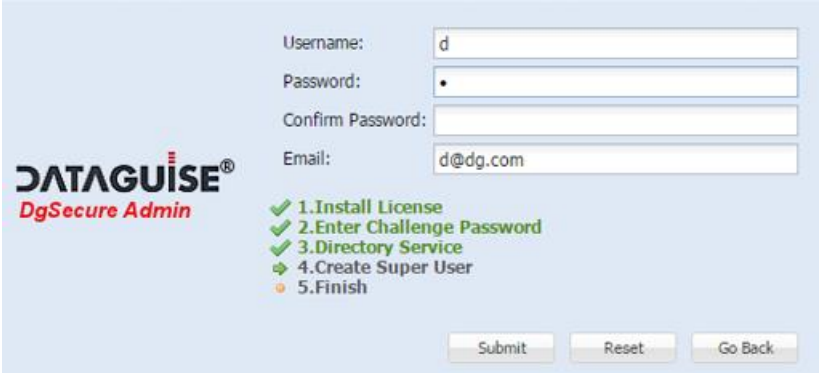
vi. Steps "3" and "4" depend upon the directory type selected for authentication.

a. DB Authentication

i. Step 3

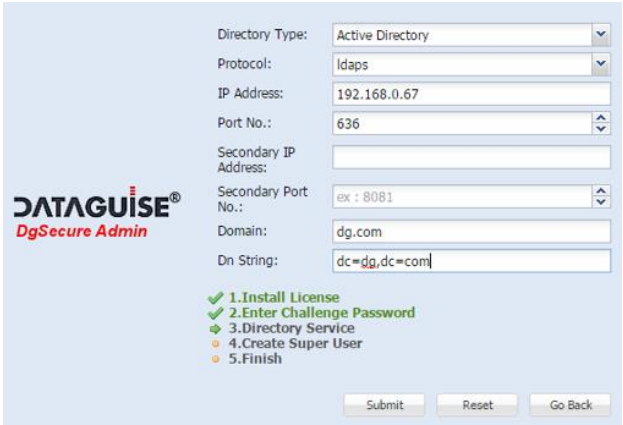


ii. Step 4

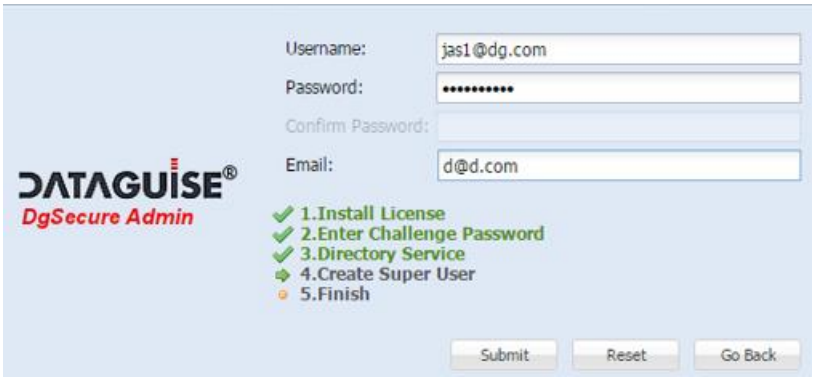


b. Active Directory

i. Step 3

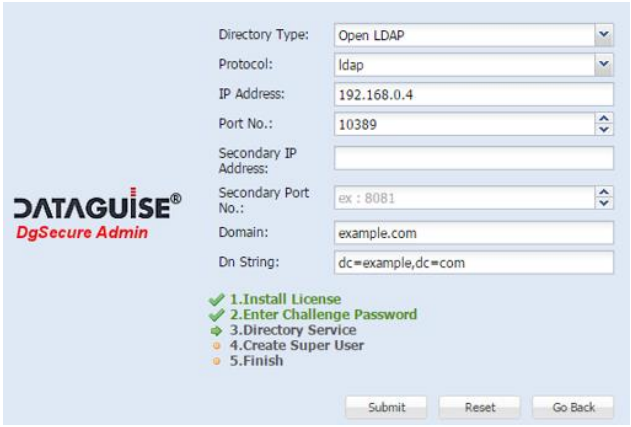


ii. Step 4



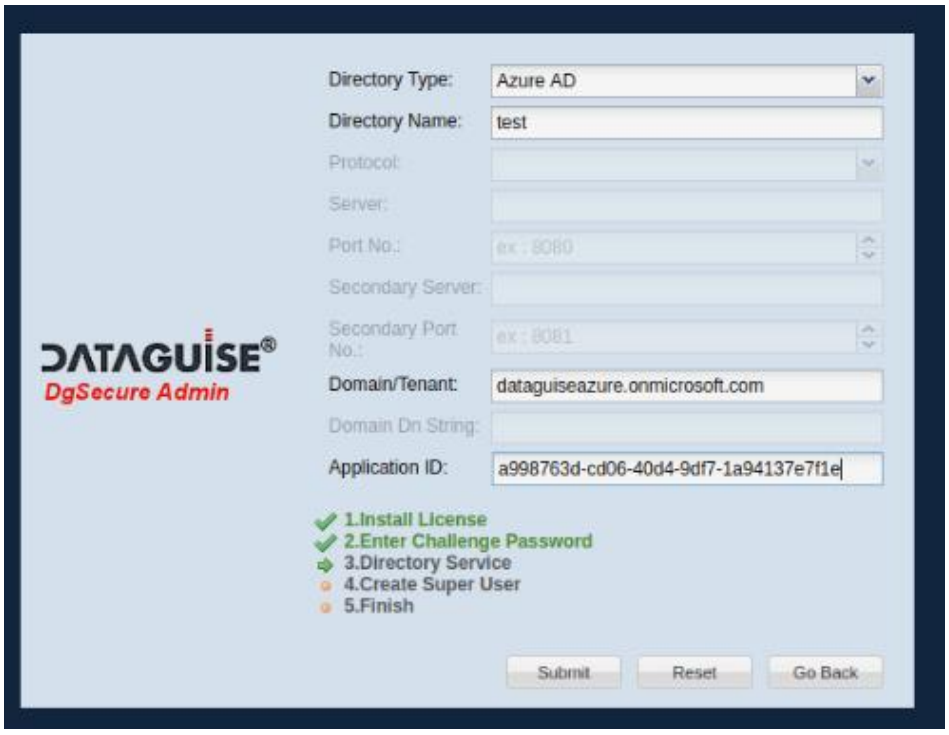
c. Open LDAP

i. Step 3



d. Azure Active Directory

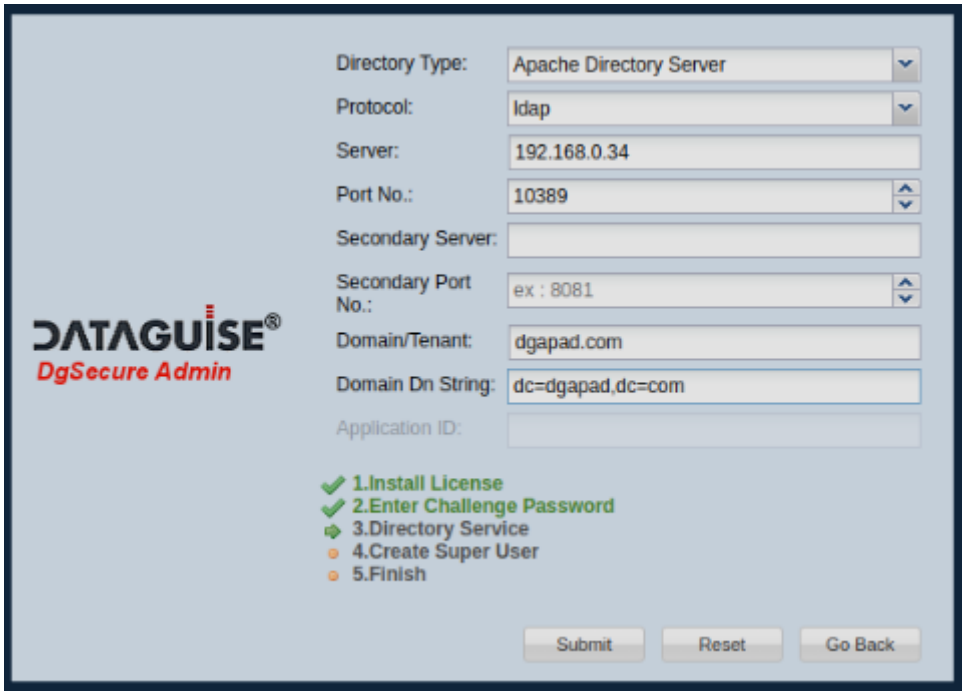
i. Step 3



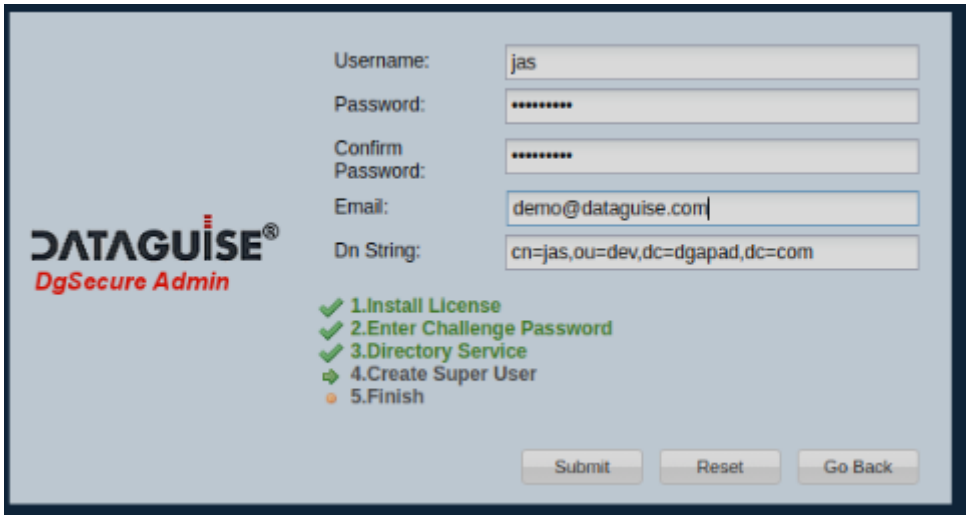
ii. Step 4

e. Apache Directory Server

i. Step 3



ii. Step 4



Enter a **Username** and **Password** for the DgSecure super user. The super user must be an authorized user in the service, and the **Username** and **Password** you enter here must match the user name and password listed in the directory service.

Enter a DN String, a concatenated string of the superuser username and the domain:

uid=<*superuser*>,dc=<*domain*>,dc=<*tld*>

where

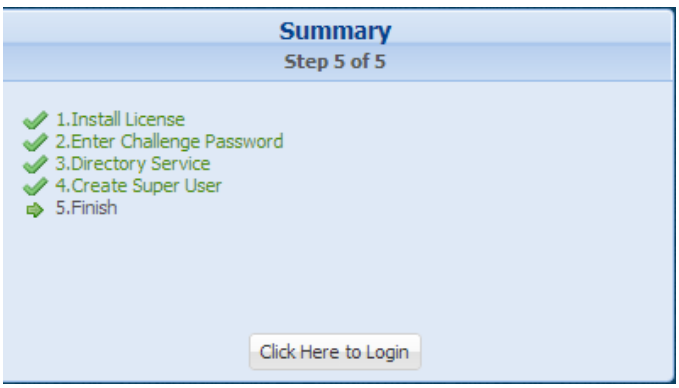
<*superuser*> The user id of the superuser.

<*domain*> The second level domain.

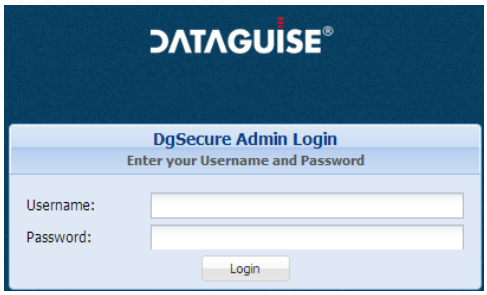
<*tld*> The top-level domain.

Click **Submit** when this is completed.

1. The Summary screen is displayed. Click here to Login.



2. The login screen for DgSecure Admin is displayed. Enter the login ID and password for the super user you created.

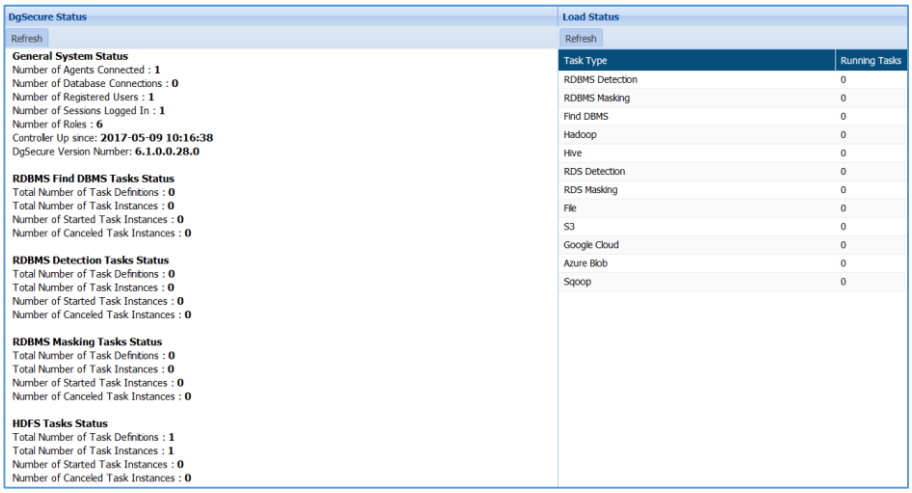


3. The DgSecure Admin console is displayed.

3.2 DgAdmin

The DgSecure Admin console provides tools for monitoring DgSecure operations and managing licenses, users, and IDPs. Only those with superuser access, usually only the DgSecure Administrator, has access to DgAdmin.

The DgAdmin landing page is **DgSecure Status Page**. This page offers a high-level overview of DgSecure's tasks (HDFS, Files, Find-it, Search-it, and Masker) and general system metrics. Below is an example of what the page looks like. The same four metrics are shown for each available task type.



From the landing page, use the left-hand menu to navigate the application. This chapter is divided into three sections. Section 4.1 covers administering users including managing role, preferences, and notifications. Section 4.2 cover administering the IDPs. Finally, section 4.3 covers miscellaneous functions.

3.2.1 Administer Users

DgSecure users must be authorized users on the directory service specified in the Directory Service Management page (for more information see section 4.1.8).

To authenticate users, the DgSecure host machine must have access to a directory service. This could be your company's enterprise directory service, or it could be a dedicated directory service created specifically for DgSecure users. DgSecure users log into the application using the credentials specified in the directory server, and the DgSecure controller connects to the server and verifies that the user credentials are valid. After this initial authentication, DgSecure applies any restrictions associated with the user's role.

3.2.2 Add users

➢ **To add DgSecure users manually**

1. In DgAdmin, choose User Management > Users. The **User Management** page is displayed.

2. Click **Add User**. A new row in the **User Management** page is displayed. Fields with sample values appear under the column headings.
    a. Enter the **User Name** as it is displayed in the directory service.
    b. Enter the user's **First Name** and **Last Name**.
    c. Enter the user's company **Email ID**, including the domain. For example: John.Doe@company.com
    d. Enter any notes you wish in the **Other Info** field. For example, you might want to note the department or business unit where the user works.
    e. Select a **Role** from the drop-down menu. For more information about roles, or to add a role, see section 4.1.5
3. Click **Update** to save the information to the table.

3.2.3 Inactivate users

You can suspend a user's access without removing him or her from the list of authorized users.

➢ **To inactivate a user**

1. In DgAdmin, choose User Management > Users. The **User Management** page is displayed.
2. Select the user you want to inactivate.
3. Click inside the cell in the **Active** column to reveal a drop-down menu.
4. Choose **No** from the drop-down menu.

The user's access to the product is now suspended. Should the user try to log in to the product, he or she will get an error message similar to the one that appears when an incorrect login or password is submitted.

3.2.4 Export users

➢ **To export users to a CSV file**

1. In DgAdmin, choose User Management > Users. The **User Management** page is displayed.
2. Under User Export/Import, click **Export to CSV File**. The File Download dialog box appears.
3. Save the file to your computer.

3.2.5 Import users

You can import user information from a CSV file, or from LDAP.

CSV File Import

For CSV file import, the file should follow the format and content guidelines described below. Users uploaded from a CSV file are added to any users already in the table. No existing users are overwritten.

The CSV file must have a header line with field names that correspond to those in the User Management table. The fields must meet these criteria:

- User names are case-insensitive.
- If you do not enter permissions, the default value for Role is DEFAULT_USER, and the default value for Active is No.
- Field order is not required, but the file must contain all column headers described below.

Column Name	Required?	Notes
User ID	Yes	The values in this column must match values in the directory service that DgSecure uses for authentication purposes.
Active	No	Either Yes or No. If the field is left blank or contains an invalid value, DgSecure defaults to No.
Role	No	The value must be a valid role—i.e., either a DgSecure predefined role or a custom role. The predefined roles are: SUPER_ADMIN DEFAULT_USER CONNECTION_ADMIN TASK_DESIGNER

Column Name	Required?	Notes
		TASK_EXEC UTOR ANALYST **NOTE: If the field is left blank or contains an invalid value, DgSecure defaults to DEFAULT_USER.**
Email ID	**No**	**You can leave this field blank.**
First Name	**No**	**You can leave this field blank.**
Last Name	**No**	**You can leave this field blank.**
Other Info	**No**	**You can leave this field blank.**

➢ **To import users from a CSV file**

1. In DgAdmin, choose User Management > Users. The User Management panel is displayed.
2. In the User Export/Import panel, enter the path to the CSV file, or click **Browse** to navigate to it.
2. Click **Import**. Once the users have been uploaded, a confirmation message appears. If any errors occurred during the process, another pop-up message lists the errors and gives the CSV line number where each occurred.
3. Click **OK** to dismiss the pop-up message.
4. In the user table at the top of the screen, click **Refresh**. The newly-imported users are displayed in the table.

LDAP Import

You can import groups from LDAP, and all users in an imported group will be imported into DgSecure. Roles can be assigned to groups, and all users in the group so assigned will get the roles.

3.2.6 Manage Roles and Permissions

Create and edit user roles that control access to connections, clusters, tasks, and products. Use roles to define permissions for DgSecure users. To modify a role's permissions, click on it and navigate to the permission panel you'd like to alter. There are several categories of permissions, each with its own panel: DBMS, File Connections, Custom Function, Task Definitions, Masker Template, Product Access and Cluster Access.

When you define a role, you can grant permissions that apply to specific objects, such as a particular database connection or masking task. You can also grant permissions that apply to groups of objects.

Owner Access permissions apply to any object that a user creates and "owns". Owner Access permissions are usually fully enabled, allowing a user to Create, Read, Update, Delete, or Execute anything that he or she has created.

Full Access permissions apply either globally or to large categories of objects. For example, Full Access permissions in the Connection Permissions panel apply to every database connection that has been or will ever be defined. Full Access permissions in the Task Permissions panel are subdivided into categories: Full Access for Search-It, Full Access for Find-It, Full Access for Masker, and Full Access for Search Files.

3.2.6.1 Predefined roles

DgSecure provides you with a set of default roles. You can also create additional roles as required.

Default Roles	Description
SUPER_ADMIN	The SUPER_ADMIN has the ability to access the DgSecure Admin UI and perform all operations allowed within DgSecure Admin. At least one SUPER_ADMIN is required. Multiple users can be assigned the SUPER_ADMIN role.
DEFAULT_USER	The Default User role is automatically assigned to each new user, unless you explicitly assign a different role. It has the following permissions: Access to all products. Owner Access for connections. The user has Create, Read, Update, and Delete permissions on the connections that he or she has defined. Owner Access for tasks. The user has Create, Read, Update, Delete, and Execute permissions on the Search-It, Find-It, Search Files, and Masker tasks that he or she has defined. This is a minimum set of permissions that gives users full control over the connections and tasks that they themselves create while limiting their impact on other users. For security reasons, we recommend that you keep the DEFAULT_USER permissions as minimal as possible, and that you develop new roles to capture more extensive permissions.

C ONNECTION_ADMIN	The Connection Administrator role has full CRUD control over all connections in the system. The Connection Administrator is typically someone who is either a DBA or has detailed knowledge of and access to the various databases of interest in the organization. A person in this role is responsible for maintaining the connections that will be used in different tasks.
TASK_DESIGNER	The Task Designer creates tasks for locating, searching, and masking data stores. This person has read access to the connections created by the Connection Administrator and CRUD access to all task definitions. The Task Designer also has execute permissions on all tasks, making the Task Designer capable of performing the Task Executor role. The expectation, however, is that the Task Designer will execute the tasks only to get the task definitions stabilized and ready for production. After that, the Task Designer informs the Task Executor and the Analyst that the task is ready to be run.
TASK_EXECUTOR	The Task Executor has read access to connections and task definitions and execute permission on all task definitions. The Task Executor does not have CRUD permissions on either connections or tasks.
ANALYST	The Analyst has read permissions on task definitions and connections. This gives the Analyst read permissions on Task Results. The Analyst can view and analyze the results of various runs but cannot modify the definitions or connections.

3.2.6.2 Create roles

- **To create a role**

1. In DgAdmin, choose User Management > Roles. The **Role Management** page is displayed.
2. Click **Add Role**. The Add Role dialog box is displayed.
3. Enter a Role Name and brief Role Description, the number of users if known, and activate the role if necessary.
4. Specify the permissions associated with the role. There are several categories of permissions, each with its own panel: DBMS Connections, File Connections, Custom Function, Task/Template Definitions, Product Access and Cluster Access.

Permissions can be global (e.g. all files connections), ownership based (e.g. any files connection that user creates), or specific to an object. Certain roles may have ownership and object specific permissions.

Once a level of permission is granted, mark the 'CRUD' permissions.

When you're finished, in the top left corner of the edited panel, click **Save**.

3.2.6.3 Edit role permissions

➢ **To edit a role's permissions**

1. Select the role to display its permissions in the panels below.
2. Scroll to a permissions panel and make any changes you want.
3. When you are finished, in the upper left corner of the panel click **Save**.
4. Repeat the above in other permissions panels, as necessary.

3.2.6.4 Deactivate a role

This procedure removes a role from the Role drop-down menu in the User Management screen, so that it can no longer be assigned to users.

Before inactivating a role, check the Users column to see if it is currently in use. If there are users assigned to the role, manually switch them to a new role in the User Management screen before proceeding.

➢ **To deactivate a role**

1. Select the role you want to deactivate.
2. Click inside the cell in the **Active** column to reveal a drop-down menu.
3. Choose **No** from the drop-down menu.

The role is now deactivated and will no longer appear in the Role drop-down menu in the User Management. If at some point you want to reactivate it, simply change the value in the Active column to Yes.

3.2.6.5 Find roles and permissions

All the permissions panels, except for the Edit Product Access Permissions panel, have a search tool to help you quickly locate the item you want.

Type a word in the Search box. The tool begins searching the list with your first keystroke. If it finds any matches, it marks them in yellow highlighter.

There are two search options you can select:

- Click **Regular Expression** to search using a regular expression rather than a key word or phrase. (Enter the regular expression in the search box, just as you would a key word.)
- Click **Case Sensitive** to limit search results to words that exactly duplicate the letter case of the word or phrase you enter in the search

box. For example, a case sensitive search on the key word "OracleDB1" will locate only OracleDB1 and not oracledb1, oracleDB1, or Oracledb1.

3.2.7 Settings

The **Settings** page in DgAdmin is accessed from the menu on the left side of the screen. It allows users to automatically export the results of a HDFS task instance, set the session timeout interval, and manage the controller IP service for DgSecure's high availability functionality. For the HDFS Results export, users can specify which results should be exported as well as the desired destination directory.

3.2.7.1 Export Hadoop Results

To set the Hadoop Results Export, select **Export Hadoop Results** and click **Edit**.

Set destination path

1. Click the **Enable Results Export** checkbox.
2. Enter the appropriate destination path. An error message ("Destination path does not exist) appears if a non-existent pathway is entered. Note that the destination folder must be on the same machine as the Controller.

Select columns

Select the columns which should be automatically exported. Choose from the following options:

- Task Instance ID
- Task Name
- File Path
- Regex Group (Sensitive Data Group)
- Regex Count (Hit Count)
- Masking
- Encryption
- Incremental

To select all the columns, Click the checkbox next to **Column Name**.

3.2.7.2 Session Timeout Interval

Session timeout is automatically set to 15 minutes. Time is set in minutes. To change the session timeout interval, select **Session Timeout** and click **Edit.**

3.2.7.3 Controller Service IP

This is the virtual IP used for DgSecure high availability. For more information on this functionality, please see section 3.6.

3.2.7.4 Enable Remediation Workflow

There are two settings under this option which allow you to enable and queue the task for remediation workflow.

1. **Enable Remediation Workflow**: This setting allows you to enable the remediation workflow for all the tasks. Click **YES** to enable the remediation workflow and **NO** to disable it.
2. **Auto Queue For Remediation**: This setting allows you to enable the remediation workflow for all the tasks i.e. once the task is executed, it will automatically executed for remediation workflow as well. Click **YES** to enable the auto queuing and **NO** to disable it.

3.2.7.5 Expire DGCL Session

Controls automatic logout for DGCL. When set to "Yes", DGCL logs the user out after a certain period of inactivity (a period of time controlled under the session timeout preference). When set to "No", a user remains logged into DGCL until logging himself out.

3.2.7.6 Detection Batch Size

Default batch size for auto discovery tasks is set at 30. To change the default batch size, click **Auto Discovery Batch Size** and click **Edit**.

3.2.7.7 Result Link in Notification

If a user has subscribed to task completion notifications and this preference is turned on, the email notification includes a direct to view task instance results.

3.2.7.8 DgSecure Monitor Alerts

This setting enables DgSecure Monitor to retrieve alerts from Eagle and determines the interval at which it retrieves alerts. When the setting is disabled, DgSecure Monitor does not retrieve alerts from Eagle.

3.2.7.9 LDAP Synchronization

Set the frequency (in hours) at which DgSecure's access control list synchronizes with LDAP. Optionally, click "Ldap Manual Sync" button to synch DgSecure's ACL with LDAP manually.

3.2.7.10 Default Masking Output Directory

Default masking output directory. This directory is used when no domain or alternative directory is set.

3.2.7.11 ControllerId

This setting shows the controller ID for the DgSecure controller environment. The ID controller is used to ensure only permitted DgSecure controllers can connect to certain

cloud IDPs when more than one controller is being used. In order to allow this controller ID to connect with the cloud IDP, add the controller ID to the IDP's white list located in the config.properties file (see chapter 3.4 or 3.5). The controller ID must be between 8 and 40 characters long and consist only of alphanumeric characters.

3.2.7.12 Ingest Task Status Poll Interval

Set the frequency (in seconds) at which ingest task statuses are updated.

3.2.7.13 Export JVM Utilization Report

Enables the JVM utilization report for the controller. When the report is enabled, users can set the report's destination path, update frequency (in minutes), and stop time (in hours).

3.2.7.14 Display Field Name in Structured Summary Results

This setting allows you to display the name of the field in the **Structured Summary Result** screen. The default setting value is **Field Number.** Set this property as **Yes** to display the field name in the screen.

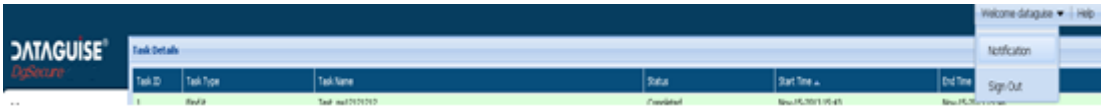
3.2.8 Get Email Notices of Events

You can setup emails to notify selected recipients about specific events.

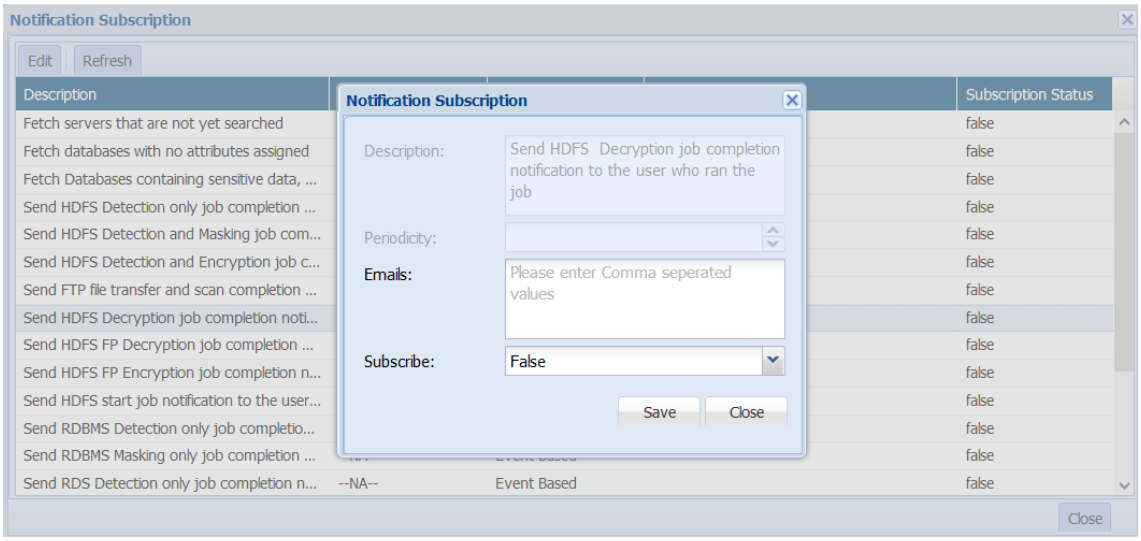
Use the Event-Based Notifications panel to alert the originating user via email when certain default discovery, masking, and encryption, file transfer and scan tasks are finished.

➢ **Set up email event notification**

3. Go to DgAdmin, choose Notifications. The Notification Management panel is displayed.
4. Select either a time-based event or an event-based event that you want to notify a user about. Click **SMTP Server Configuration**. The **SMTP Server Configuration** dialog box is displayed.
5. Enter an email user name, password, SMTP host name, and port number. Select a connection type. Click **Save**.
6. Click **Add Roles and More Details.** Choose a role and shareability (user or full). This allows users with this role to subscribe to the notification.
7. If the user would like to subscribe to this notification, they must sign into DgSecure and click on Welcome <*username*> dropdown in the top right corner of the screen. Select **Notification**.



8. For notification subscriptions, select the notification and click on edit button. Then, the Notification Subscription dialog box will appear.



5. In the Notification Subscription dialog box, enter the required details and change subscribe status to **True** to activate the notification. If the user does not want to receive notification emails, then change subscribe status to **False**.
6. Modify the fields displayed and click **Save**.

*Note: Email notifications can be setup for all tasks related to HDFS, Hive, DBMS, Azure, GCS, and AWS S3.

3.2.9 Authenticate users

To authenticate users, the DgSecure host machine must have access to a directory service. This could be your company's enterprise directory service, or it could be a dedicated directory service created specifically for DgSecure users. DgSecure users log into the application using the credentials specified in the directory server, and the DgSecure controller connects to the server and verifies that the user credentials are valid. After this initial authentication, DgSecure applies any restrictions associated with the user's role.

Before switching to a new directory server, you should confirm that all DgSecure users listed on the DgAdmin User Management screen are also listed in the directory server.

➢ **To modify the directory service**

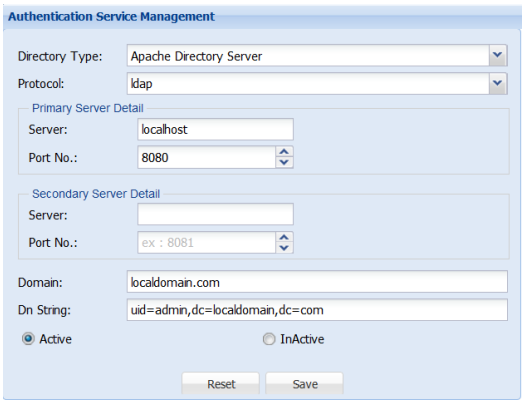
1. In DgAdmin, choose Authentication. The DgSecure Directory Service panel is displayed.
2. Enter the new information directly into the fields described below.

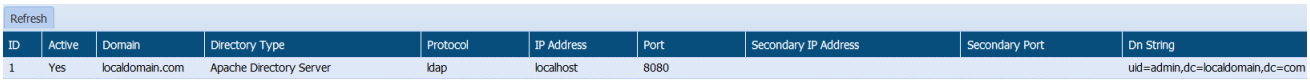
Directory type	Protocol	IP address	Port No.	Domain
Active Directory	Ldap or ldaps	ldap://<*server_ip_address*> **For example:** ldap://162.148.3.4	<port_number> The port number of the server you use to access the directory service.	<domain>.<tld> The domain name of the server hosting the directory service.
OpenDS (Sun Java)	Dn	DN://<*string*> For example, dn=testuser,dc=mydomain,dc=com; dn=testuser,dc=mydomain,dc=com		

7. Primary Server Details (IP address & port number). Required.
8. Secondary Server Details (IP address & port number). Specifies the back-up server in case of fail-over.
9. DN String, for example: dn=testuser, dc=mydomain,dc=com
10. Click **Active** or **Inactive**. (At least one authentication service must be active)
11. Click **Save**.

A pop-up message confirms that the information was saved to the repository.

When Active Directory is the directory service, DgSecure supports the use of multiple domains. One domain is set when DgSecure is initially configured. In order to use multiple domains, navigate to the Authentication page in DgAdmin, after initial configuration, and enter the details of the domain to be added.





3.2.10 LDAP Object Class Management

The LDAP Object Class Management page is used to properly configure the LDAP Object Viewer on DgSecure's ACL Management page. The LDAP Object Viewer is used to add users to DgSecure's access control list (ACL), which determines a user's decryption rights. This page contains 4 pre-defined object classes per cluster, when a cluster is configured with LDAP or LDAPS.

When a Hadoop cluster is configured with LDAP or LDAPS, four pre-defined object classes appear on this page. These four pre-defined object classes appear for every Hadoop cluster configured with LDAP or LDAPS. The cluster to which a particular object class belongs is listed in the "Cluster" column on the page. The values of each of the 4 pre-defined object classes are editable.

S. No	Object Class	Values
1	group	Search DN: $BASEDN Filter: member=(&(memberOf=$DN)(objectClass=user)) Search Scope: SUBTREE_LEVEL Collection: Yes
2	container	Search DN: $DN Filter: (objectClass=*) Search Scope: ONE_LEVEL Collection: Yes
3	organizationalUnit	Search DN: $DN Filter: (objectClass=*) Search Scope: ONE_LEVEL Collection: Yes
4	default (any object class for which specific properties have not been defined above. This will not be configurable).	Search DN:$DN Filter: (objectClass=*) Search Scope: ONE_LEVEL Collection: Yes

Pre-defined object classes can be edited, but not deleted.

Add an object class:

1. Click **Add Object Class**
2. Enter an object class name (e.g. top)
3. Enter a search Dn (e.g. objectClass=*) (e.g. ou=dev,dc=dg,dc=com)

4. Enter filter (e.g. objectClass=*)
5. Select scope

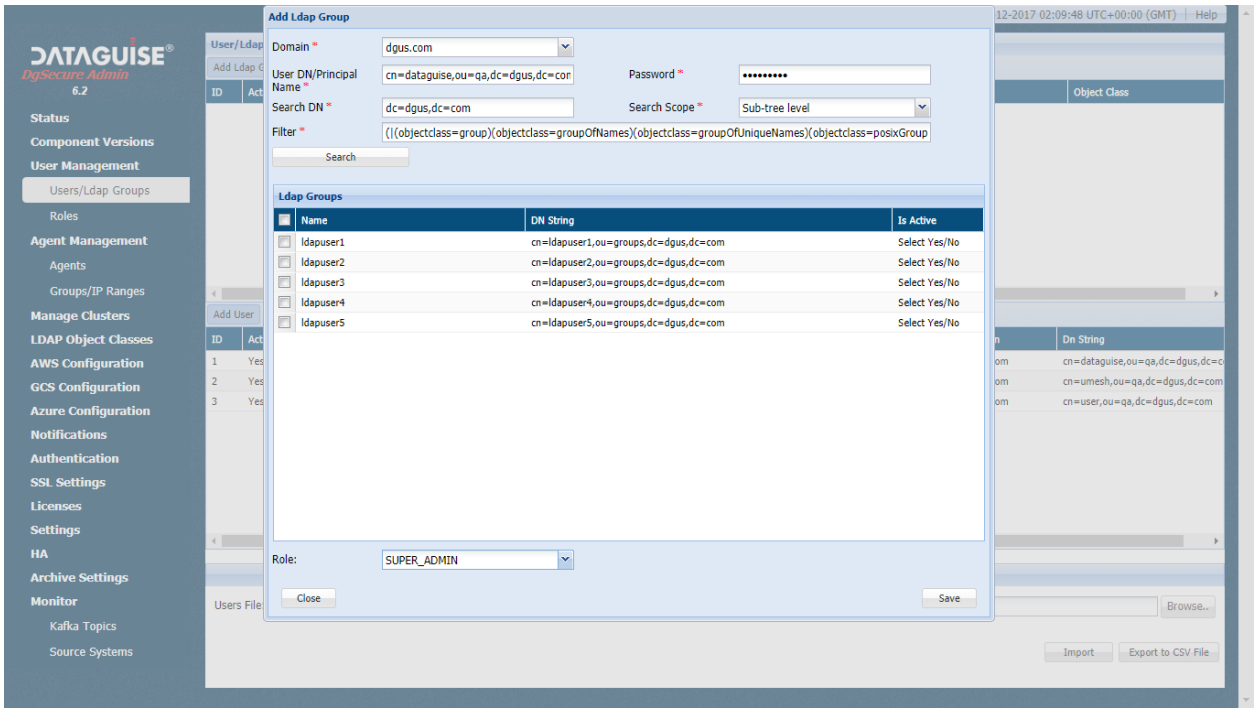
   One-level scope - This value is used to indicate searching all entries one level under the base DN - but not including the base DN and not including any entries under that one level under the base DN.
   Sub-tree scope - This value is used to indicate searching of all entries at all levels under and including the specified base DN.
   Object-level scope - This value is used to indicate searching only the entry at the base DN, resulting in only that entry being returned.

6. Select "Yes" or "No" for collection. Currently, this value is disabled.
7. Apply to a specific Hadoop cluster.
8. Click **Save**.

3.2.11 Import Users/Groups from LDAP/AD

DgSecure offers role and group management based on the LDAP group that specific users belong to. This allows the Super Admin to assign a Role (defined in section 4.1.5) to groups within the company's LDAP or AD.

Under "User Management", click on "Users/Ldap Groups" (this will be "Users" if DB Authentication was used at the time of installation). The below screen will appear when you click on "Add Ldap Group". Here, the user can search for the LDAP Groups by providing the details required and then assigning the LDAP Group to the role (at the bottom of this pop-up).



The group is then successfully imported to DgSecure. Now the users that are part of the group will be able to login to the DgSecure application with the role assigned to that group. In case a user is part of multiple groups in LDAP, and those groups are assigned different roles in DgSecure, the UNION of roles is granted to the user. One group can be assigned only one role. Selected users can be added from the groups.

In case of openLdap as the authentication type, the DN string of the user is required to login into DgSecure. In case of Active Directory as the authentication type, user enters the samAccountName or the userPrincipalName to login into the application.

The users that are logged in through a group are virtual/shadow users and will not be displayed in the Users List. However, these users will be counted as licensed users and will have impact on the licensing.

Shadow/Virtual users can be added at any stage as a defined DgSecure User from the Add User option. This manually created user will override the group user and the assigned role will be given preference.

3.3 Administer IDPs

In this step, enter the necessary information to allow DgSecure controller to connect to the desired IDPs (Discover Databases IDP, Files IDP, Masker IDP, HDFS IDP, Monitoring IDP, Hive IDP, and/or Cloud IDP). Select only those IDPs appropriate for your installation.

You can define the connection to an IDP before physically installing it, but you must know the IP address of the prospective host machine and the port number that will be used.

The default port numbers for the DgSecure IDPs are:

- HDFS IDP – 8111
- Discover Databases IDP – 8889
- Files IDP – 8082
- Masker IDP – 8888
- Hive IDP – 9980
- Hadoop Control IDP – 8089
- Cloud IDP – 8111
- DSAR - 1433

If you choose to use a different port number, you will need to update the IDP's properties file. For instructions, see chapter 3.

*** To define which machines run DgSecure tasks, define specific IDPs. An IDP is composed of a name, hostname or IP address, a port number, and an assigned IDP type.

3.3.1 Create IDPs

1. In DgSecure Admin, choose **IDP Management > IDPs**. The **IDP Management** panel is displayed.
2. Create a record for each IDP you want to connect to, click **Add IDP**. The Add / Edit IDP dialog box is displayed.
    a) Enter a descriptive name for the IDP.
    b) Enter the hostname or IP address of the IDP's host machine.
    c) Enter the port number the IDP will use to send and receive information.
    d) Select the IDP type from the drop-down menu.
3. Click **Save**.
4. The next step depends on the IDP.
    a. If you are creating an HDFS IDP (aka Hadoop Data IDP), Hive, Flume, Hadoop Control IDP (aka Hadoop Control IDP), or Cloud IDP:
        i. Click **Test Connection**.
    i. Choose **Clusters** from the side menu.
    ii. On the **Cluster Management** page, select **Add Cluster.** In the dialog box, select the IDP.
    iii. For HDFS IDPs, optionally set a SSL type.

    *High availability (HA) is available for the HDFS IDP. In order to set up HA, install 2 IDPs and set one IDP as the secondary IDP on the cluster. For more information, see section 3.6.

    2. If you are creating an IDP,
    i. Choose **Groups/ IP Ranges**
    ii. Add an IP range by clicking **Add Range** and entering the IP range details into the dialog box that appears. Click **Save**.
    iii. Choose the recently created IP range and click the **Add** button in the bottom panel. Select the IDP. Only one type of each IDP can be assigned as the primary IDP to a particular range.

    *High availability is available for the Discover and Files IDPs. In order to set up HA, install 2 IDPs and set one IDP as the secondary IDP on the cluster. For more information, see section 3.6.

NOTE: 1) While creating IDP of any type, a default IP Range corresponding to IP from IDP or IP obtained from hostname of IDP will be created.

In case of single agent of any type, IP address field of On-Premises RDMS connection parameters will be auto mapped to corresponding IDP, i.e. all the connections will be auto mapped to corresponding type of IDP. No need to define or edit any IP range.

2) While adding another agent of same type user need to explicitly define the IP Range for both new and existing IDPs. Because in that case first agent will handle only the connections having ip address same as that of IDP only.

Note: First agent will not handle all other IPs not belonging to other agent ip range, i.e. it will work the similar way IP ranges were working prior to this new change.

3.3.2 Edit IDPs

1. In DgAdmin, choose IDP Management > IDPs. The IDP Management panel is displayed.
12. Select the entry you want to edit. Click the **Edit IDP** button.
13. Make any needed changes to the IDP details.
14. Click **Save.**

3.3.3 Delete IDPs

1. In DgAdmin, choose IDP Management > IDPs. The IDP Management panel is displayed.
15. Select the entry you want to remove. The **Delete IDP** button is enabled.
16. Click **Delete IDP**. A pop-up message asks you to confirm your intention to delete the entry.
17. Click **Yes**. The entry is removed.

3.3.4 Manage Clusters/Fileshare

DgSecure can connect to multiple clusters. Once connected, DgSecure can discover, mask, and encrypt sensitive data on the cluster. The **Manage Clusters/Fileshare** page in DgAdmin allows users to manage their Hadoop IDPs and their relationship to specific clusters. To navigate to this page, click **Clusters** on the left-side menu.

To add a cluster:

1. Click **Add Cluster/Fileshare**.
2. Enter the name in the **Clusters/Fileshare Name** text box.
3. Select the type of the cluster/fileshare in the **Clusters/Fileshare Type** drop-down. The available types are: Hadoop, Files, Azure Data, AWS S3 and Google Big Query.

4. Select the location in the **Location** drop-down.
5. This field will be available when you select **Hadoop** in the **Clusters/Fileshare Type** drop-down. Select the primary and secondary IDP.
6. This field will be available when you select **Hadoop** in the **Clusters/Fileshare Type** drop-down. Select the Hive, Flume, Sqoop and DgWalker IDP.
    18. This field will be available when you select **Hadoop** in the **Clusters/Fileshare Type** drop-down. Enter LDAP Parameter details in order to utilize LDAP/LDAPS users for DgSecure's ACL management.
7. Select cluster location.
8. This field will be available when you select **Files** in the **Clusters/Fileshare Type** drop-down. **Select Files IDP.**
9. This field will be available when you select **Files** in the **Clusters/Fileshare Type** drop-down. Select **S3 Cloud IDP**.
10. This field will be available when you select **Azure Data** in the **Clusters/Fileshare Type** drop-down. Select **Azure Data IDP**.
11. This field will be available when you select **Azure Data** in the **Clusters/Fileshare Type** drop-down. Select **Azure Cloud IDP**.
12. This field will be available when you select **Azure Data** in the **Clusters/Fileshare Type** drop-down. Enter Azure AD Parameter details in order to utilize Azure users for DgSecure's ACL management.
13. This field will be available when you select **AWS S3** in the **Clusters/Fileshare Type** drop-down. Select **S3 IDP**.
14. This field will be available when you select **AWS S3** in the **Clusters/Fileshare Type** drop-down. Select **S3 Cloud IDP**.
15. This field will be available when you select **AWS S3** in the **Clusters/Fileshare Type** drop-down. Select **DgWalker IDP**.
16. This field will be available when you select **Google BigQuery** in the **Clusters/Fileshare Type** drop-down. Select **Big Query IDP**.
17. Click **Save**.

To edit a cluster:

1. Click **Edit Cluster/File Share**. The **Add / Edit Clusters/Fileshare** dialog box appears.
19. Edit the details that you want to edit.
20. Click **Save**.

To delete a cluster or fileshare:

1. Select the cluster or file share that you want to delete.
2. Click **Delete.**

3.3.5 AWS Configuration

Using the DgSecure Cloud IDP, users can create an EMR cluster in order to access S3 data. When the cloud IDP creates, it also creates a DgSecure S3 IDP on that cluster. The S3 IDP allows DgSecure to run data detection scans on the S3 repository. Currently, only one AWS configuration is supported.

To add a cluster:

1) Click **Configure**.
2) The Add/Edit Configuration dialogue box appears. Enter configuration details:
    a) Compute Cluster Name (Required): Enter the name you would like for the cluster.
    b) Instance Type (Required): Select a cluster type.
    c) Instance Count (Required) Specify the number of compute clusters to create. *Currently, only one compute cluster is recommended per Cloud IDP.
    d) "Use Default Roles" checkbox: Check to use default AWS roles.
    e) Service Roles (Required): Only required if (D) not selected.
    f) Log URI (Required): Location of EMR logs
    g) Tags: Create labels for the machines included in the cluster.
    h) "Visible to all users" checkbox (Required): If checkbox is not selected, only the user who created the role can see the cluster.
    i) Key Name: AWS keypair to securely access the cluster with ssh.
    j) Release Label: Select EMR version number from dropdown, this is set to 5.19 by default.
    k) AMI Version: Amazon version deployed
3) Click **Save**.

To Provision a cluster:

1. Select the AWS Configuration
2. Click **Provision Compute Cluster**

To Delete a cluster:

1. Select the AWS Configuration

2. Click **Destroy Compute Cluster**

3.3.6 GCS Configuration

Using the DgSecure Cloud IDP, users can create an Dataproc cluster in order to access GCS data. When the cloud IDP creates, it also creates a DgSecure GCS IDP on that cluster. The GCS IDP allows DgSecure to run data detection scans on the GCS repository. Currently, only one GCS configuration is supported.

To add a cluster:

1. Click Configure.
2. The Add/Edit Configuration dialogue box appears. Enter configuration details:
    a) Compute Cluster Name (Required): Enter the name you would like for the cluster.
    b) Master node machine type (Required): Select a machine type for the master machine. Please see Google's documentation to ensure selection of the ideal instance type.
    c) Worker node machine type (Required): Select a machine type for the worker machine(s). Please see Google's documentation to ensure selection of the ideal instance type.
    d) Worker node count (Required): Sets the number of worker nodes to be created.
    e) Subnet ID: Defines the subnetworks TCP/IP address. Specify this value when the cluster must be created on a specific subnet.
3. Click **Save**.

To Provision a cluster:

1. Select the GCS Configuration.
2. Click **Provision Compute Cluster**

To Delete a cluster:

1. Select the GCS Configuration
2. Click **Destroy Compute Cluster**

3.4 Other Admin functions

This section covers license management, and SSL certification.

3.4.1 Manage Licenses

To manage licenses, you can:

- Activate or deactivate a license by checking or unchecking the box in the Active column.
- Delete a license that has expired or is no longer required.

➢ **To view your installed license**

1. In DgAdmin, choose Licenses. The **License Management** page is displayed.
2. Existing licenses are listed in the **Installed Licenses** panel. For details, click a license. Details are displayed in the License Details panel.

3.4.1.1 Install a license

➢ **To install a license**

1. Copy the license (.lic) file to your computer.
2. In DgAdmin, choose Licenses. The **License Management** page is displayed.
4. In the License Management panel, enter the path to the license file, or click **Browse** to navigate to it.
5. Click **Install**.

Note: If you add a new license and a new product to your DgSecure environment, in DgAdmin, choose User Management > Roles. Review each user's permissions and modify user's permissions as needed to enable them to access the new product. (The new product is not visible to users who don't have the necessary permissions to access it.)

3.4.1.2 Activate and deactivate licenses

When you install a new license, by default it supersedes any previously-installed licenses, and the older licenses are deactivated. In some instances, an older license may support features not supported by a new license. When that is the case, you can choose to reactivate the older license. To view the net terms of multiple active licenses, click Effective License.

➢ **To activate a license**

1. In DgAdmin, choose Licenses. The **Licenses** page is displayed.
2. In the **Installed Licenses** panel, locate the license that you want to activate.
6. In the **Active** column, click the checkmark.

7. Click **Save**. A pop-up message confirms the change.

➢ **To deactivate a license**

1. In DgAdmin, choose Licenses. The **Licenses** page is displayed.
2. Locate the license in the **Installed Licenses** panel.

8. De-select the checkbox in the **Active** column.
9. Click **Save**. A pop-up message confirms the change.

3.4.1.3 Delete a license

If a license has expired or has been superseded by a newer license, you can delete it.

➢ **To delete a license**

1. In DgAdmin, choose Licenses. The **Licenses** page is displayed.
2. In the **Installed Licenses** panel, select the license that you want to delete.
10. Click **Delete License**. A pop-up message asks you to confirm your intention to delete the license.
11. A pop-up message confirms that the license was deleted.

3.4.1.4 Manage multiple licenses

If your DgSecure installation uses multiple active licenses, DgSecure treats them as the building blocks of a single, site-wide license called the "effective license". To determine the effective license, DgSecure applies the following rules.

License Component	Effect when licenses are combined
Start Time	The later start time applies to all licensed products.
Days	The greater number of days applies to all licensed products.
Users Connections Controllers IDPs	The greater number of users, connections, controllers, and IDPs applies to all licensed products.

For example, if a DgSecure installation uses the following licenses:

```
License 1
Product(s): Discover
Start Time: 06/01/2012
Days: 30
Users: 5
Connections: 50
Controllers: 1
IDPs: 2
License 2
```

```
Product(s): Discover MF
Start Time: 06/15/2012
Days: 60
Users: 10
Connections: 10
Controllers: 5
IDPs: 10
```

Then DgSecure treats these two licenses as a single license with the following characteristics:

```
Product(s): Discover, Discover MF
Start Time: 06/15/2012
Days: 60
Users: 10
Connections: 50
Controllers: 5
IDPs: 10
```

3.4.1.5 Protect license files

When you install a license, the DgSecure controller reads the license details and creates a copy of the file in the webapps/license directory within the DgSecure product folder.

It is important to make sure that this folder is not altered in any way. If the folder (or any of the files within it) are moved or deleted, DgSecure will not function correctly.

3.4.2 Modify SSL Certificates

You can modify the Secure Socket Layer (SSL) certificates created when DgSecure was installed.

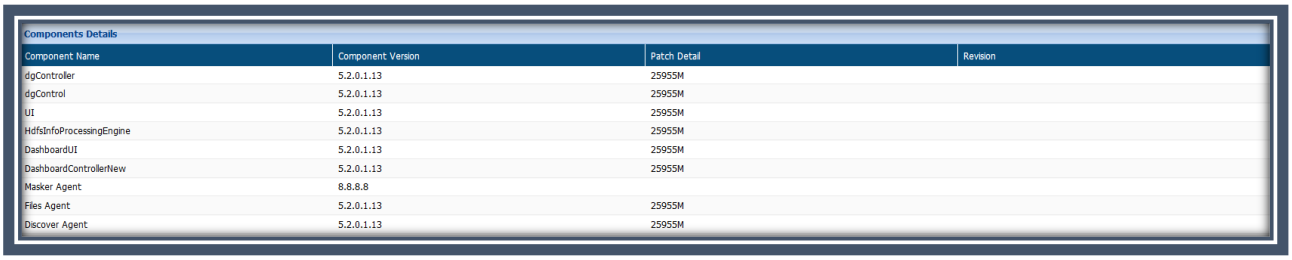
➢ **To modify SSL certificates**

1. In DgAdmin, choose Settings. The **SSL Management** panel is displayed.
12. Select the certificates you want to enable: Repository, Directory Service, HDFS IDP, or Enable All.

3.4.3 View Component Versions

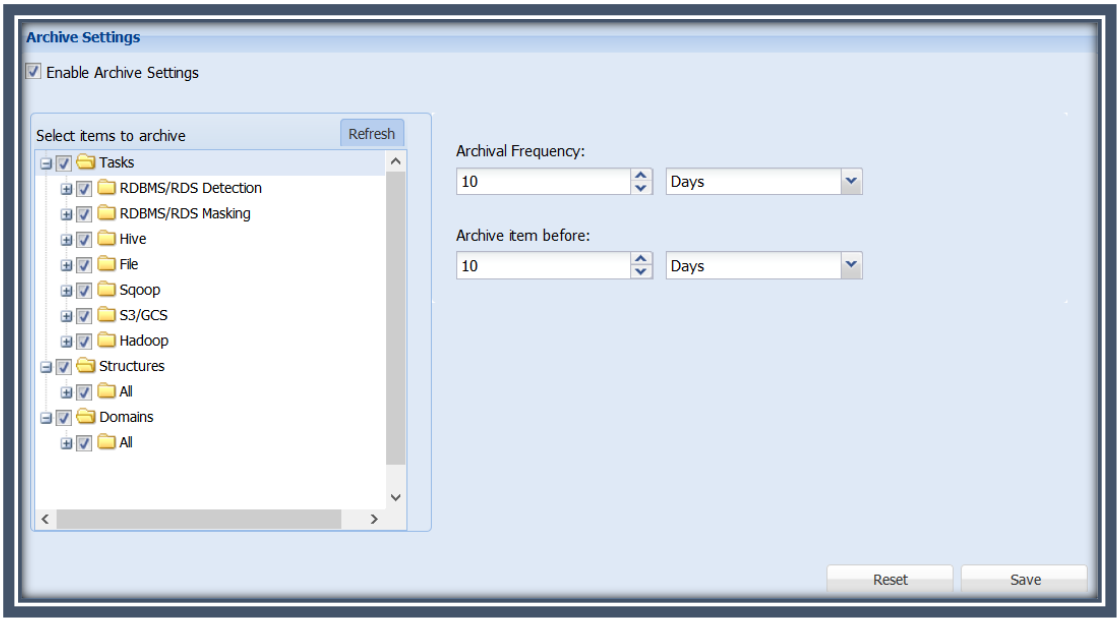
You can view the version and patch number for any installed DgSecure component on the **Component Version** page. Navigate to this screen using the menu on the left side of the screen.

Example:



3.4.4 Archive Tasks, Structures, Domains, and RDBMS Connections

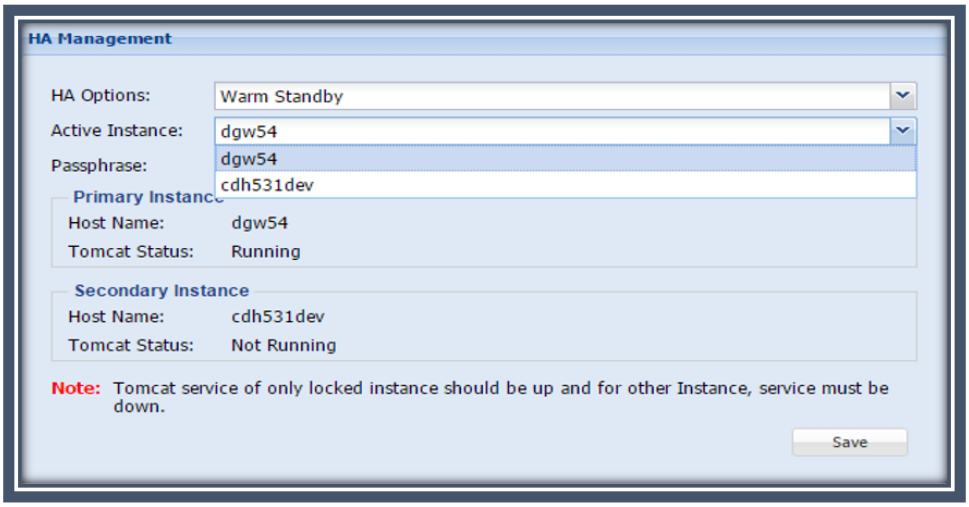
You can enable archiving for tasks, Hadoop structures, domains, and or RDBMS connections on the **Archive Settings** page. Set the frequency at which items are archived using the "Archival Frequency" field. Set the parameters for which items should be archived using the "Archive Item Before" field. For instance, if "Archival Frequency" were set to 10 days and the "Archive Item Before" were set to 5 days, DgSecure would archive any selected item older than 5 days every 10 days. The period of time can be measured in either days or hours.



3.4.5 Warm Standby

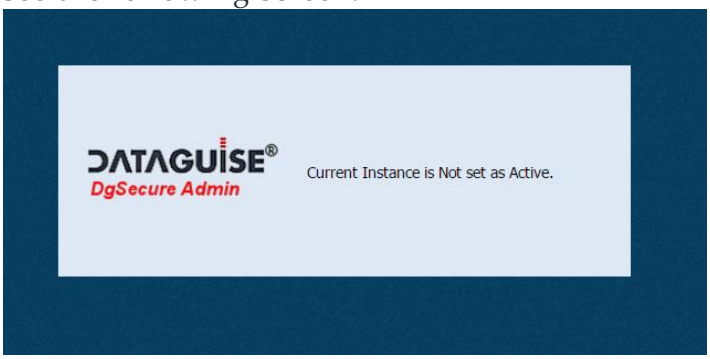
Warm Standby configuration requires that 2 DgSecure Instances are pointing at the same backend database (metadata repository). One important condition for Warm Standby to work is that only one of the Tomcat Services should be up at any time.

When warm standby is enabled by selecting "Warm Standby" as the HA option, the user can select which instance should be considered active. Select the passphrase that will be used when switching active instances. Click "Save".



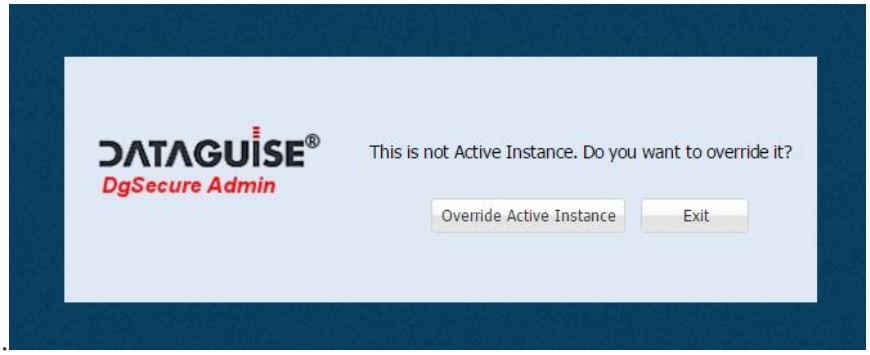
Switching Active Instance:

1. If the user tries to bring up the Secondary DgSecure, the user will see the following screen:

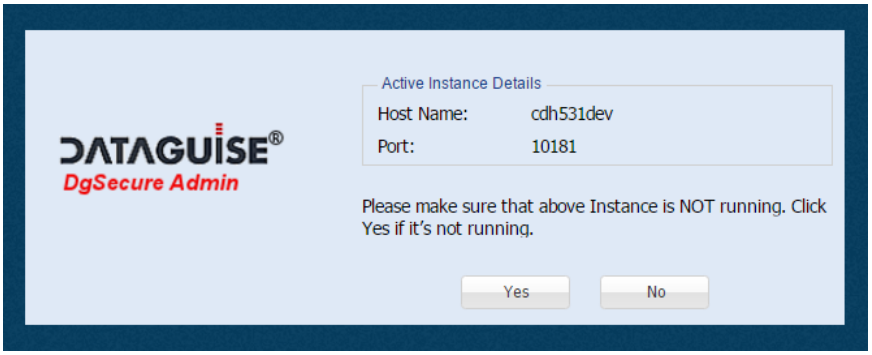


2. When DgSecure is starting up, it checks the DBMS table in the metadata repository – it verifies if the "Warm Standby" choice is selected.
3. It then checks to see if another instance is Active.

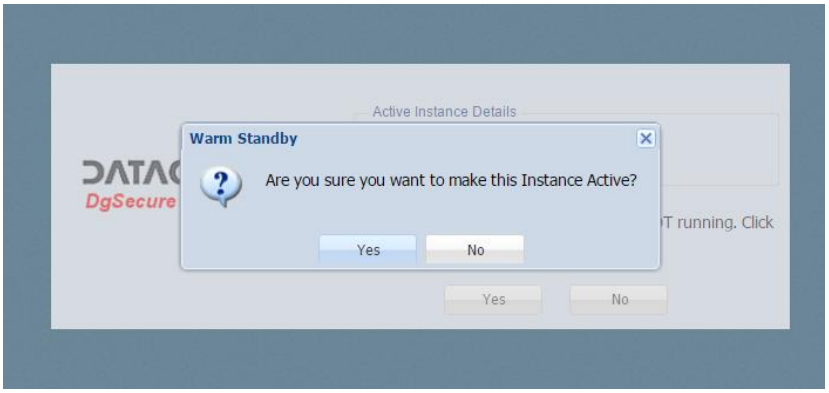
4. If it finds another instance is Active, it gives the below warning. Click “Override Active Instance.”



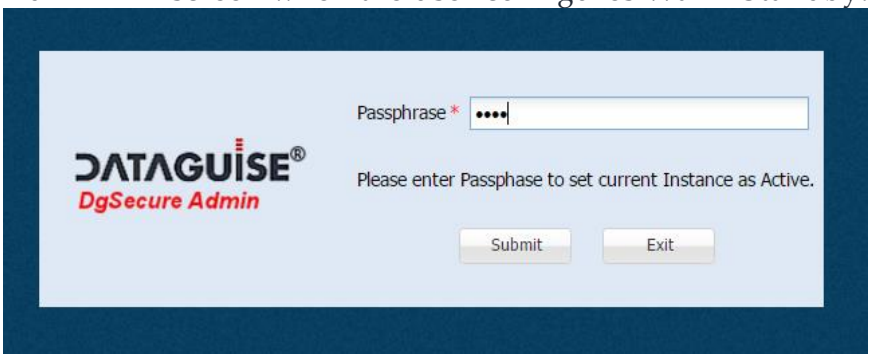
5. If Choice 1 is chosen, a message displays asking the user to confirm the instance is not running. Click “Yes”.



6. DgSecure asks for another verification that the user wants to switch active instances. Click “Yes”.



7. It will then ask to enter the passphrase. It needs to be set from the Admin HA screen when the user configures Warm Standby.



After authenticating, switch the Active Instance and user will be redirected to login screen.

Notes:

1. User should not be able to log into main UI of Inactive DgSecure Instance. Instead they will get an error message "Please configure this instance as the Active one in DgAdmin."
2. If Active Instance is switched, DgSecure invalidates all the active sessions. This avoids multiple login sessions from both instances.

3.4.6 Manage Source Systems

Use this page to turn the DgSecure monitoring on or off for specific source systems.

Add:

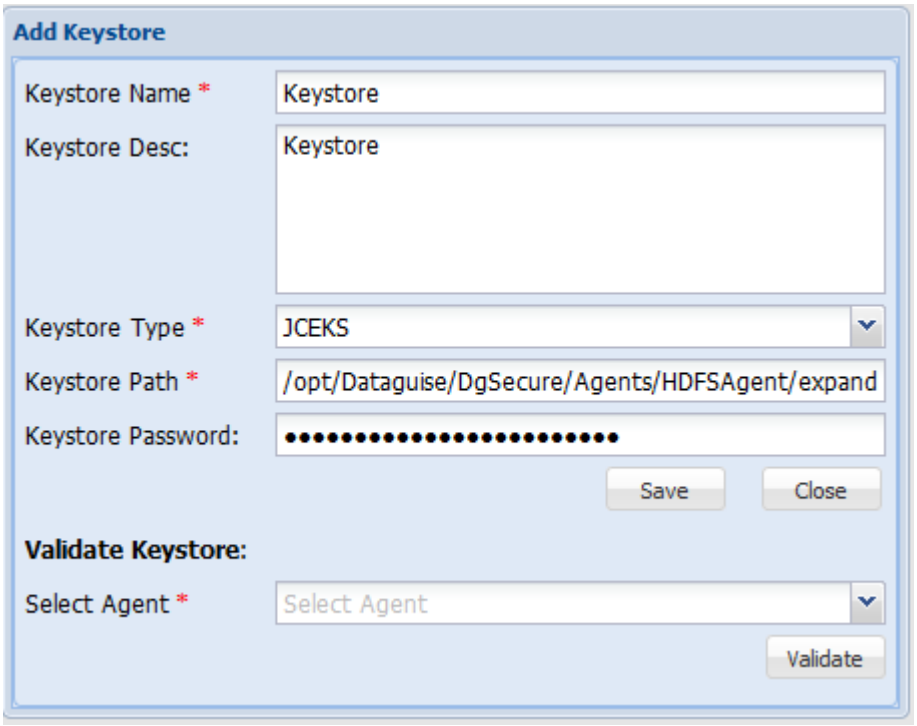
1. Click **Add Source Systems**.
2. When the dialog-box appears, enter the source system details.
3. Click **Save**.

Start/Stop

1. Optionally, Click **Start All** to begin monitoring or **Stop All to cease monitoring activity.**
2. Alternatively, select a source system and start/stop monitoring on it.

3.4.7 Manage Keystores

Use this page to manage the Keystores that will be used for encryption and decryption in HDFS. Below is an example of a sample keystore:



3.5 Tableau

Tableau is a reporting tool that DgSecure has integrated with to provide comprehensive and flexible reporting on sensitive data analysis.

3.5.1 Prerequisites

Ensure that the following requirements are met before setting up Tableau:

- The user must have Permissions and Grants to read and access the servers on which DgSecure and Tableau are installed. The required permissions are:

  PostrgeSQL: Execute the permissions from the file available at the following location:

  ```
  /etc/postgresql/10/main/pg_hba.conf
  ```

  MySQL: Provide the following grant to the user:

  ```
  GRANT ALL PRIVILEGES ON . TO 'root'@'<TARGET SERVER IP>' IDENTIFIED BY '<PASSWORD>' WITH GRANT OPTION ;
  ```

- Drivers should be installed to the target Tableau server to enable reporting.
- For MySQL and PostgreSQL backend, Tableau Server and Controller IP have to be manually configured to the databases.

  PostgreSQL: Add the IPs to the file at the following location:

`/etc/postgresql/10/main/pg_hba.conf`

MySQL: Provide permissions for the required IPs through the MySQL client.

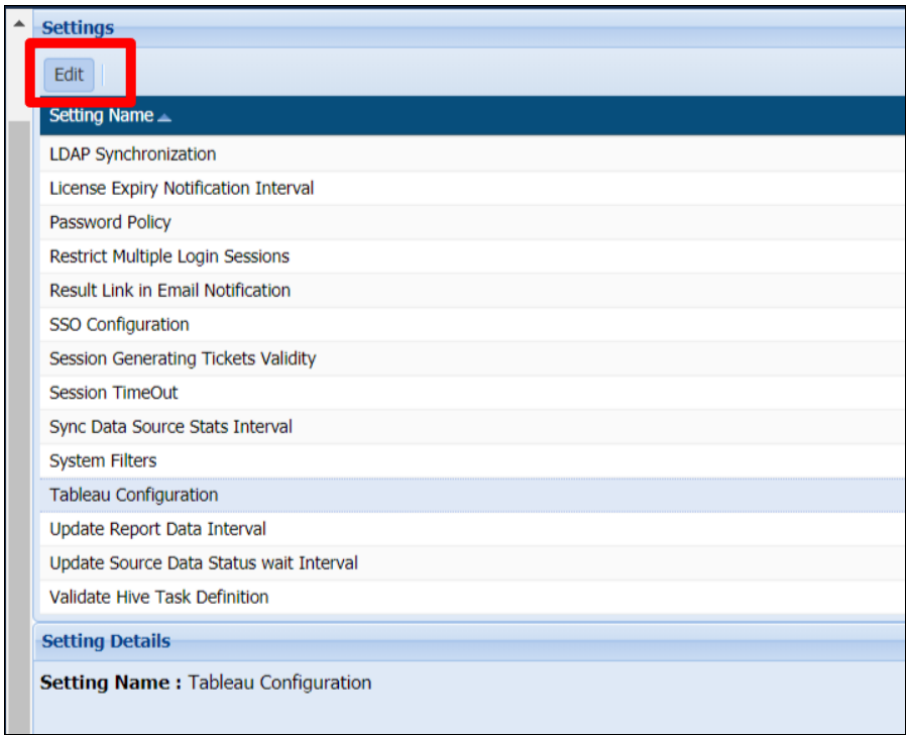
- Supply the hostname details to the machine where DgSecure accessed as well as on the machine where DgSecure is installed, (`eg. hostname : http://xen192-tableau-centos`) in the file located at: (Login using root)

  `/etc/ hosts`

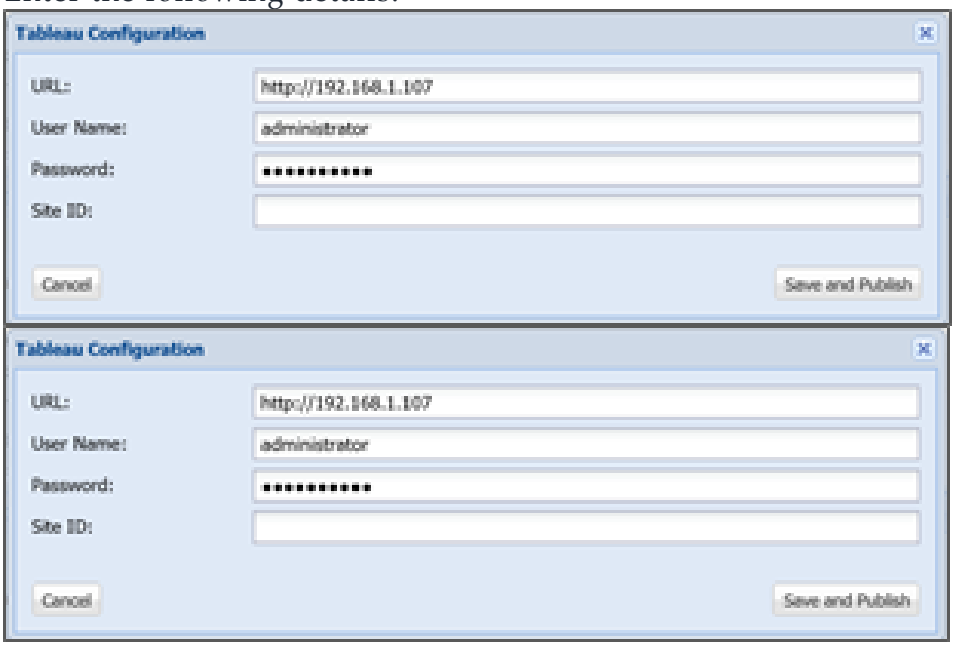
3.5.2 Configure Tableau

Perform the following steps to configure Tableau to DgSecure:

1. Go to **DgSecure Admin>Settings>Tableau Configuration**. Click **Edit**.



2. Enter the following details:



URL: Provide the Server IP/ Host name of the tableau server.

User name: Provide the user name.

Password: Provide the password.

Site ID: Provide the Site ID where you want to publish the report on the tableau server. You can publish the report under your particular site. However, if you will not provide the Site ID, the report will be published under the default site.

NOTE: If you are providing the server IP as URL, then hostname details need not be specified in the host file.

3. Save and publish.

NOTE: DgSecure Admin overwrites any existing reports. Ensure that a backup has been taken from the Tableau server if any older reports are required.

Appendix A: Using Master-Slave Controllers

If you want to run DgSecure on more than one controller, but use only a single set of policies, you can export them from a master DgSecure Controller and import them into one or more secondary Controllers.

Before your export and import policies, verify that your policies are correctly defined.

Note: You cannot export or import preloaded (pre-existing) policies.

To export DgSecure policies:

1. Login to DgSecure on the first or primary controller containing the policies you want to export.
2. Choose DgPolicy > Export/Import > Policies. The Policy Import / Export panel is displayed.

To…	Click…
Export all existing Policies	Full Export
Export Policies created since the last export	Incremental Export

3. You're prompted to save the policydetails.json file. Save it to your preferred location.

To import policies into the second controller, you also need to modify the policySlave properties setting found in DgConnection.properties on that server.

To enable a second DgSecure Controller to import policies:

1. Login to the server hosting the second controller.
2. Navigate to the directory where Tomcat was installed. If you installed Tomcat with DgSecure, the root path is:

   `C:/Program Files/Dataguise/DgSecure`

3. Navigate to the hibernate folder under Tomcat.

   `<InstallationPath>/tomcat9/webapps/dgcontroller/WEB-INF/classes/com/dataguise/hibernate/`

4. Edit DgConnection.properties. Find the parameter policySlave.
5. Change the value of policySlave=false to policySlave=true.

6. Save your changes and exit.
7. Open the Services Console and restart the Apache Tomcat Service.

To import DgSecure policies:

1. Login to DgSecure on the secondary controller into which you want to import the policies you created on the primary DgSecure Controller.
2. Copy the policydetails.json file you saved from the primary Controller to the secondary Controller.
3. Choose DgPolicy > Export/Import > Policies.
4. Click Browse. Navigate to the location where you copied the policydetails.json file. Open the file.
5. Click Import. The policies are saved to the controller.

Appendix B: InstallDgSecure.sh parameters

When installing DgSecure on Linux from the command prompt, run DgSecure-<version>-linux-installer.run. You can enter parameters in InstallDgSecure.sh; the install flow and prompts exactly mirrors the described GUI sequence. The installer must have administrative privileges to install successfully. Below is the list of important parameters, the values of which the user needs to mention carefully to have DgSecure installed successfully. After saving the parameters in this file, just select the file and press enter to install DgSecure.

1. **Enter the location where the installer has been placed**.
   # *INSTALLER NAME
   InstallerName="./DgSecure-4.5.0.6a-linux-x64-installer.run"

2. **Path where DgSecure will be installed.**
   # ----------INSTALLATION DIRECTORY----------
   # *INSTALLATION PATH
   InstallPath="/opt/Dataguise"

3. **Keeping ssloption = 1 will install the application with SSL Enabled.**
   # ----------COMPONENTS SELECTION OPTION (SSL AND ACTIVE DIRECTORY)----------
   # *IF USER WANT TO INSTALL SSL OPTION, 1 = YES 0 = NO
   ssloption="1"

4. **Below are SSL Certificate Details**
   # ----------SSL CERTIFICATE----------
   # *FULL NAME
   FullName="dataguise"
   # *ALIAS NAME
   AliasName="dataguise"
   Unit=""
   Organization=""
   Locality=""
   State=""
   # *COUNTRY CODE TWO CHARS ONLY
   CountryCode="US"
   # *PASSWORD ATLEAST 6 CHARS
   SslPassword="dataguise"

5. **If DatabaseType="PostgreSQL", application will be installed with Postgres.**
   If DatabaseType="MySQL", application will be installed with MySQL.

```
# ----------DATABASE TYPE SELECTION----------
# DATABASE TYPE PostgreSQL OR MySQL
DatabaseType="PostgreSQL"
```

6. **Postgres Details when application is installed with Postgres**

```
# ----------PostgreSQL CONFIGURATION----------
# *POSTGRESQL PORT DEFAULT=5432
PostgreSQLPort="5432"
# *POSTGRESQL PASSWORD
PostgreSQLPassword="dataguise"
# ENABLE SSL, 1 = YES 0 = NO
pgssloption="1"
```

7. **MySQL Details when application is installed with MySQL**

```
# ----------MySQL CONFIGURATION----------
# *MySQL PORT DEFAULT 3306
MySQLPort="3306"
# *MySQL HOST
MySQLHost="localhost"
# *MySQL USER NAME
MySQLUser="root"
# *MySQL PASSWORD
MySQLPassword="root"
```

8. **Below are TOMCAT configuration details**

```
# ----------TOMCAT CONFIGURATION----------
# *TOMCAT PORT, DEFAULT=10181
TomcatPort="10181"
```

9. **Mention IDP type (i.e CDH531 or MaprR401) See support matrix for supported distributions**

```
# ----------HDFSIDP Type----------
# HDFSIDPType MapR301 OR MapR211 OR CDH4
HDFSIDPType="MapR401"
```

10. **Mention the path where MapR client is installed**

```
# ----------MAPR CLIENT----------
# MapR CLIENT PATH
MapRClientPath=""
```

Appendix C: Snappy Compression

Typically, DgSecure does not require any additional libraries to read compressed files. However, there are occasions where it may be necessary to install independently a library to read the Snappy compressed file structure. In such situations, follow the instructions below.

Verify Snappy compressed file structure (IDP on Linux): In order to verify the structure of a Snappy compressed file when the HDFS IDP is installed on Linux, ensure that native snappy library libraries are installed (these files typically install as part of Apache Hadoop). If they are not installed, download Snappy libraries and copy the library files to $HADOOP_HOME/lib/native. Alternatively, just install Hadoop-0.20-native package.

Run HDFS tasks on Snappy compressed file structure (IDP on Windows): DgSecure does not currently support structure verification for Snappy compressed files when the HDFS IDP is installed on Windows. In order to run an HDFS task on a structured Snappy compressed file, uncheck the “Verify” checkbox on the structure management page.

Appendix D: Updating Credentials on DgSecure Repository Database

When installing DgSecure and selecting DB authentication a DB user account is created (e.g. on SQL Server) and the encrypted service account name and password are stored in the DgConnection.properties file.

When password aging is instituted for DB accounts for security purposes, the service account password may require changes on periodic basis. When the account updates, the credentials need to be updated in DgSecure as well. Update the DB authentication credentials in each of the following files below. Use DGCL to generate the encrypted entry.

DGCL Command:

encrypt " <stringToEncrypt>";

Copy the encrypted value (e.g. JWFH+XG1aqytPIlaUbr3fQ==) into the following files.

<InstallationPath>/DgSecure/tomcat9/webapps/dgControl/WEB-INF/classes/com/dataguise/dgcontrol/hibernate/DgConnection.properties

<InstallationPath>/DgSecure/tomcat9/webapps/dgcontroller/WEB-INF/classes/com/dataguise/hibernate/DgConnection.properties

<InstallationPath>/DgSecure/tomcat9/webapps/dgDashboardControllerNew/WEB-INF/classes/com/dataguise/dashboard/database/DgConnection.properties

<InstallationPath>/DgSecure/tomcat9/webapps/dgHdfsInfoProcessingEngine/WEB-INF/classes/com/dataguise/config/dgConnection.properties

<InstallationPath>/DgSecure/DgSettings.properties

Appendix E: Active Directory with SSL

There are two cases addressed here. The first case considers how to configure DgSecure with AD and SSL on a fresh install. The second case addresses how to change the authentication method on an existing installation.

CASE 1

Step 1: First install the build.

Step 2: After installing the build, give the path of the keystore and password

(in encrypted format) in “DgConnection.properties” file.

```
/opt/Dataguise/DgSecure/tomcat9/webapps/dgcontroller/WEB-
INF/classes/com/dataguise/hibernate/DgConnection.properti
es
```

Step 3: Update below mentioned properties in “DgConnection.properties” files.

```
postgresTrustedStoreLocation = /opt/Dataguise/DgSecure
/DgCertificate/DgTestCertificate.jks

postgresTrustedStorePassword =40z/H6qLk8eSl09PKlyLfg==
```

****Note:** By default, “dataguise” encrypted password is mentioned in above mentioned property.

Step 4: We can encrypt the “postgresTrustedStorePassword” using dgcl with the below mentioned command:

```
encrypt “dataguise”;
```

Step 5: Restart tomcat service

Step 6: Open DgSecure Admin page for installing the license

Step 7: After the license is installed, enter Secure AD detail

Note: For LDAPS protocol default port is 636. For LDAP protocol default port is 389

Step 8: Now, enter the user details in next page

Now, you can successfully log into DgSecure with Secure Active Directory details.

CASE 2

Step 1: First install the build

Step 2: After installing the build, give the path of the keystore and password (in encrypted format) in "DgConnection.properties" file.

```
/opt/Dataguise/DgSecure/tomcat9/webapps/dgcontroller/WEB-INF/classes/com/dataguise/hibernate/DgConnection.properties
```

Step 3: Update below mentioned properties in "DgConnection.properties" files.

```
postgresTrustedStoreLocation = C\:\\Program
```

```
Files\\Dataguise\\DgSecure\\DgCertificate\\DgTestCertificate.jks

postgresTrustedStorePassword =40z/H6qLk8eSl09PKlyLfg==
```

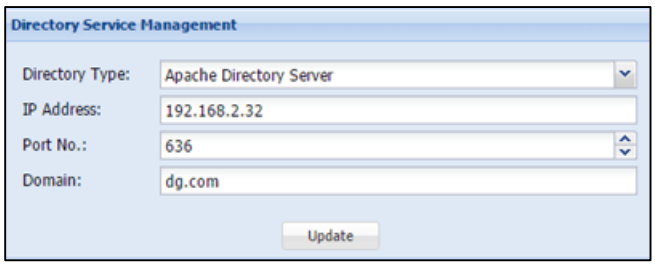
Note: By default, "dataguise" encrypted password is mentioned in above mentioned property.

Step 4: We can encrypt the "postgresTrustedStorePassword" using dgcl with the below mentioned command:

```
encrypt "dataguise";
```

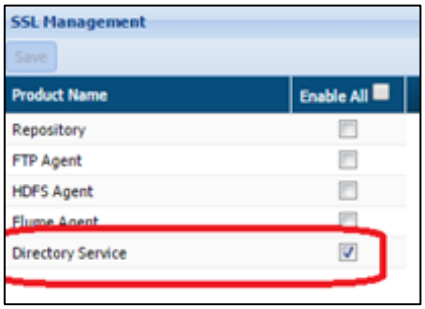
Step 5: Restart tomcat service

Step 6: Go to DgSecure Admin > "Authentication page" and update the new AD details



Note: For LDAPS protocol default port is 636. For LDAP protocol default port is 389

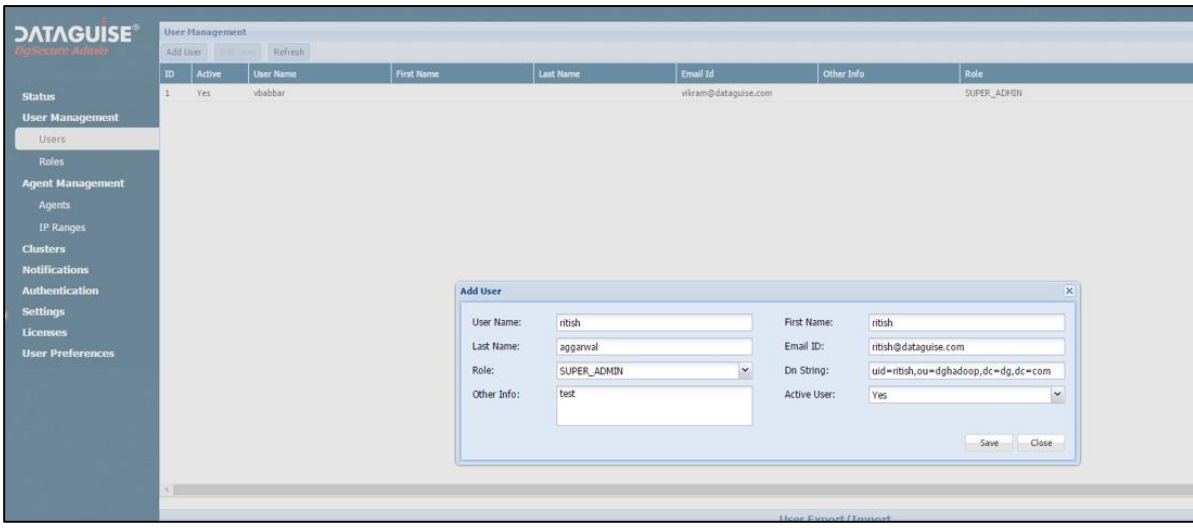
Step 7: If we are using the secure AD, then we have to check "Directory Service" checkbox from DgSecure Admin > Settings page.



Note: When we checked the Directory Service checkbox then "LDAPS" protocol value saved in backend.

If user did not check the Directory Service checkbox, then "LDAP" protocol value saved in backend.

Step 8: Go to DgSecure Admin > “Users page”, and add/update Active Directory user details



Step 9: Restart the tomcat service.

Step 10: Now login with updated authentication method or with new user details.

Appendix F: Enable Database Logging for Monitoring

In order to monitor Oracle, Teradata, or SQL Server using DgSecure Monitor, database logging must be turned on in each respective database that needs to be monitored.

Oracle:

1. Check if audit_trail is enabled or not:
2. `Show parameter audit;`
3. If auditing is not already enabled, initialize auditing in Oracle:

```
Alter system set AUDIT_TRAIL=db, extended
scope=spfile;
```

****NOTE:**
The database initialization parameter **AUDIT_TRAIL** enables and disables auditing. The default setting for this parameter is **NONE**, which means that no auditing will be performed. When `AUDIT_TRAIL = db:`, all audit records are directed to the database audit trail (the **SYS.AUD$** table), except for records that are always written to the operating system audit trail. Use this setting for a general database for manageability. When `AUDIT_TRAIL = db, extended:` The database performs all actions of `AUDIT_TRAIL=db`, in addition to populating the SQL bind and SQL text CLOB-type columns of the `SYS.AUD$` table, when available. These two columns are populated only when this parameter is specified.

4. DE initializing auditing in Oracle

```
ALTER SYSTEM SET AUDIT_TRAIL=NONE SCOPE=SPFILE
```

5. Enable Auditing

   a. By table

   ```
   Audit insert, update, select, delete on
   USERNAME.TABLENAME by access
   ```

   b. By user

   ```
   AUDIT ALL BY mannattest BY ACCESS;
   AUDIT SELECT TABLE, UPDATE TABLE, INSERT TABLE,
   DELETE TABLE BY mannattest BY ACCESS;
   AUDIT EXECUTE PROCEDURE BY mannattest BY ACCESS
   ```

6. Check that audit logging is turned on

```
SELECT * FROM DBA_AUDIT_TRAIL;
```

Teradata:

1. Connect to Teradata using Teradata BTEQ.

2. Check if DBQL is already enabled or not:

```
show query logging on all;
SELECT * FROM DBC.DBQLRulesV;
```

3. If DBQL is enabled, no further action is required. If DBQL is not enabled, grant permissions to your admin account:

```
grant execute on DBC.DBQLAccessMacro to dbc;
```

4. Begin logging:

```
begin query logging with objects, sql, usecount,
utilityinfo LIMIT SQLTEXT=0 on all;
```

```
begin query logging with objects, sql limit threshold = 5
elapsedsec and sqltext=0 on VIEWPOINT;
```

5. Check that logging is turned on:

```
SELECT * FROM DBC.DBQLRulesV;
select * from DBC.QryLog
```

**Note:

WITH Logging Options

1. WITH ALL:

- One default row per query in DBQLogTbl that includes the first 200 characters of the SQL statement, unless you define LIMIT SQLTEXT=0.

2. WITH SQL:

- A default row in DBQLogTbl.
- The entire SQL statement for each request for each user being logged. Large statements can cause multiple rows to be written to log the full query text.

**NOTE:

DBQL is limited to logging information about base tables and logging direct SQL statements. Macros, views and triggers do not result in complete logging information.

If you set LIMIT SQLTEXT=0 when you specify the WITH SQL option, you avoid duplicate SQL in the DBQLogTbl.

SQL Server:

****NOTE:**

When a trace file is created, we can deactivate the trace using sql queries, but the file has to be deleted manually. So, user should either have access to the file system or if the logging is deactivated then the trace file with a different name or at different location should be created.

Giving the location of the trace file

The location of the trace file is to be saved in the file properties.config in the variable sqlserver.file.location

File path: `{Installation Path}\webapps\DgLogReader\WEB-INF\classes\com\dataguise\common\configproperties.config`

1. Check if SQL Trace is currently turned on:

```
SELECT * FROM sys.traces
```

2. Create a trace:

```
DECLARE
      @TraceID          INT,
     @maxSize           BIGINT
     SET @maxsize = 20

EXEC sp_trace_create
      @traceid          = @TraceID OUTPUT,
     @options           = 2, --@optionValue = 2 -- 2 is for
     creating more files when one file is filled completely
      @tracefile        = N'D:\RDSDBDATA\Log\createYourOwnTrace',
      @maxfilesize      = @maxsize
```

***Note:**

1. @TraceID should always be INT
2. @maxSize should always be BIGINT and set to maximum size permissable in the user's system
3. @tracefile is the location where there trace file will be created

3. Set new trace events. Trace events need to be created for TextData, HostName, LoginName, StartTime, ObjectName, and DatabaseName. SQL's ID for each of these column types are

   1 - TextData

8 - HostName

11 - LoginName

14 - StartTime

34 - ObjectName

35 – DatabaseName

The event ID should be 114 (this is the default ID for Audit Schema Object Access Events). Set the trace events with the following commands:

```
exec sp_trace_setevent
       @traceid = 2
       ,@eventid = 114
       ,@columnid = 1
       ,@on = 1
exec sp_trace_setevent
       @traceid = 2
       ,@eventid = 114
       ,@columnid = 8
       ,@on = 1
exec sp_trace_setevent
       @traceid = 2
       ,@eventid = 114
       ,@columnid = 11
       ,@on = 1
exec sp_trace_setevent
       @traceid = 2
       ,@eventid = 114
       ,@columnid = 14
       ,@on = 1
exec sp_trace_setevent
     @traceid = 2
     ,@eventid = 114
       ,@columnid = 34
     ,@on = 1
exec sp_trace_setevent
     @traceid = 2
     ,@eventid = 114
     ,@columnid = 35
       ,@on = 1
```

4. Activate the trace. Set the status to "1" to activate the trace. Set new trace events.

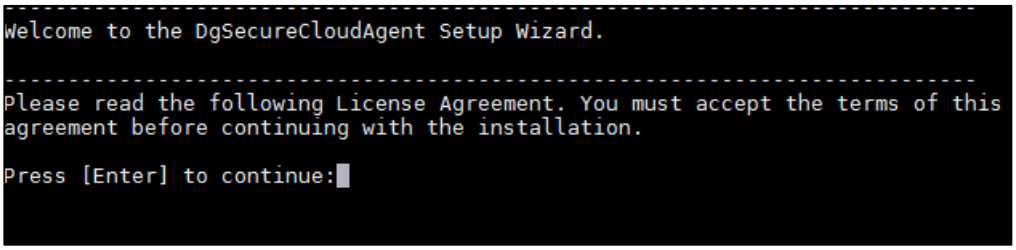
```
exec sp_trace_setstatus
      @traceid = 2,
      @status =  1 --Activate trace
```

Note: To deactivate a trace, set the status to “0”. To delete a trace, set the status to “2”.

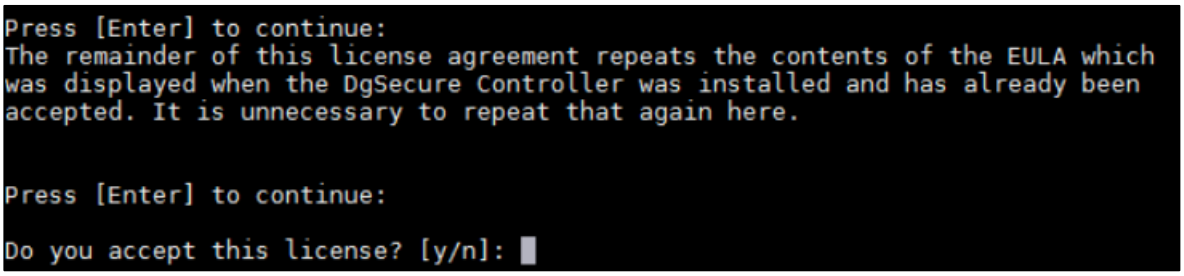
Appendix G: Cloud IDP Command Line Install

1. In the files that you received from Dataguise, locate the installation files. Copy the executable file **DgSecureCloudIDP-linux-x64-installer.run** to the machine on which you will install the IDP.

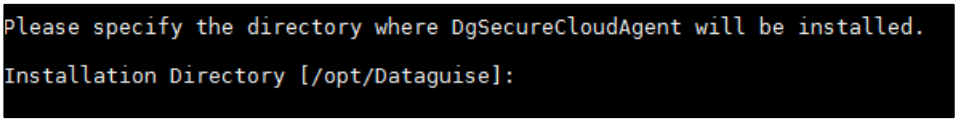
2. Run the installation file **DgSecureCloudIDP-linux-x64-installer.run**. The Setup Wizard is displayed. Press enter to continue.



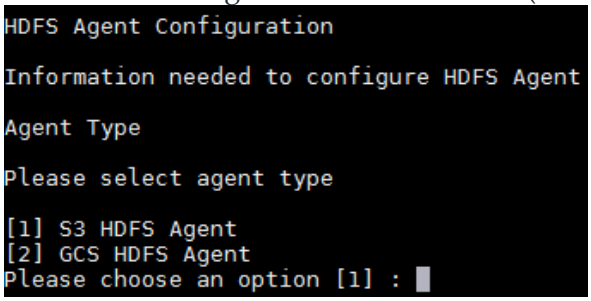
3. Accept the License agreement.



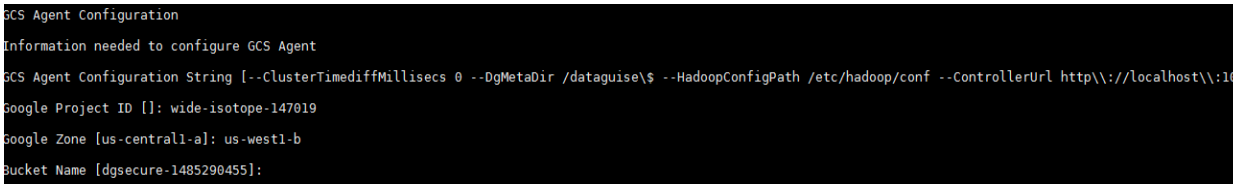
4. Install the utility in the default directory (/opt/Dataguise) or select a different location.



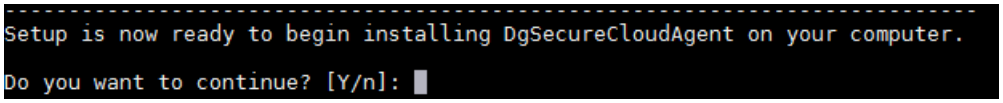
5. Select the configuration for the IDP (S3 or GCS).



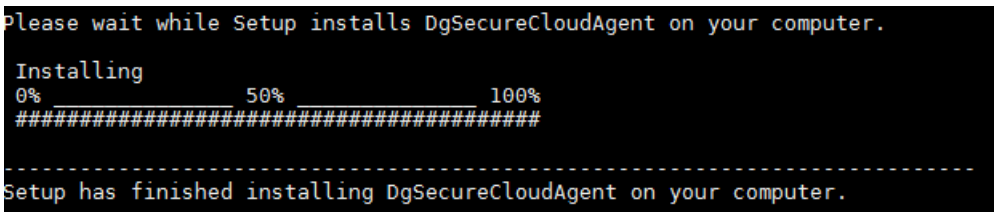
6. Enter the IDP's configuration details. (GCS example shown).



7. The Cloud IDP is now ready for installation. Click Next.



8. When installation is complete, Click Finish to exit the installer.



9. Check the IDP's property file to ensure the IDP is properly configured. For more details on IDP configuration for S3, please see chapter 3.4. For more details on IDP configuration for GCS, please see chapter 3.5.

Appendix H: SSL Type between HDFS IDP and Controller

SSL Type is set to No SSL

No change needed

SSL Type is set to 1-way SSL

For 1 way SSL we need to create keystore file at the server level (IDP in our case) and certificate is generated using this keystore file. The location of this keystore file and its password (encrypted form) is then added in the jetty.properties file of the IDP. The corresponding certificate location is specified under dgcontroller.properties file to be added under the java trust store of the client (controller in our case).

IDP Changes

- The Keystore and crt files needs to be created at the IDP side using the following 2 commands respectively:

  keytool -genkey

  -dname "CN=FULL_NAME, OU=UNIT, O=ORG, L=LOCALITY, ST=STATE, C=US"

  -alias "ALIAS_NAME"

  -keystore "kserver.keystore"

  -storepass storepassword

  -keypass

  -keyalg RSA -sigalg SHA1withRSA

  keytool -export

  -alias "ALIAS_NAME"

  -keystore kserver.keystore

  -storepass storepassword

  -keypass keypass

  -rfc -file kserver.crt

  modify /<InstallationPath>/Dataguise/DgSecure/IDPs/HDFSIDP/jetty-embedded.properties

keyStorePath =/path/to/kserver.keystore

keyManagerPassword = encrypted form of keypass

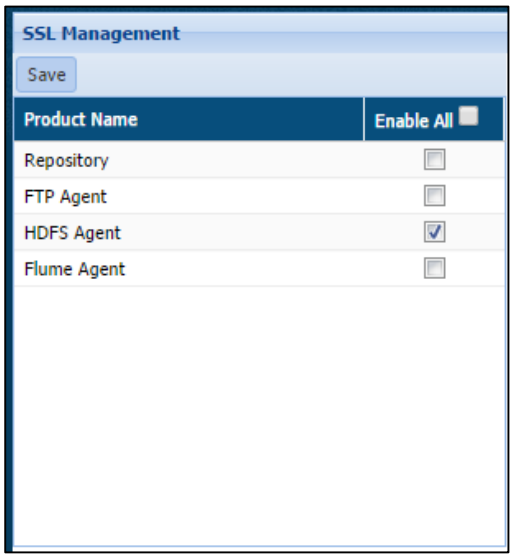
keyStorePassword = encrypted form of storepassword

needClientAuth = N

sslEnabled = Y

Controller Changes:

- The kserver.crt file generated at the IDP side needs to be copied on the machine where controller is installed.
- The path of this file needs to be specified under the variable "pathToCertKey" under the tomcat9\webapps\dgcontroller\WEB-INF\classes\ dgController.properties file.
- Restart tomcat.
- Login to DgSecure Admin and enable SSL under the SSL Setting tab for HDFS IDP.



- Confirm the IDP is working by doing Test Connection on that IDP under the IDP tab in DgSecure Admin.

SSL Type is set to 2-way SSL

For 2-way SSL, since the authentication needs to be done at both client and server end, we need to generate a set of 2 keystore files, one for the server(IDP) and another for the client(controller) for the server level authentication and similarly a set of 2 keystore files for the client level authentication. The location of paths of these keystore files with their corresponding passwords are then defined under the IDP and controller properties files respectively. These keystores files are then added under

trust manager and key manager of the SSL context at the controller side to perform a successful 2 way secure SSL communication between IDP and controller.

IDP Changes

- We shall be creating 2 keystore files for IDP side (server level) and 2 keystore files for controller side (Client level).

  IDP side:

  keytool -genkey

  -dname "CN=FULL_NAME, OU=UNIT, O=ORG, L=LOCALITY, ST=STATE, C=US"

  -alias "ALIAS_NAME"

  -keystore "kserver.keystore"

  -storepass storepassword

  -keypass keypass

  -keyalg RSA

  -sigalg SHA1withRSA

  keytool -export

  -alias "ALIAS_NAME"

  -keystore kserver.keystore

  -storepass storepassword

  -keypass keypass

  -rfc -file kserver.crt

  keytool -import

  -alias "ALIAS_NAME"

  -file kserver.crt

  -storepass storepassword

  -keypass keypass

  -keystore tclient.keystore

- **tclient.keystore** generated here needs to be copied wherever the controller is installed.

- Please note when creating the keys, make sure to keep the value of **keypass** and **storepassword** the same.

Controller side:

keytool -genkey

-dname "CN=FULL_NAME, OU=UNIT, O=ORG, L=LOCALITY, ST=STATE, C=US"

-alias "ALIAS_NAME"

-keystore "kclient.keystore"

-storepass storepassword

-keypass keypass

-keyalg RSA

-sigalg SHA1withRSA

keytool -export

-alias "ALIAS_NAME"

-keystore kclient.keystore

-storepass storepassword

-keypass keypass

-rfc -file kclient.crt

keytool -import

-alias "ALIAS_NAME"

-file kclient.crt

-storepass storepassword

-keypass keypass

-keystore tserver.keystore

- Please note when creating the keys, make sure to keep the value of **keypass** and **storepassword** the same.

Modify /<InstallationPath>/Dataguise/DgSecure/IDPs/HDFSIDP/jetty-embedded.properties

keyStorePath = /<InstallationPath>/kserver.keystore

keyManagerPassword = encrypted form of keypass of kserver.keystore

keyStorePassword = encrypted form of storepassword of kserver.keystore

trustStorePath = /path/to/tserver.keystore

trustStorePassword = encrypted form of storepassword of tserver.keystore

needClientAuth = Y

sslEnabled = Y

****Note:**

All the keystore passwords used for creating these keystores files needs to be encrypted using the following dgcl command and then encrypted passwords are placed under the properties file under controller and IDP accordingly.

- Command for encrypting password in dgcl is

  encrypt password "passwordString";

  //"passwordString" means any password that needs to be encrypted

Controller Changes:

- The **kclient.keystore** and **tclient.keystore** files created at the IDP level needs to be copied wherever the controller is installed.

- The location of these keystore files and their respective passwords need to be specified under \<InstallationPath>\tomcat9\webapps\dgcontroller\WEB-INF\classes\dgController.properties file under following parameter

  serverKeyStoreFileLocation = /path/to/kclient.keystore

  keyStorepasswordServer= encrypted form of kclient.keystore's storepassword

  clientKeyStoreFileLocation = /path/to/tclient.keystore

  keyStorepasswordClient= encrypted form of tclient.keystore's storepassword

- Restart tomcat.
- Login to DgSecure Admin and enable SSL under SSL Setting tab for HDFS IDP.

- Confirm the IDP is working by doing Test Connection on that IDP under the IDP tab in DgSecure Admin.

Client Side Configuration:

There are two ways a client could be configured for SSL:

1. **With CA (Certification Authority) Authentication**

   Client authenticates the certificate sent by the server by matching the CA that have signed the certificate with the list of CAs available at client side.

****Note:**

In DgSecure, if user needs to use self-signed certificates (which is not recommended due to security risks), then such certificates should be imported on the client side trust store (JVM Default trust store) at following location:

$JAVA_HOME/JRE/lib/security/cacerts or

$JRE_HOME/lib/security/cacerts

We can use following keytool command to import certificate in JVM trust store.

keytool -import -alias <alias name> -keystore <keystorelocation>/cacerts -file <certificate file>

2. **Without CA Authentication**

   No client side CA authentication takes place. This type of communication is not recommended since it could cause MITM attack. Because, attacker can impersonate Client or Server in this case since no CA authentication is happening.

   ****Note:**

   It is recommended to update JVM to its latest version because of updates in cypher algorithms and some updates in security feature happened in recent times due to advent of attacks like DROWN. Also, it is essential due to some conflicts between encryption algorithms that are used in old java version and updated algorithms used on server side.

Appendix I: Enabling Spark in the HDFS IDP

The Hadoop Data IDP (HDFS IDP) now works with Spark in Spark-enabled clusters for Detection. Protection is not enabled to use Spark in 6.5.0 for Spark-enabled clusters. Both Detection and Protection are supported as before, with MapReduce, whether the cluster is Spark-enabled or not.

To enable Spark integration, the following steps are to be performed.

At Install Time

Step 1: Installer asks for location where HDFS IDP has to be installed:

```
------------------------------------------------------------------------
Please specify the directory where DgSecureHDFSAgent will be installed.

Installation Directory [/opt/Dataguise]:
```

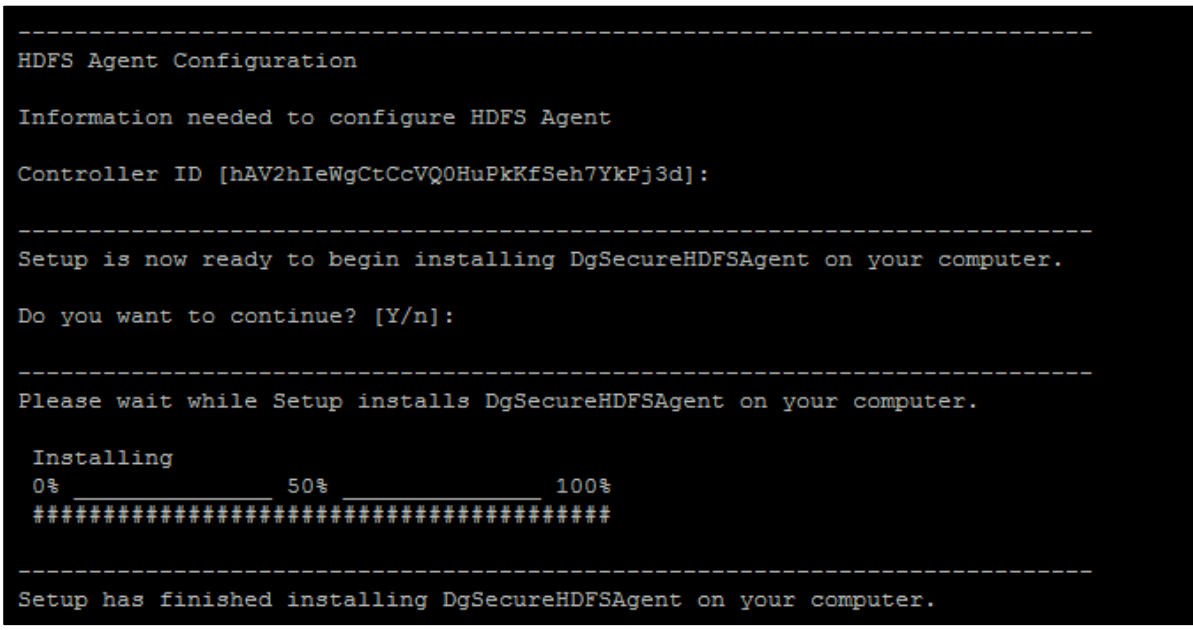
Step 2: Select Option 6 ("Spark") shown below

```
------------------------------------------------------------------------
HDFS Agent Selection

Select option to deploy HDFS Agent

[1] MapR
[2] Cloudera
[3] Hortonworks
[4] Pivotal
[5] EMR
[6] Spark
[7] Local
Please choose an option [1] : 6
```

Step 3: Select Hortonworks or EMR options as below

```
Spark

[1] Spark 2.0.2 HW-2.4: HDFS Agent compatible with Spark-2.0.2 HW-2.4 will be deployed for use by DgSecure
[2] Spark 2.0.2 EMR/S3: HDFS Agent compatible with Spark-2.0.2 will be deployed for use by DgSecure.
Please choose an option [1] : 2
```

Step 4: Proceed as normal with installation



In the HDFSIDPConfig.properties File:

For Spark with Hortonworks, the **distro** property will be set to **spark** by the installer.

For EMR, the **distro** property will be set to **EMR** and the **s3filesystem** property to **s3a.** If the user needs to use EMR for Hadoop HDFS instead of for S3 processing, the **s3filesystem** property needs to be set to **hdfs,** and the IDP needs to be restarted.

In the jetty-embedded.properties File (Hortonworks only)

The appropriate Hortonworks version number needs to be set, as follows (particular version # below is an example.)

-Dspark.driver.extraJavaOptions=-Dhdp.version=2.4.2.0-258

Appendix J: Single Sign On and Single Sign Out

DgSecure supports Single Sign-On using SAML. To configure Single Sign-On and Single Sign-Out please perform the following steps:

1. Verify that the property “pathtoSSLCert” exists in dgController.properties file under below mentioned path:-

   [Installed Directory]/Dataguise/DgSecure/tomcat8/webapps/dgcontroller/WEB-INF/classes/dgController.properties

   If not, then add below mentioned property in dgController.properties

   pathtoSSLCert= [Installed Directory]/Dataguise/DgSecure/DgCertificate/DgTestCertificate.cer

   Certificate Path:

   [Installed Directory]/Dataguise/DgSecure/DgCertificate/DgTestCertificate.cer

2. Verify DgCertificate and Generate Private Key

   Path for Certificate - [Installed Directory]/Dataguise/DgSecure/DgCertificate/Commands

   Path for Verify JKS file and get the value of “keyAlias”, ‘keystorePass’ &’ keyPass’

   cat /[Installed Directory]/Dataguise/DgSecure/tomcat8/conf/server.xml

   <Connector port="10182" protocol="org.apache.coyote.http11.Http11NioProtocol"

   maxThreads="150" SSLEnabled="true" scheme="https" secure="true"

   clientAuth="false" sslEnabledProtocols="TLSv1.2"

   keystoreFile="/[Installed Directory]/Dataguise/DgSecure/DgCertificate/DgTestCertificate.jks" keystorePass="dataguise" keyPass="dataguise" keyAlias="dataguise"

   ciphers="TLS_ECDHE_RSA_WITH_AES_128_CBC_SHA256,TLS_ECDHE_RSA_WITH_AES_128_CBC_SHA,TLS_ECDHE_RSA_WITH_AES_256_CBC_SHA384,TLS_ECDHE_RSA_WITH_AES_256_CBC_SHA,TLS_RSA_WITH_AES_128_CBC_SHA256,TLS_RSA_WITH_AES_128_CBC_SHA,TLS_RSA_WITH_AES_256_CBC_SHA256,TLS_RSA_WITH_AES_256_CBC_SHA" server="Server" />

Go to below path

/[Installed Directory]/**Dataguise/DgSecure/DgCertificate** and execute below mentioned command:-

Sample Command:-

keytool -importkeystore -srckeystore keystore.jks -destkeystore keystore.p12 -deststoretype PKCS12 -srcalias dataguise -srcstorepass dataguise -srckeypass dataguise -deststorepass dataguise -destkeypass dataguise

Working Command:-

keytool -importkeystore -srckeystore DgTestCertificate.jks -destkeystore keystore.p12 -deststoretype PKCS12 -srcalias dataguise -srcstorepass dataguise -srckeypass dataguise -deststorepass dataguise -destkeypass dataguise

Please verify the private key with below command:-

openssl pkcs12 -in keystore.p12 -nocerts -nodes -out private.key

Enter Import Password:

MAC verified OK

Delete the highlighted content from the private key and save it

[ec2-user@ip-10-141-240-77 DgCertificate]$ cat private.key

Bag Attributes

friendlyName: dataguise

localKeyID: 54 69 6D 65 20 31 35 30 39 35 33 30 37 31 30 35 33 37

Key Attributes: <No Attributes>

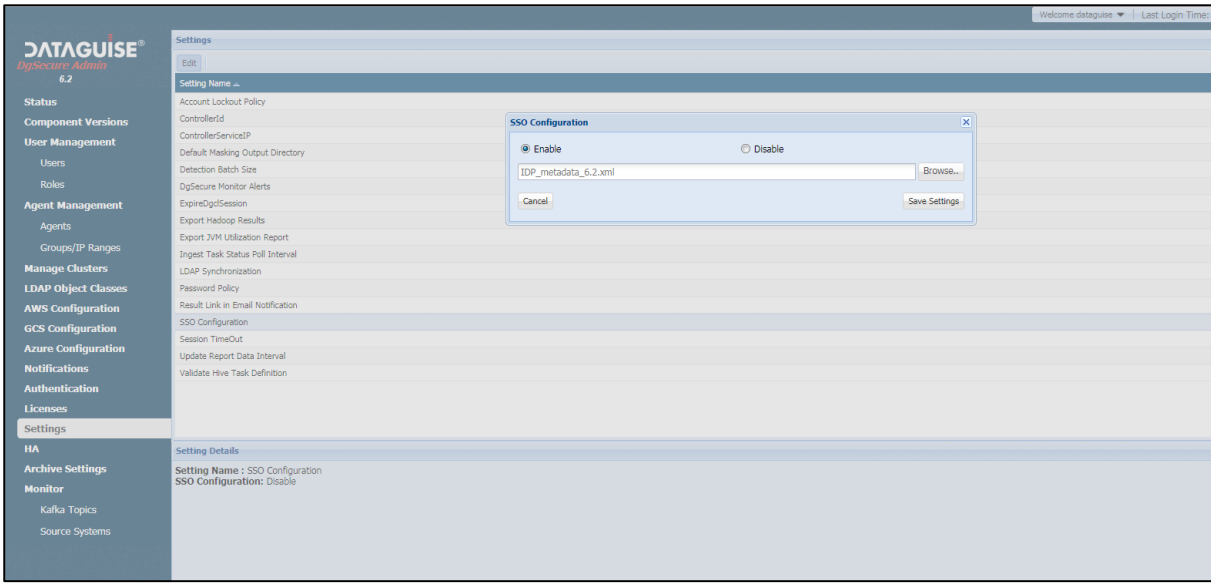
-----BEGIN PRIVATE KEY-----

--key content----

-----END PRIVATE KEY-----

Steps to configure from DgSecure Admin:

- Go to DgSecure Admin Settings
- Select SSO
- Click enable
- Browse path to the Open AM metadata xml



- Click save settings.
  Please contact Dataguise Support or Professional Services for more details on configuring SAML-based SSO with DgSecure 7.2.0.

****Note:** If using email as authentication type, make sure the property basic is not set to uid.

Appendix K: Create a Temporary Directory

A parameter "sys_temp_dir" is available, which can be used to set temporary directory where installer can write, create and execute files during installation, the syntax is:

<InstallerName>--sys_temp_dir "<Absolute_path_of_Custom_temp_folder>"

For installation on Linux environment minimum permission requirement is 755 for the folder -"<Absolute_path_of_Custom_temp_folder>".

Appendix L: Key Management Options

DgSecure for Hadoop encryption is compatible with a variety of key management systems. One option is to use the Java keystore which installs with DgSecure. Since the keystore installs with DgSecure, it does not require any configuration. Another option is to use a third-party keystore that supports the Key Management Interoperability Protocol (KMIP). Currently, DgSecure encryption can run using either Safenet or RSA as the key management system. Other KMIP-compliant KMSs can be integrated based on customer request.

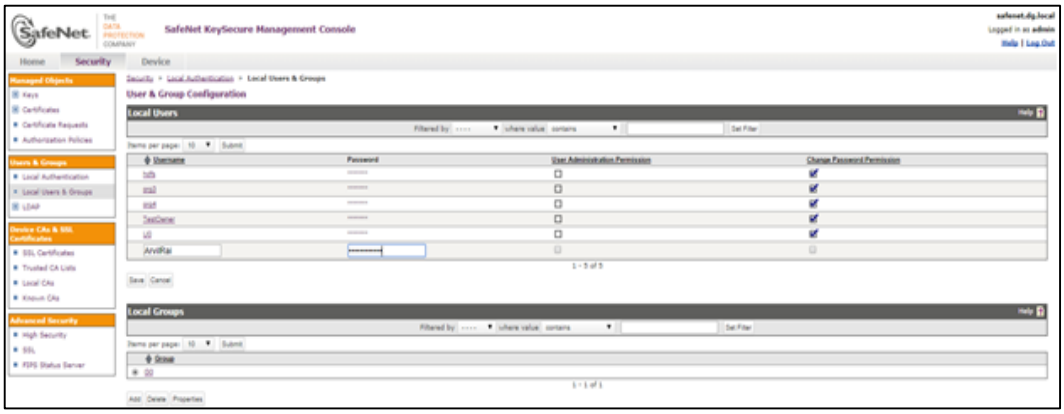
SAFENET

These instructions illustrate how to configure the Safenet JCE provider on both the server and client. Original documentation from Safenet is found in chapter 10 of Safenet's "ProtectApp-JCE, Version 6.6.0."

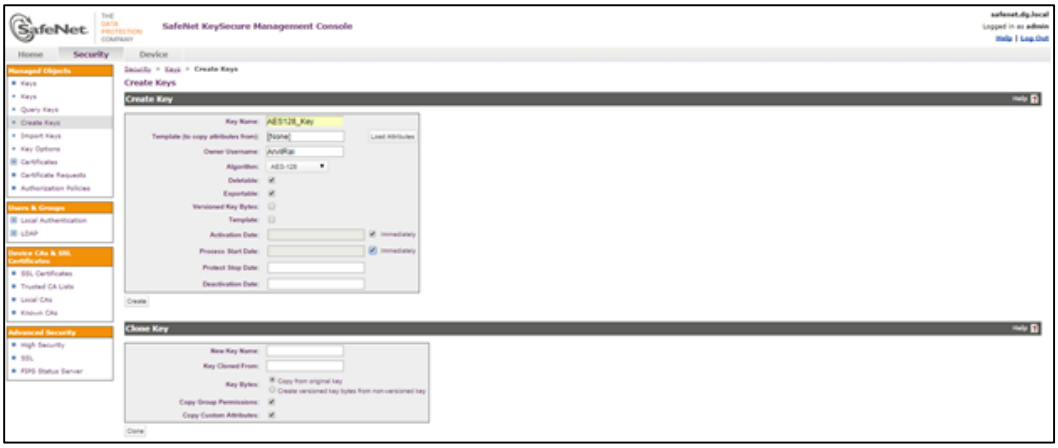
Server-Side Configuration

Step 1: Creating a User & Keys

- Create User

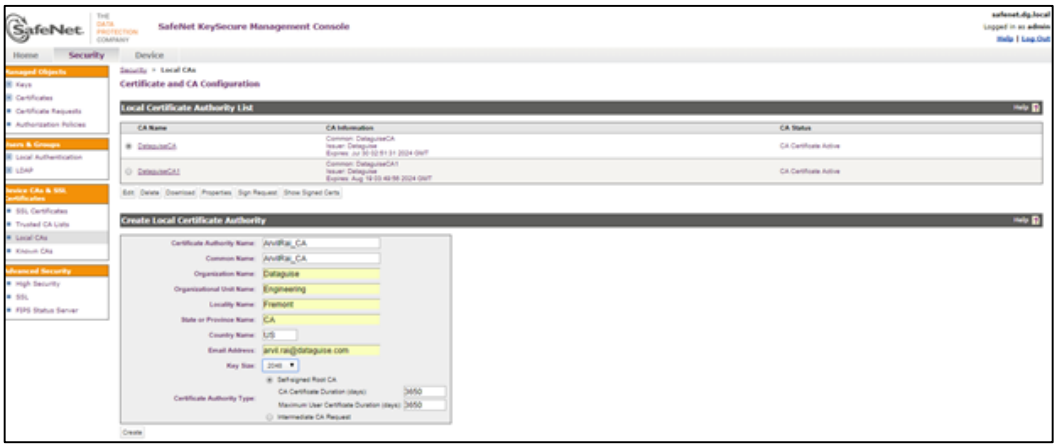


- Create Keys [AES128_Key, AES256_Key]. Associate keys with the owner / newly created user.

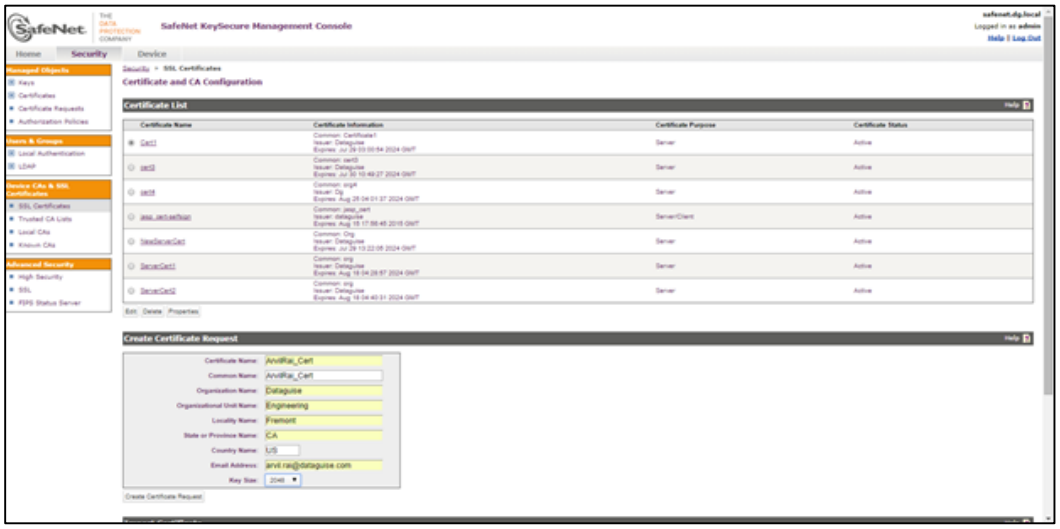


Step 2 Create a Local Certificate Authority (CA)

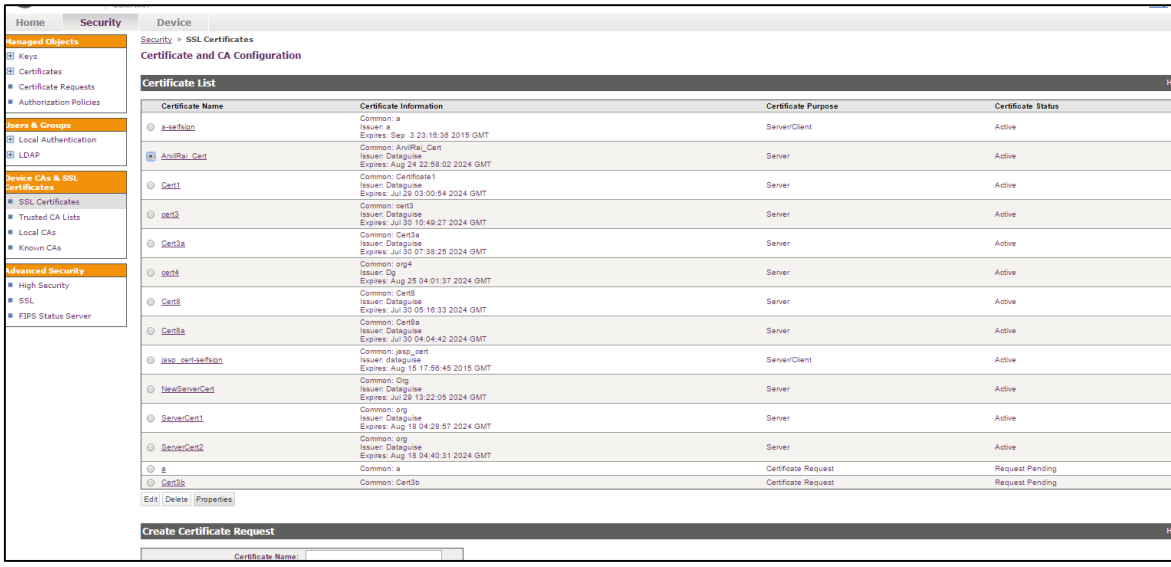
1. Navigate to the Create Local Certificate Authority section (Security, Certificates & CAs, Local CAs). Enter the values shown below to create a new local CA. Click Create.



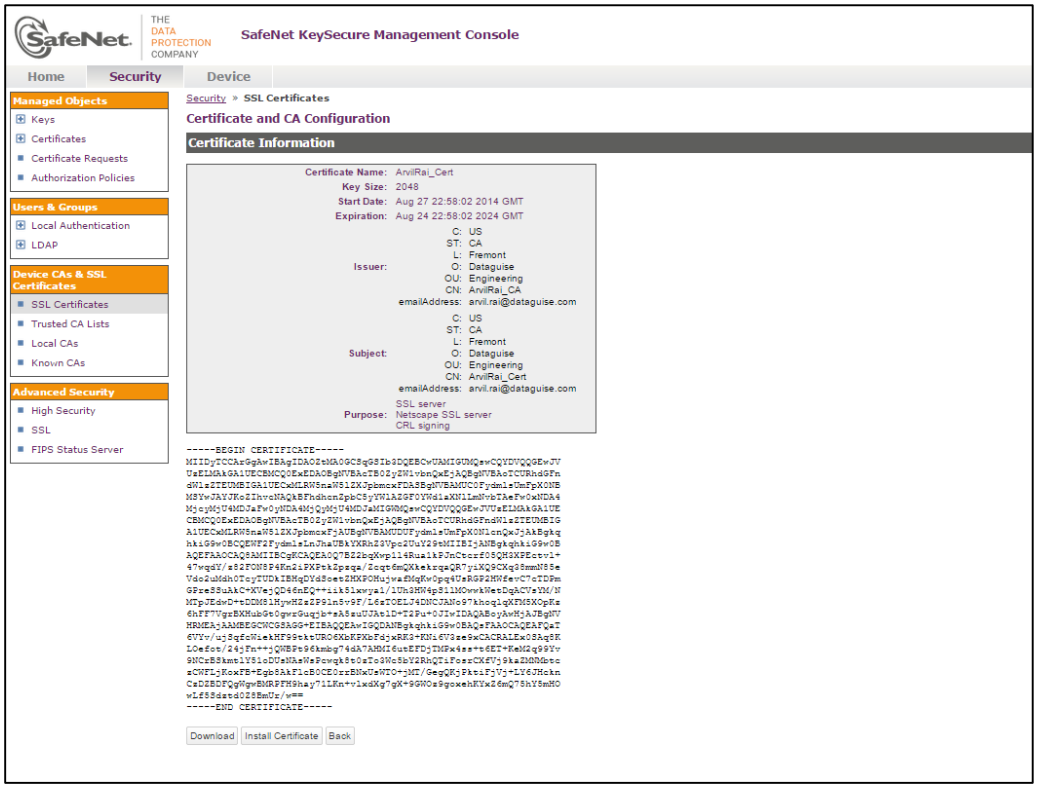
2. Navigate to the *Create Certificate Request* section (Security, Certificates & CAs, Certificates). Enter the values shown below to create a request. Click Create Certificate Request.



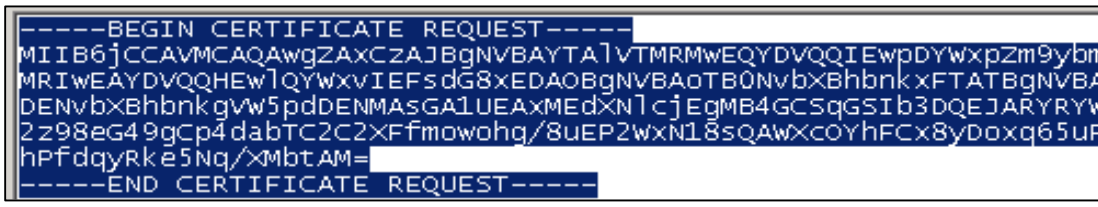
3. Select the new certificate request from the *Certificate List* section (located above the *Create Certificate Request* section).



4. Click **Properties**.

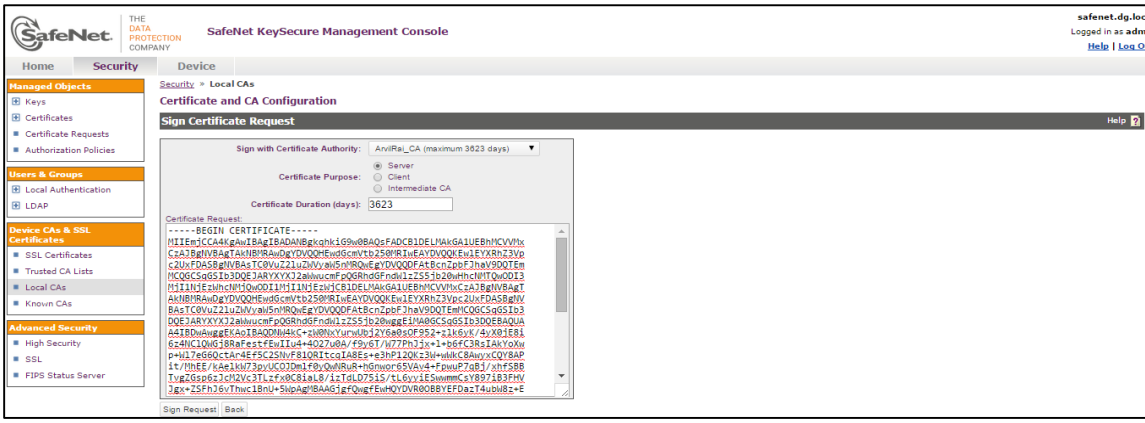


5. Copy the actual request (example below). Include the header and footer.



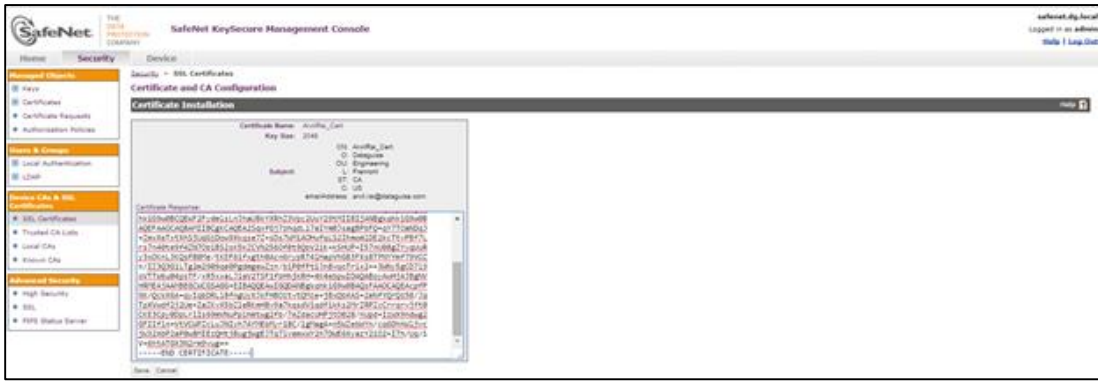
6. Navigate back to the *Local Certificate Authority List* section (Security, Certificates & CAs, Local CAs). Select your new local CA and click **Sign Request**.

7. Select **Server** as the *Certificate Purpose* and paste the certificate request into the **Certificate Request** field.



8. Click **Sign Request**. This will take you to the *CA Certificate Information* section. Copy the actual certificate (at bottom). Include the header and footer.

9. Navigate back to the *Certificate List* section (Security, Certificates & CAs, Certificates). Select your certificate request and click **Properties**. Click Install Certificate. Paste the actual certificate, as shown below. Click **Save**.



10. Navigate to the *KMIP Server Settings* section (Device, NAE Server, KMIP). Click **Edit**.

11. Check *Use SSL* and select your new server certificate in the **Server Certificate** field. Click **Save.**

12. Navigate back to the Local Certificate Authority List section (Security, Certificates & CAs, Local CAs). Select your new CA and click **Download**. Place the CA certificate on your client.

13. Move the certificate from the download location to *Java_Home>/lib/security.*
14. Open a command prompt on your client, navigate to *Java_Home>/lib/security.*
15. Install the CA certificate into the cacerts keystore using the command below. Follow the prompts as shown.

    keytool -keystore cacerts -import -alias NewLocalCA -file NewLocalCA.crt.cer Enter keystore password: changeit
    ...
    Trust this certificate?[no]: yes
    Certificate was added to keystore

****Note:**

The value for the -file option must reflect your actual filename. The keystore password must reflect your actual keystore password.

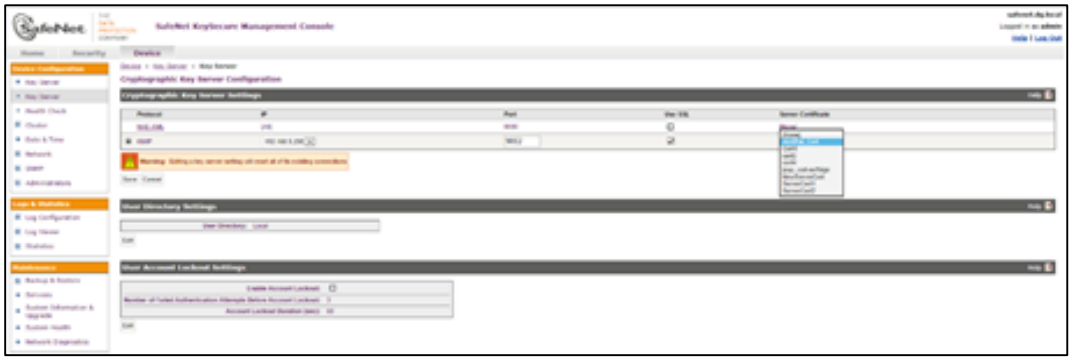
16. Update the following parameters in your IngrianNAE.properties file:

- Protocol=ssl

- Key_Store_Location=< <path to Java Home>/lib/security/cacerts.

- Key_Store_Password=<password>

-Username=<username>



Client-Side Configuration

Step 1: Create a Client Certificate

1. On the client, open a command prompt and navigate to <Java_Home>/lib/security.
2. From the command line, create a new client keystore to use with DgSecure. ArvilRai_keystore]
3. Command: keytool -keystore ArvilRai_keystore_1 -genkey -alias ArvilRai_alias_1 -keyalg RSA

   Enter keystore password: Dataguise123
   What is your first and last name?
   [Unknown]: Arvil Rai
   What is the name of your organizational unit?
   [Unknown]: ArvilRai
   What is the name of your organization?
   [Unknown]: Dataguise
   What is the name of your City or Locality?
   [Unknown]: Fremont
   What is the name of your State or Province?
   [Unknown]: CA
   What is the two-letter country code for this unit?
   [Unknown]: US
   Is CN= Arvil Rai, OU= ArvilRai, O= Dataguise, L= Fremont, ST= CA, C=US correct?
   [No]: yes
   Enter key password for <ccert> Press 'Enter' without entering any password

4. Generate a client certificate request using the public/private key that was created in your new keystore.

   Command: keytool -keystore ArvilRai_keystore_1-certreq -alias ArvilRai_alias_1 -file ArvilRai_alias_1.csr

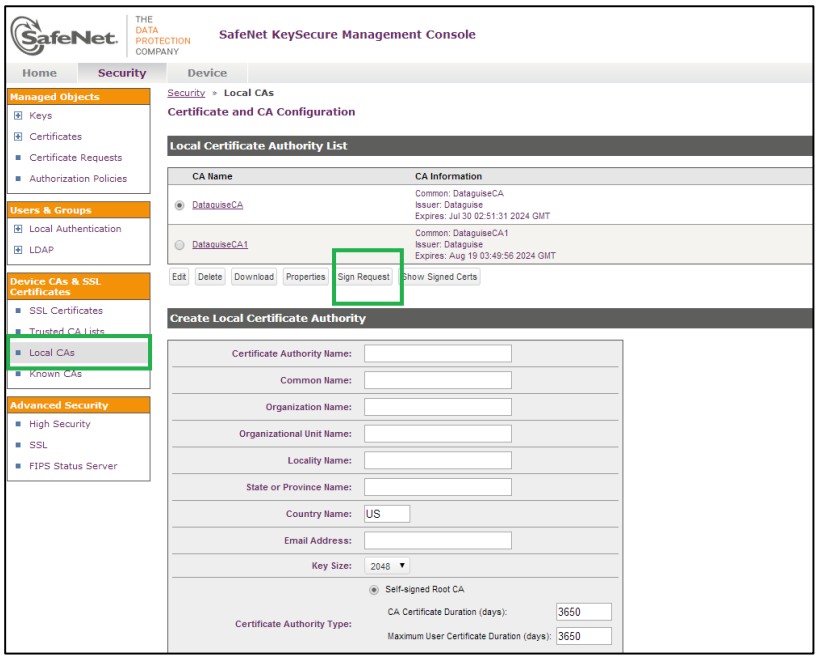
   Enter keystore password: <password>

5. Open the client certificate request file and copy the actual request. Include the header and footer. The certificate is created in <Java_Home>/lib/security.

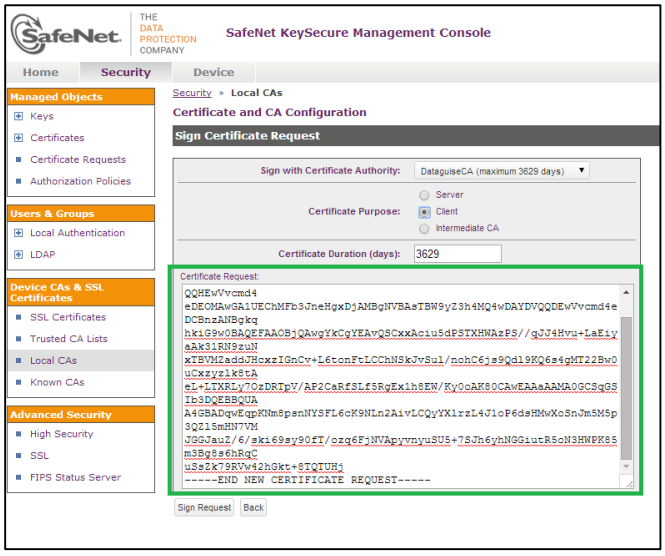
Command: vi ccert.csr

```
-----BEGIN CERTIFICATE REQUEST-----
MIIB6jCCAVMCAQAwgZAxCzAJBgNVBAYTAlVTMRMwEQYDVQQIEwpDYWxpZm9ybmlh
MRIwEAYDVQQHEwlQYWxvIEFsdG8xEDAOBgNVBAoTB0NvbXBhbnkxFTATBgNVBAsT
DENvbXBhbnkgVW5pdDENMAsGA1UEAxMEdXNlcjEgMB4GCSqGSIb3DQEJARYRYWRk
2z98eG49gCp4dabTC2C2XFfmowohg/8uEP2wxN18sQAWxcOYhFCx8yDoxq65uFcb
hPfdqyRke5Nq/XMbtAM=
-----END CERTIFICATE REQUEST-----
```

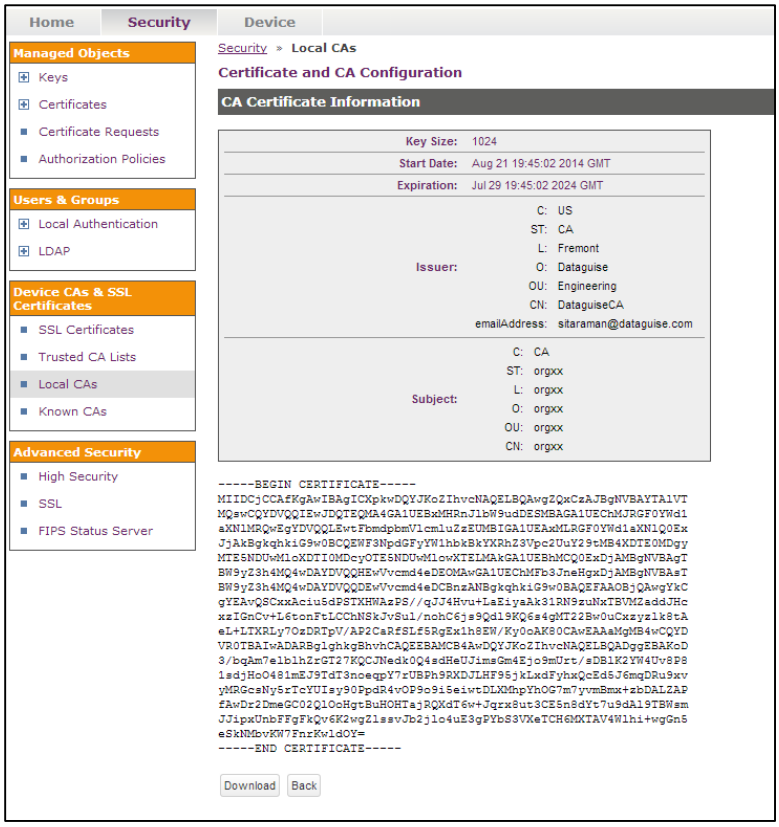
6. Log in to the Management Console on the server machine as an administrator with Certificate Authority and NAE Server and Navigate to the Local Certificate Authority List section (Security, Certificates & CA's, and Local CAs). Select NewLocalCA and click **Sign Request**. (NewLocalCA is the CA created in step 2.)



7. Select **Certificate Purpose** Client and paste the certificate request into the **Certificate Request** field, as shown below.



8. Click **Sign Request**. This will take you to the CA Certificate Information section.

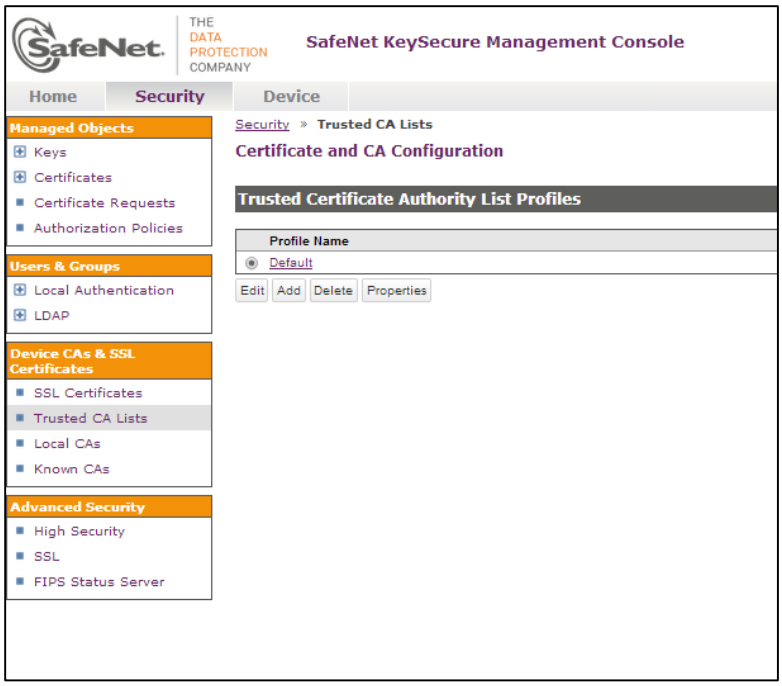


9. Your Client Certificate is now created.

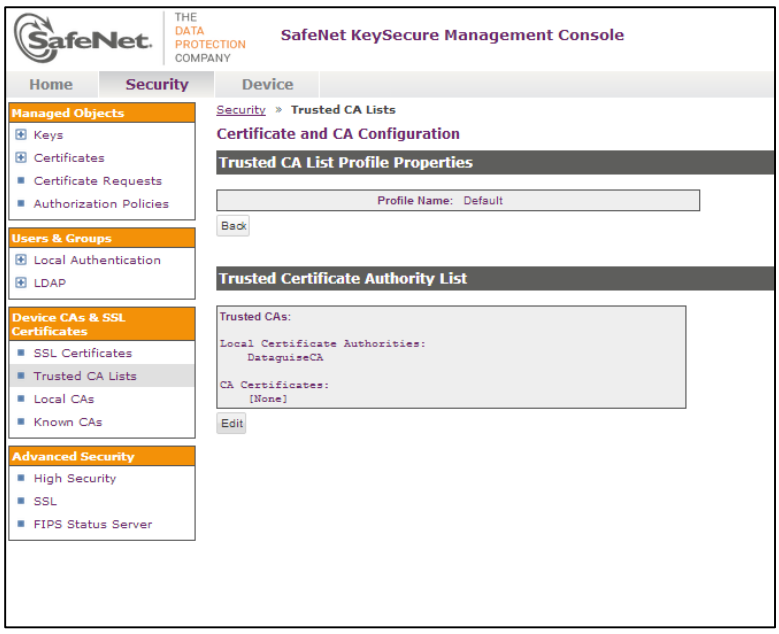
Step 2: Download & Import Client Certificate

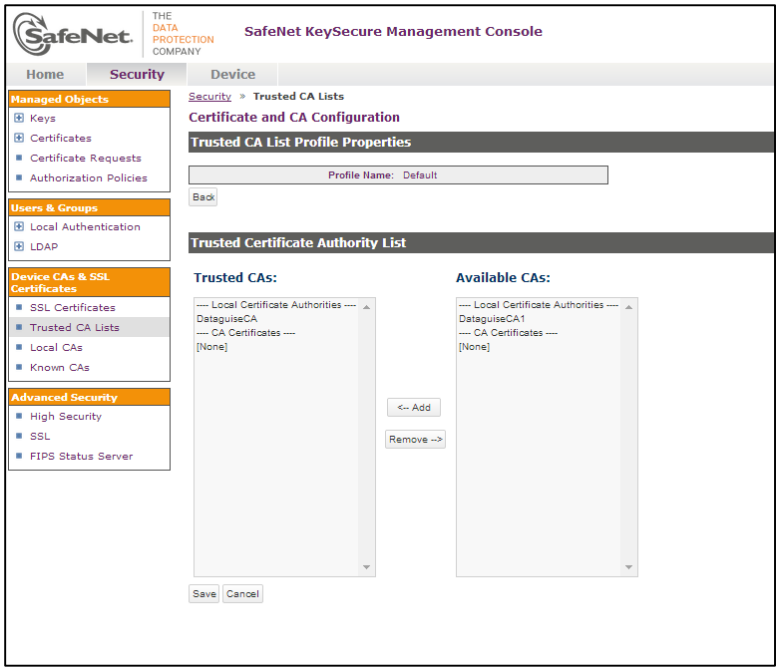
1. On the client machine, click Download to download your new client certificate (signed.crt) to your client

2. Return to the Local Certificate Authority List section (click Back on the CA Certificate Information section).
3. Select NewLocal CA and click Download
4. Navigate to the Trusted Certificate Authority List Profiles section (Security, Certificates & CAs, Trusted CA Lists). Select Profile Name Default and click Properties



5. Click **Edit** in the Trusted Certificate Authority List section. Add the CA to the Trusted CA list.





6. On the client, move the client certificate and the CA certificate from the download location to <Java_Home>/lib/security.

7. Open a command prompt on your client and navigate to <Java_Home>/lib/security

8. Import the CA certificate (NewLocalCA) into the client keystore.

   $ keytool -keystore clientcerts -import -alias NewLocalCA -file NewLocalCA.crt
   Enter keystore password:
   Owner: Issuer:
   Serial number: 0
   Valid from: mm/dd/yy hh:mm until: mm/dd/yy hh:mm
   Certificate fingerprints:
   MD5: F0:2D:2F:ED:55:31:6F:F0:A6:E4:AA:37:1F:83:E7:FA
   SHA1:8A:08:61:AB:73:32:E2:18:0E:B7:8D:69:2E:91:A6:24
   Trust this certificate? [no]: yes
   Certificate was added to keystore

9. Import the signed client certificate into the client keystore file

   Command: keytool -keystore ArvilRai_keystore_1 -import -alias ArvilRai_alias_1 -file signed.crt
   Enter keystore password
   Certificate reply was installed in keystore

10. Verify that the certificates are correctly installed. You will need to see a certificate chain length of 2, and two certificates: the client certificate and the CA.

Command: keytool -keystore ArvilRai_keystore_1 -alias ArvilRai_alias_1 -list –v
Enter keystore password:
Alias name: ccert
Creation date: Oct 19, 2006
Entry type: keyEntry
Certificate chain length: 2
Certificate[1]:
Owner: CN=clientcerts, OU=clientcerts, O=clientcerts, L=clientcerts, ST=clientcerts, C=US Issuer: EMAILADDRESS=NewLocalCA@NewLocalCA.com, CN=NewLocalCA, OU=NewLocalCA, O=NewLocalCA, L=NewLocalCA, ST=NewLocalCA, C=US
Serial number: 17
Valid from: 10/18/06 9:11 AM until: 10/7/16 9:11 AM
Certificate fingerprints:
MD5: 7B:73:86:91:A6:E7:6C:60:0C:28:FA:E2:AF:03:0A:ED
SHA1: 26:63:4C:59:EB:80:A9:16:C9:DB:E4:D4:1D:C0:1A:BD:F3
Certificate[2]:
Owner: EMAILADDRESS=NewLocalCA@NewLocalCA.com, CN=NewLocalCA, OU=NewLocalCA, O=NewLocalCA, L=NewLocalCA, ST=NewLocalCA, C=US Issuer: EMAILADDRESS=NewLocalCA@NewLocalCA.com, CN=NewLocalCA, OU=NewLocalCA, O=NewLocalCA, L=NewLocalCA, ST=NewLocalCA, C=US
Serial number: 0
Valid from: 10/9/06 5:13 PM until: 10/7/16 5:13 PM
Certificate fingerprints:
MD5: F0:2D:2F:ED:55:31:6F:F0:A6:E4:AA:37:1F:83:E7:FA
SHA1: 8A:08:61:AB:73:32:E2:18:0E:B7:8D:69:2E:91:A6:24:BC

11. Update the following parameters in the IngrianNAE.properties file:

- Key_Store_Location/usr/lib/jvm/java-1.6.0-openjdk-1.6.0.0.x86_64/jre/lib/ext/orgx
- Key_Store_Password=changeit

- Client_Cert_Alias=ccert3
- Client_Cert_Passphrase=

****Note:** The Client_Cert_Passphrase parameter should be set to no value.

12. Return to the Management Console and navigate to the KMIP Server Authentication Settings section

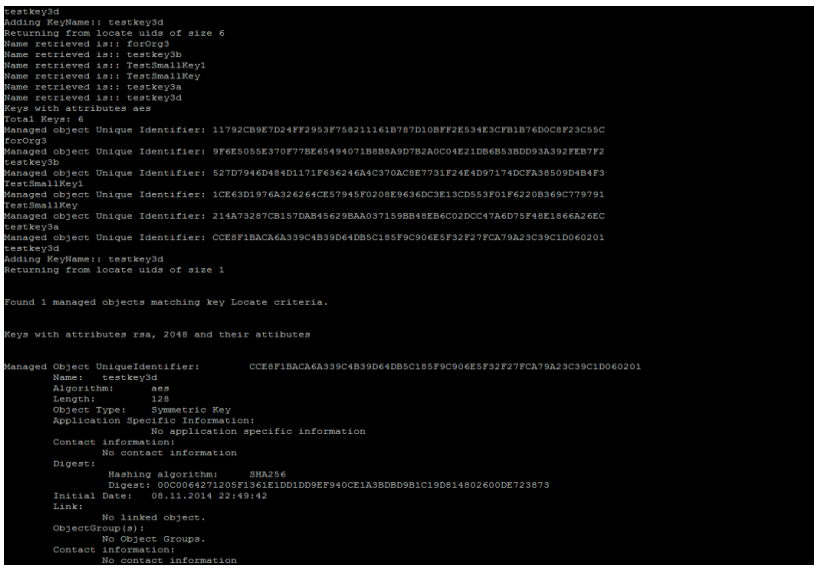
    (Device Management, NAE Server) and enter the following values:

    - Client Certificate Authentication: Used for SSL Session only
    - Trusted CA List Profile: Default
    - The CA that signed the certificate must be a member of the Trusted CA List Profile.

13. Be sure to update the HDFSIDPConfig.properties file with Safenet's NAE.Properties Location. For more information, refer to Section 3.1.1 *HDFS IDP*.

Step 3: Verify Safenet JCE Provider Configuration

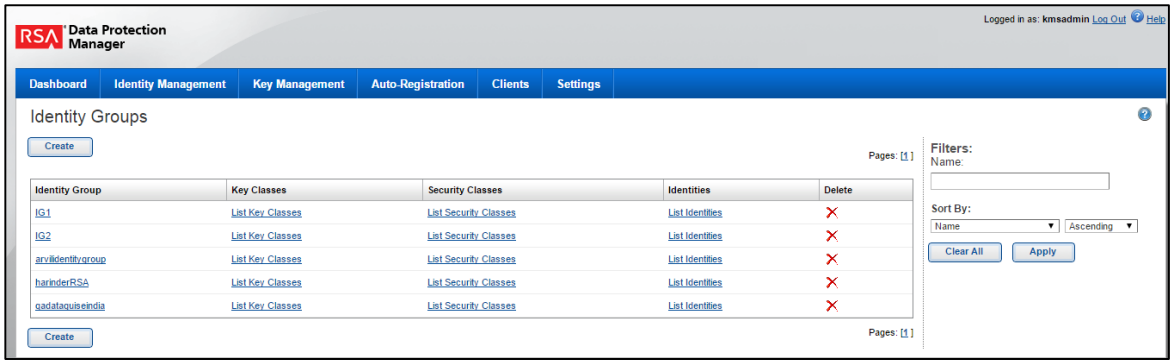
1. Create and Execute an encryption task from DgSecure.
2. View catalina.out log file.
3. If you find exceptions related to KMIP, this means the client has not been configured correctly. Verify each of the above steps again.
4. If the client was configured successfully, something similar to what is shown below should be seen in the log file.



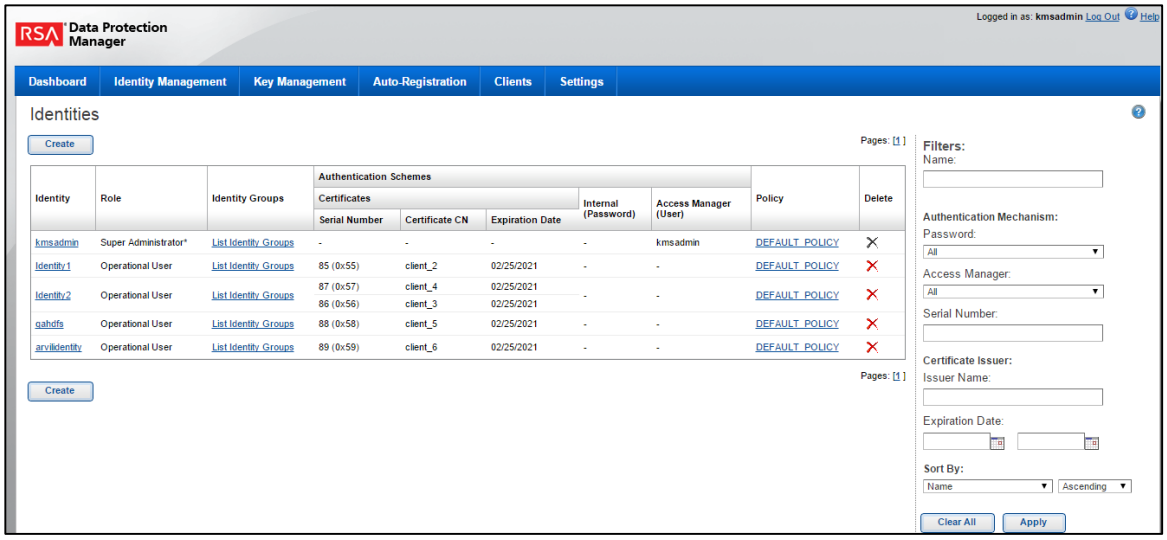
Pre-Requisites

Ensure you have all your client, server and root PKI certificates you obtained from your certificate vendor and they are loaded on to DPM via the Appliance Console GUI.

1. Go to IdentityManagement -> Identity Groups ->Create. Create a new Identity Group. Click Save.

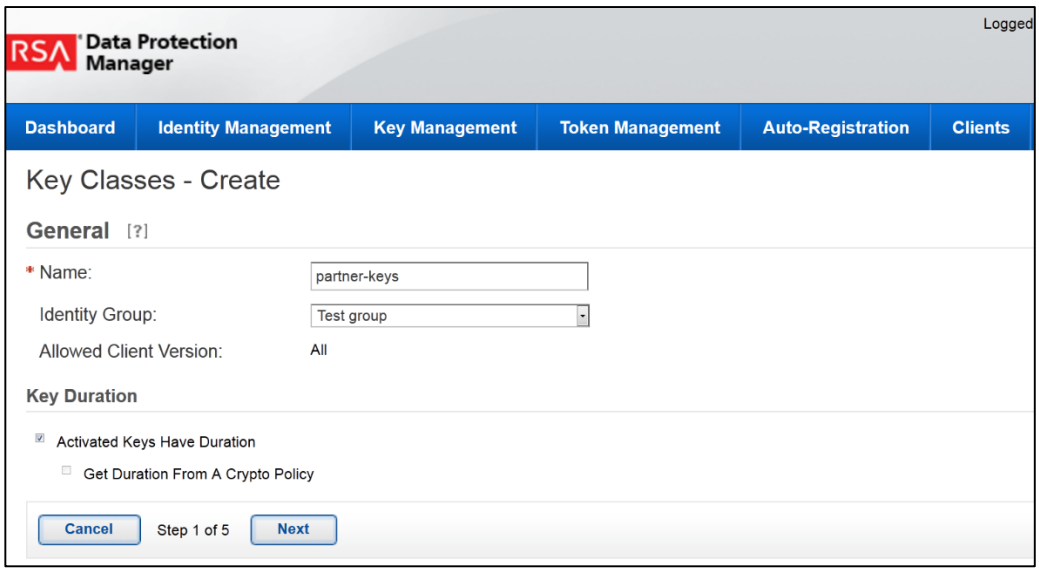


2. Go to Identity Management -> Identities -> Create. Create a new Identity, associating it with the recently created Identity Group. Select "Operational User" as the role and ".cer" as the file extension. Place the corresponding .p12 file in the location as specified in HDFSIDPConfig.properties entry (discussed in the DgSecure Configuratio section). Leave the internal and Access manager fields blank. Click **Save**.

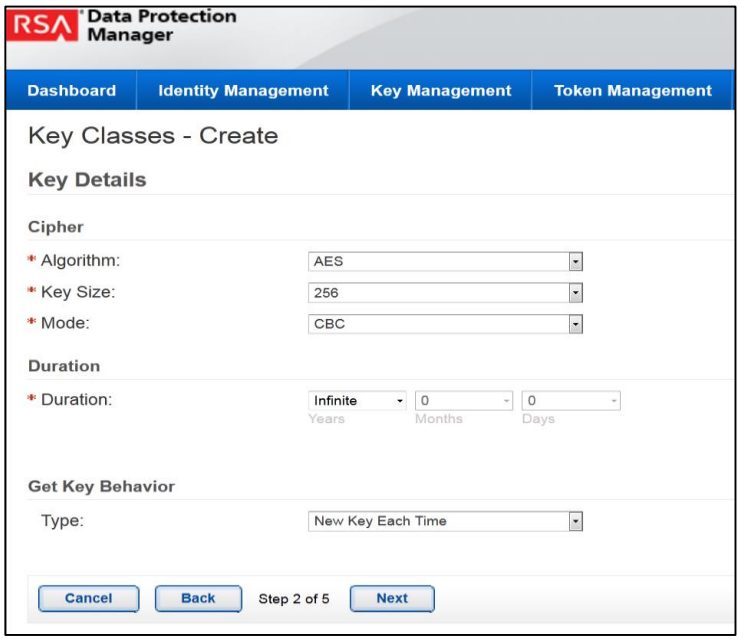


3. Go to Key Management -> Key Classes. Name a Key Class and associate it with the recently created Identity Group. Click the Activated Keys

Have Duration check box. Leave Get Duration from a Crypto Policy unchecked. Click Next.

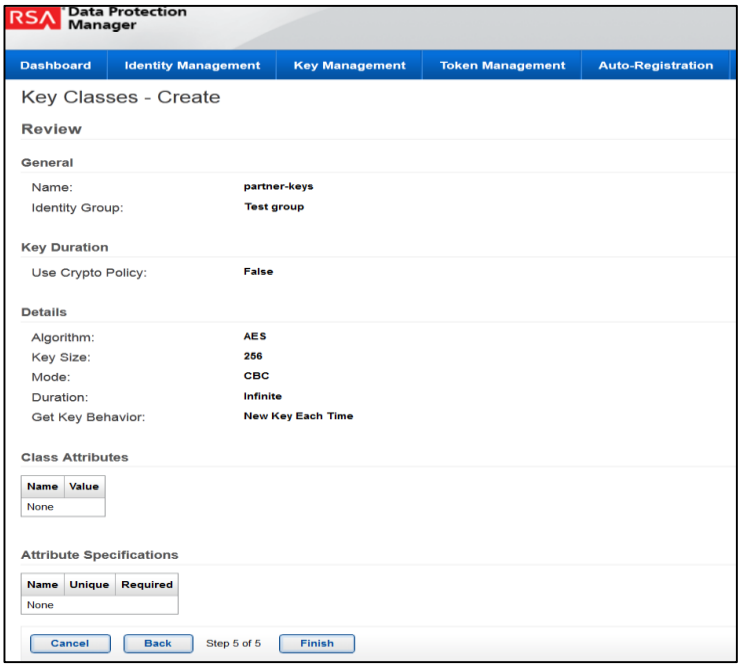


4. Select Algorithm AES, Keysize 128 or 256, and mode CBC.



- Algorithm: AES
- Key Size: 128 or 256
- Mode: CBC
- Duration: User discretion
- Key Behavior: Use Current Key
- Check Allow: Auto-generation box

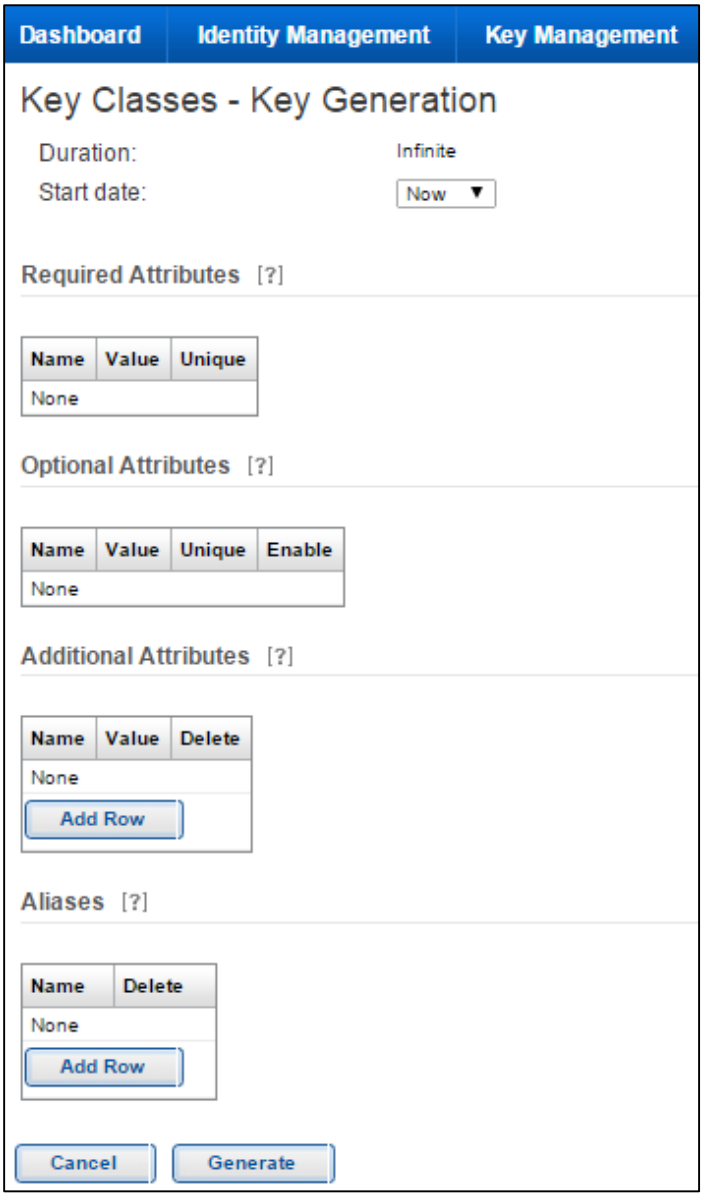
- Click the **Next** button and enter any attributes if desired untill you get to step 5 for review



5. Click Finish.
6. Go to KeyManagement -> KeyClasses. Against the key class for which you want to generate a key, click the yellow key icon in the column titled "Generate key".

   Leave the attributes blank.

   Click **Add Row** and enter an alias. Click Generate.



7. To use this key, you must have the corresponding Identity in

   `client.identity_name=Identity2`

   in rsaconfig.properties. Have the corresponding .p12 (corresponding to the .cer file that you added earlier while creating the identity in RSA web console) in your local client configuration in the appropriate place as mentioned in rsaconfig.properties.

   Add the just created alias in the list of aliases in HDFSIDPConfig.properties (see below):

   ```
   rsa.key.aliases=KeyAlias1420533992581,KeyAlias1419064002219,Ke
   yAlias1419063766020,KeyAlias1419063661538,KeyAlias141906237154
   3,Sitkey256
   ```

DgSecure Configuration

****Note**: Configuring DgSecure to use the RSA Key Manager occurs after DgSecure has been successfully installed.

In order to follow the configuration instructions below, ensure that you have unlimited strength policy jurisdiction files in the java_home/jre/lib/security folder of tomcat's JAVA_HOME.

rsaconfig.properties Configuration

Place rsaconfig.properties file with the following contents in the location as specified in HDFSIDPConfig.properties (#RSA config properties file location rsa.config.props.location=/$JAVA_HOME/jre/lib/software/jdk1.7.0_60/jre/lib/ext/rsaconfig.properties).

Use values as appropriate for your configuration for the items in red font.

```
#rsaconfig.properties file

server.host=192.168.5.31

validate.hostname=false

protect_with_deactivated_keys=false

pki.client_keystore_file=/$JAVA_HOME/jre/lib/software/jdk1.7.0_60/jre/lib/ext/client_4.p12 #for convenience this location is the same as where you place your rsaconfig.properties file.

server.retry_delay=5000

client.registration_file=/$JAVA_HOME/jre/lib/software/jdk1.7.0_60/jre/lib/ext/client.reg #for convenience this location is the same as where you place your rsaconfig.properties file.

pki.client_keystore_expiry=15

cache.mode=DiskAndMemory

client.lockbox=false

client.actmgmt_enable=true

client.app_name=name18

client.identity_name=Identity2

pki.client_keystore_password=Password1

cache.max_time_to_live=7200

server.tls_version=TLSv1

server.connect_timeout=10000

server.read_timeout=5000

cache.write_delay=30

cache.file=/$JAVA_HOME/jre/lib//software/jdk1.7.0_60/jre/lib/ext/keycache.kmc #for convenience this location is the same as where you place your rsaconfig.properties file.

high.availability=false
```

```
secure_random.general=HMACDRBG256

client.actmgmt_poll_interval=20 m

server.request_retries=3

secure_random.iv=HMACDRBG256

client.registration=false

client.auto_update_certificate=true

server.port=443

pki.server_keystore_file=/$JAVA_HOME/jre/lib//software/jdk1.7.0_60/
jre/lib/ext/cacert.pem #for convenience this location is the same as
where you place your rsaconfig.properties file.

client.origin_info.optional_in_ciphertext=false

cache.max_keys=100
```

Place the files given by RSA (the files in the folder RSAFiles) in the location as appropriate to your configuration (replace the rsaconfig.properties with the one you created above). If you have used the default location as mentioned above, they will be placed in the same location as rsaconfig.properties.

HDFSIDPConfig.properties Configuration

Add the following entries, as appropriate, for your configuration above to HDFSIDPConfig.properties, which is located in:

<DGSecure_Install_Directory>/Dataguise/DgSecure/tomcat9/webapps/HDFSIDP/WEB-INF/classes/com/dataguise/hadoop/util

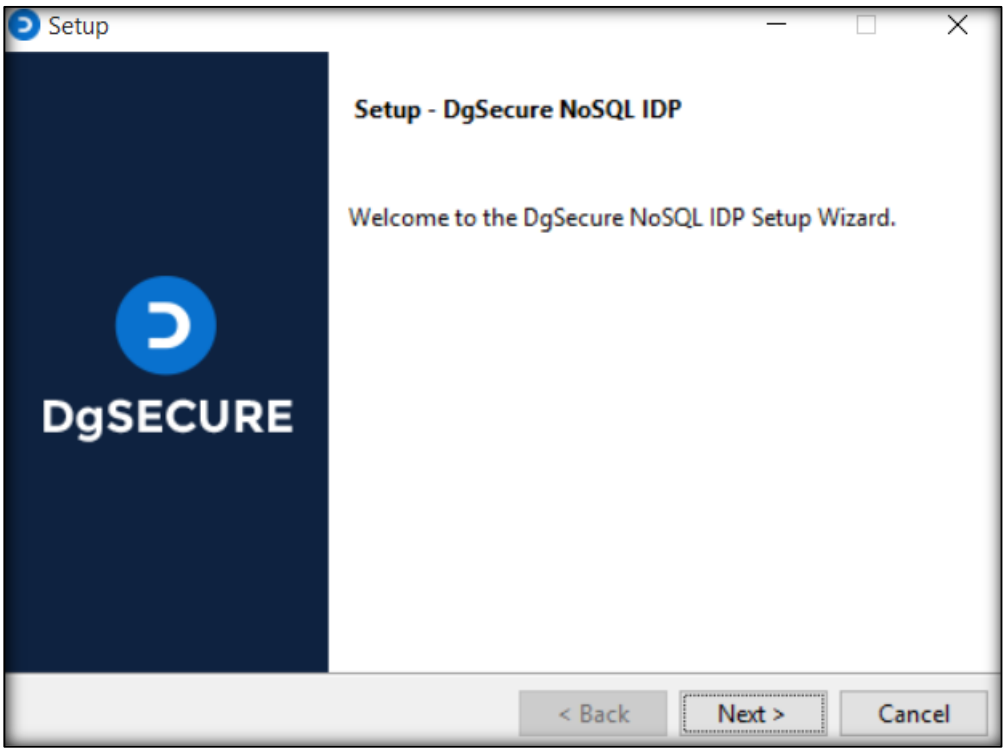
```
#Retrieve key from KMIP Server

kmip.retrieval=N  ##This must be N for RSA.

#RSA config properties file location

rsa.config.props.location=/$JAVA_HOME/jre/lib//software/jdk1.7.0_60
/jre/lib/ext/rsaconfig.properties


#RSA KeyClass

rsa.keyclass=KeyClass2


#KeyRetrieval Source Currently supported value is RSA and Other

key.retrieval.source=RSA


#KeyAliases availabe in RSA DPM. The aliases must be those belonging
to the identity as specified in client.identity_name=Identity2 in
rsaconfig.properties file.

rsa.key.aliases=KeyAlias1420533992581,KeyAlias1419064002219,KeyAlias1
419063766020,KeyAlias1419063661538,KeyAlias1419062371543,Sitkey256
```

Appendix M: NoSQL IDP

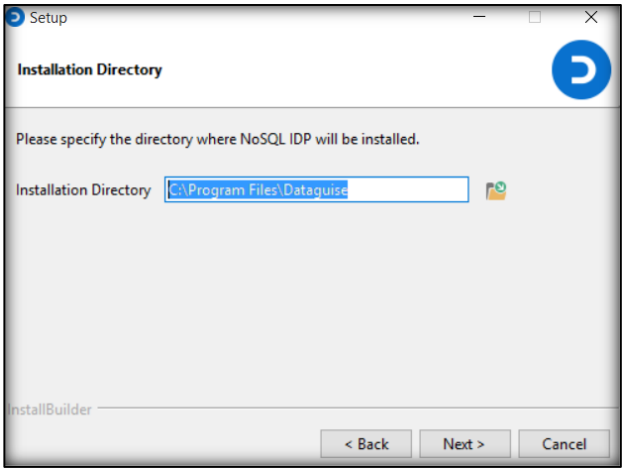
Install the NoSQL IDP

1. In the files that you received from Dataguise, locate the installation files. Copy the executable file **DgSecureNoSQLIDP-linux-x64-installer.run** to the machine on which you will install the IDP.

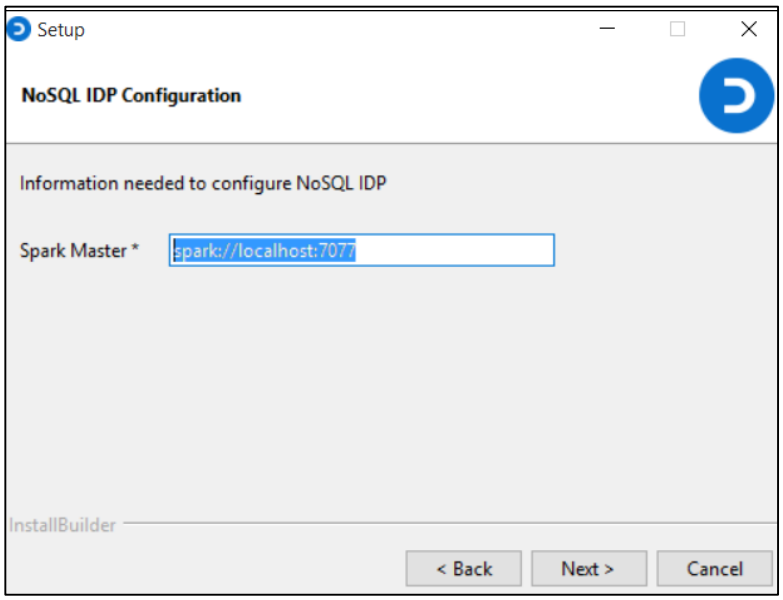
2. Run the **DgSecureNoSQLIDP-linux-installer.run.** The IDP setup wizard is displayed. Click **Next** to continue.



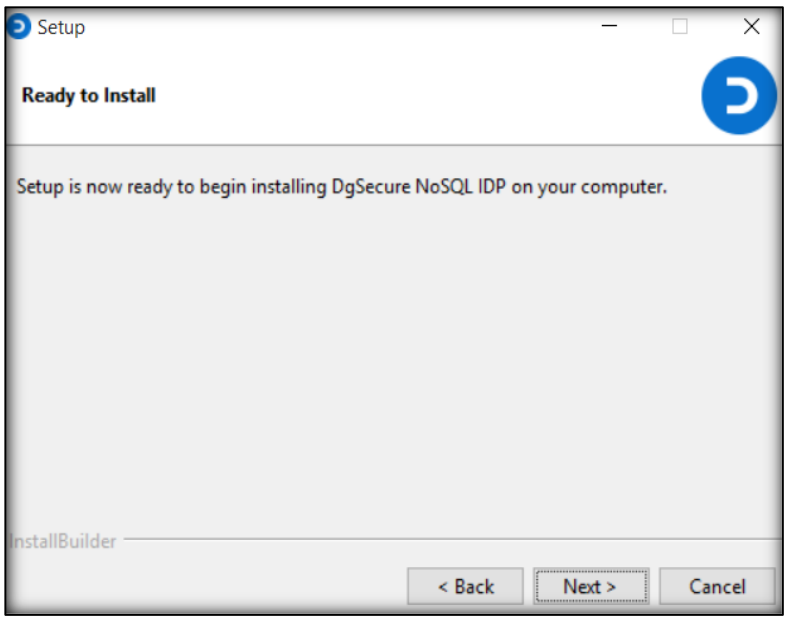
3. Accept the License Agreement. Click Next.
4. Install the IDP in the default directory (C:\Program Files\Dataguise) or select a different location. Click **Next.**

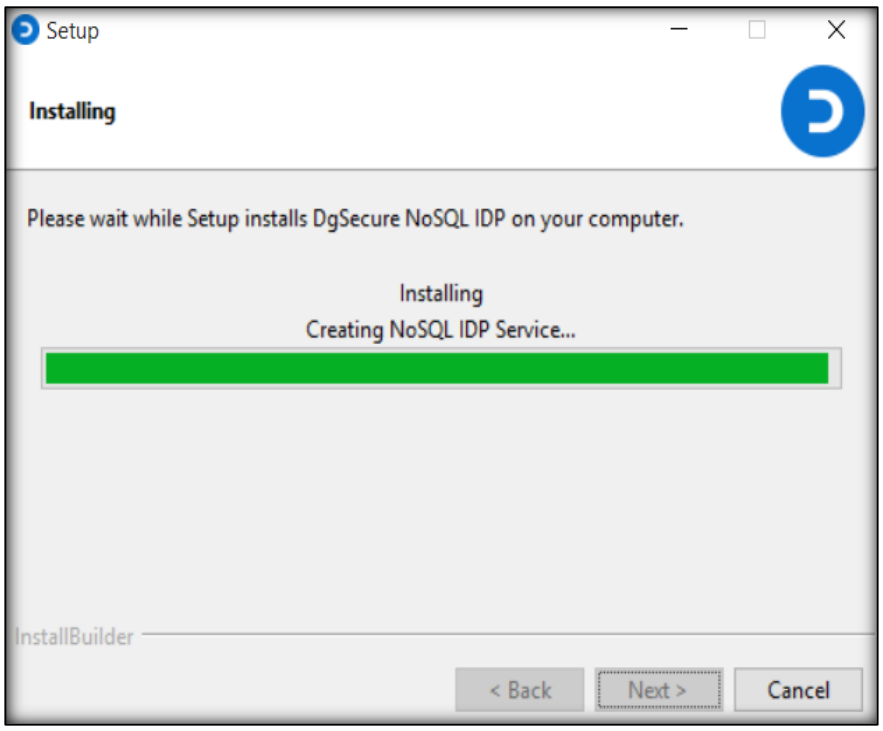


5. Enter details for Spark Configuration. Click **Next**.

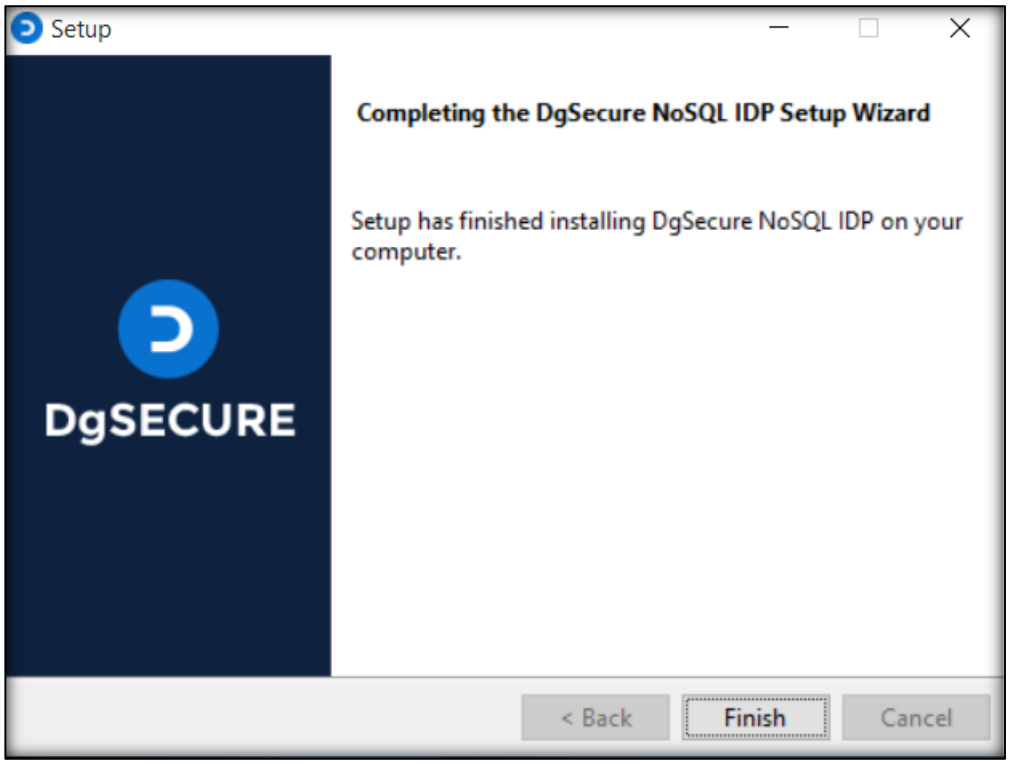


6. Click next to begin installation.



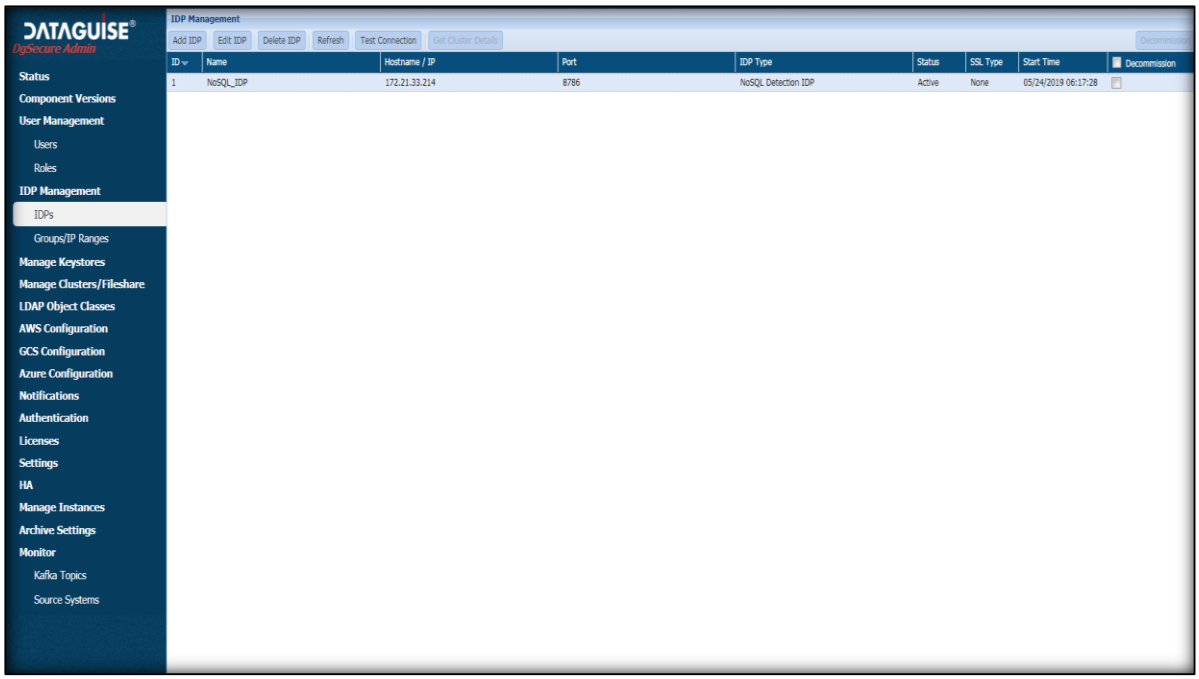


7. When installation is complete click Finish.

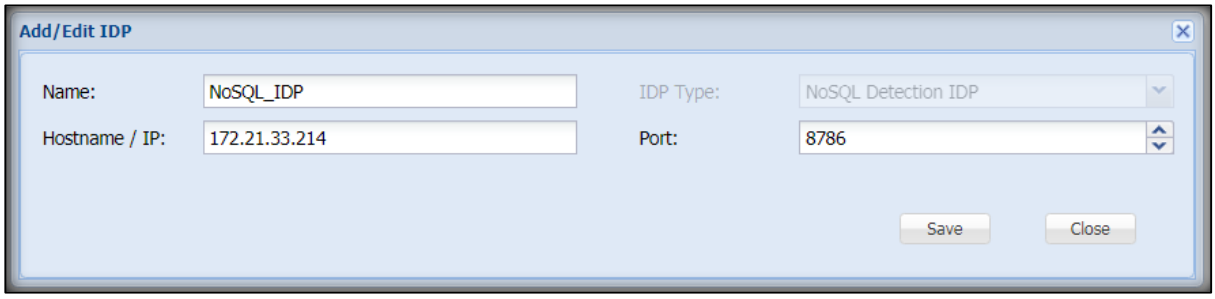


Configure the NoSQL IDP

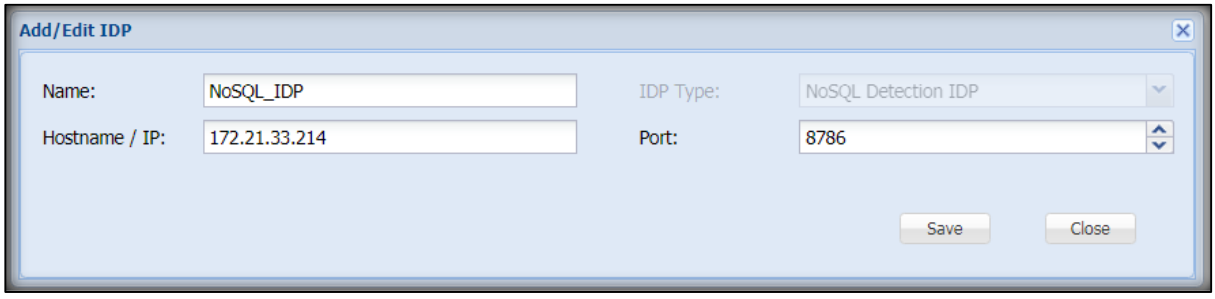
1. Go to IDPs under IDP Management in DgSecure Admin. Click Add IDP.



2. Enter Name, Hostname, and port number. Select NoSQL Detection IDP from the IDP type dropdown.

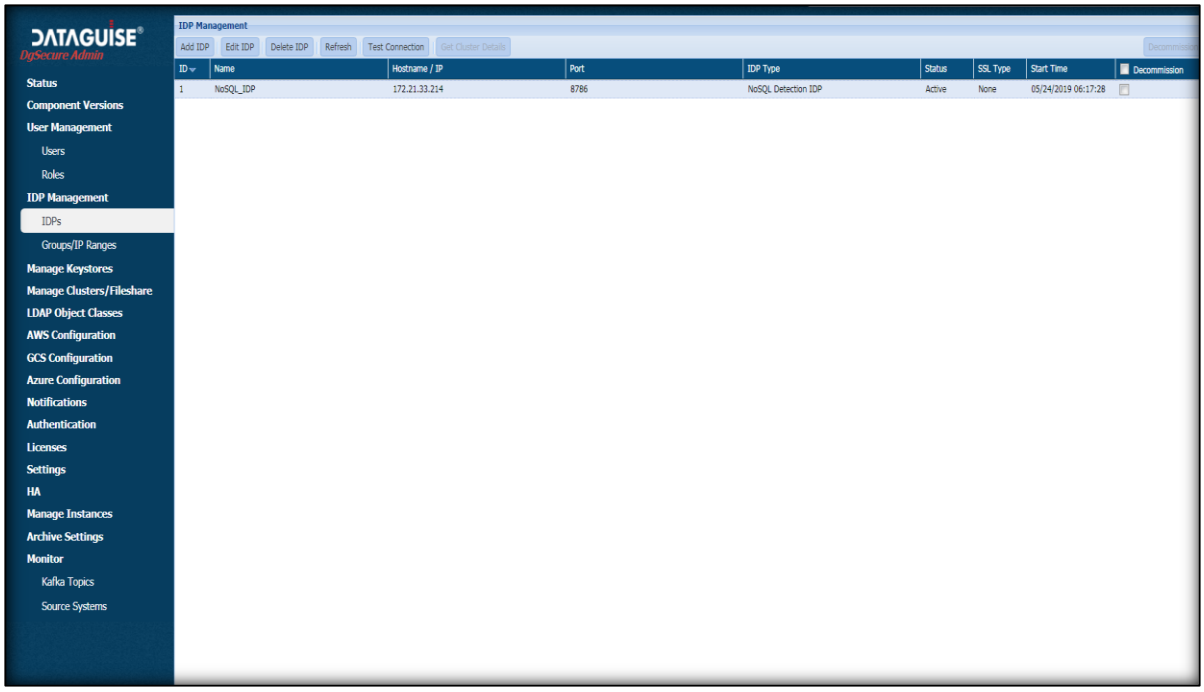


3. Save the IDP



4. The IDP will be listed under IDPs. Click Test Connection.

5. On successful connection to the created IDP, its status will become active.



Appendix N: Setup Auto-Purging

Auto-purging means removing of obsolete data automatically. This functionality has been implemented for all the components that generates log files. DgSecure allows you to auto purge the log files. You can define the size of the log files exceeding which, the log files are subjected to a roll over. The default size of the log file is 100 MB. In addition to this, you can specify the number of days after which the log files are automatically deleted. The default value is 30 days. Following are the properties:

```
<?xml version="1.0" encoding="UTF-8"?>
- <Configuration packages="" name="LoggingConfig" status="warn">
  - <Properties>
      <Property name="baseDir">logs</Property>
      <Property name="rolloverSize">100MB</Property>
      <Property name="purgeAfter">30d</Property>
      <Property name="logLevel">INFO</Property>
    </Properties>
```

Following is the name and path of the file for different IDPs, in which these properties exist:

IDP	File Name	Path
HDFS/ LFA	log4j2.xml	*<Installation_directory>/Dataguise/DgSecure/Agents/HDFSAgent/expandedArchive/WEB-INF/classes/log4j2.xml*
Cloud Agents(Azure, GCS, AWS)	log4j2.xml	*<Installation_directory>/Dataguise/DgSecure/Agents/CloudAgent/log4j2.xml*
DBMS Discover	log4j2.xml	*<Installation_directory>/Dataguise/DgSecure/Agents/DgDiscoverAgent/expandedArchive/WEB-INF/classes/log4j2.xml*
DBMS Masker	log4j2.xml	*<Installation_directory>/Dataguise/DgSecure/Agents/DgMaskerAgent/expandedArchive/WEB-INF/classes/log4j2.xml*
GDPR/Privacy	log4j2.xml	*<Installation_directory>/Dataguise/DgSecure/Agents/GDPRAgent/expandedArchive/WEB-INF/classes/log4j2.xml*

Hive	log4j2.xml	*<Install_directory>/Dataguise/DgSecure/Agents/HiveAgent/conf/log4j2.xml*
NoSQL	log4j2.xml	*<Installation_directory>/Dataguise/DgSecure/Agents/DgNoSQLAgent/expandedArchive/WEB-INF/classes/log4j2.xml*
DgWalker	log4j2.properties	*<Installation_directory>/Dataguise/DgSecure/Agents/DgWalkerAgent/log4j2.properties*
HBase	log4j2.xml	*<Install_directory>/Dataguise/DgSecure/Agents/HBaseAgent/expandedArchive/WEB-INF/classes/log4j2.xml*
Hadoop Control	log4j.properties	*<Install_directory>/Dataguise/DgSecure/Agents/MonitoringAgent/expandedArchive/WEB-INF/classes/log4j.properties*
SharePoint	Web.config	*<Install_directory>/Dataguise/logs/application.log*
Monitoring	dgsyslog4j2.xml	*<Install_directory>/Dataguise/DgSecure/tomcat9/webapps/dgcontroller/WEB-INF/classes/syslogger/dgsyslog4j2.xml*
Tomcat Component	**File Name**	**Path**
dgControl	log4j2.xml	*<Install_directory>/Dataguise/DgSecure/tomcat9/webapps/dgControl/WEB-INF/classes/log4j2.xml*
dgController	log4j2.xml	*<Install_directory>/Dataguise/DgSecure/tomcat9/webapps/dgcontroller/WEB-INF/classes/log4j2.xml*
dgHdfsInfoProcessingEngine	log4j2.xml	*<Install_directory>/Dataguise/DgSecure/tomcat9/webapps/dgHdfsInfoProcessingEngine/WEB-INF/classes/log4j2.xml*
dgUI	log4j2.xml	*<Install_directory>/Dataguise/DgSecure/tomcat9/webapps/dgUI/WEB-INF/classes/log4j2.xml*

dgDashboardUI	log4j.properties	*<Install_directory>/Dataguise/DgSecure/tomcat9/webapps/dgDashboardUI/WEB-INF/classes/log4j.properties*
dgDashboardRest	log4j.properties	*<Install_directory>/Dataguise/DgSecure/tomcat9/webapps/dgDashboardRest/WEB-INF/classes/log4j.properties*
DgLogReader	log4j.properties	*<Install_directory>/Dataguise/DgSecure/tomcat9/webapps/DgLogReader/WEB-INF/classes/log4j.properties*
DgSecureServices	log4j.properties	*<Install_directory>/Dataguise/DgSecure/tomcat9/webapps/DgSecureServices/WEB-INF/classes/log4j.properties*

Note: The user must restart the IDP/Tomcat to make the changes effective.