COHESITY

Version 4.3.0
March 2020

Cohesity Imanis Data Installation Guide

COHESITY

Table of Contents

1   About this Guide

This guide describes the process of installing Imanis Data software on your servers, collectively referred to as the Imanis Data cluster in the rest of this document. The Pre-Installation Overview chapter details the hardware and software requirements, prerequisites for primary cluster, and the prerequisites for all modes of installing Imanis Data software. The Installing Imanis Data Software chapter describes the various modes of installation in addition to the information regarding setting up Kerberos authentication, call home feature, SSL on Imanis Data cluster, and Nagios with Imanis Data installer. The information regarding actions to take installing Imanis Data software is provided in the Verifying Installation chapter. The Managing the Imanis Data Cluster chapter describes the steps to stopping and starting services, adding more nodes, and uninstalling Imanis Data software. Appendix A discusses the alternate modes of installation. While Appendix B illustrates the post installation for NOSSH-NOSUDO mode of installation.

Appendix C lists and describes the parameters of Imanis Data software installation, which includes the mandatory and optional parameters list. Appendix D outlines the installation pre-checks that should be successfully completed prior to starting the install. Appendix E is a quick reference list that needs to be completed by the administrator prior to starting the installation process. Appendix F lists the advanced command line options to manage the Imanis Data cluster. Appendix G lists the Nagios plug-ins supported by Imanis Data software. Appendix H lists the troubleshooting options should you experience any issues installing Imanis Data software and Appendix I lists the parameters for Non-HA and HA cluster. Appendix J presents the steps you need to perform for formatting and mounting disks. Appendix K and Appendix L discusses the procedures of creating and adding a swap file and disabling SeLinux respectively.

1.1   Intended Audience

This guide is meant for Imanis Data administrators and assumes that they have prior hands-on experience in installing software applications in a Linux environment. The administrator must be familiar with Linux and have some basic understanding of Hadoop concepts.

1.2   Typographical Conventions

This section describes the typographical conventions used in this guide.

CONVENTION	DESCRIPTION
`Monospace font`	▪ The text in this font indicates the text that you type. File paths and filenames, CLI commands, or information that the system displays on the screen (message text or prompts).
Bold	▪ Boldface text indicates any graphical user interface controls and column headings in tables and text.
Underlined colored text	▪ A link to a chapter or appendix, or to a glossary entry, or to the section of the document that defines the colored term. ▪ For example: Refer to Appendix C: Parameters for Imanis Data Installation

IMPORTANT: Follow the steps carefully mention in this document when installing Imanis Data software.

2   Pre-Installation Overview

This section presents the prerequisites of installing the Imanis Data software in your big data infrastructure.

NOTE: If you are installing Imanis Data software on Cohesity Storage Platform, then you must refer to the document for installing Imanis Software on Cohesity NFS Storage.

2.1   Prerequisites for Imanis Data cluster

Prior to installing the Imanis Data software, ensure the following:

1. Cluster Nodes
Cluster nodes are a list of servers on which you plan to install Imanis Data software. You must have a minimum of three cluster nodes to install Imanis Data software. These nodes are referred to as Imanis Data nodes in this guide. Following are the prerequisites of the cluster nodes:

ITEM	MINIMUM REQUIREMENT

Operating System (OS)	▪ Ubuntu 12.04, 14.04, and 16.04
	   or
	▪ Oracle Enterprise Linux 6.x and 7.x
	   or
	▪ CentOS 6.x and 7.x
	   or
	▪ Red Hat Enterprise Linux 6.x and 7.x

Supported Kernel	▪ 3.10.0-862.el7.x86_64 ▪ 3.10.0-1062.12.1.el7.x86_64 Any other kernel version between these versions is not supported. Example: 3.10.0-957.el7.x86_64 Is not supported.
File System	▪ ext3/ext4 ▪ Ensure that file is mounted with: o exec permissions

ITEM	MINIMUM REQUIREMENT
	o flock is either supported or mounted with nolock option For any other file system, contact Imanis Data Technical Support.
Disk Space	▪ 10 GB in the install directory INSTALL_DIR is the directory where installation is done by default for example, /opt/talena ▪ 50 GB in the log directory For example, INSTALL_DIR/talena-logs by default
Memory/CPU	▪ 8-16 cores ▪ 64 GB to 128 GB of RAM
Hostname	▪ The hostname should always be FQDN. So, the output of commands hostname (short hostname) and hostname -f (full hostname) should always match

NOTE: The disk space requirement for storing data in the Imanis Data cluster is different. For storing data, the datanode directory list setting (`tl_datanode_dirlist=`) in `conf/tl_installer.conf` file is used. The `conf/tl_installer.conf` file is obtained after you extract the Imanis Data self extractor.

IMPORTANT: Ensure that the difference between clock times on each of the Imanis Data nodes is not more than 30 seconds.

2. Installation Checklist

Before starting the installation process, the administrator must complete this installation checklist to speed up the overall installation process. Refer to Appendix E: Installation Checklist.

3. Imanis Data recommends Unix user with sudo privileges to install the software. This user is called the admin user and it could be root or a different user with sudo privileges. Ensure that admin user exists on all the nodes in the Imanis Data cluster and has the same password or SSH key across all Imanis Data nodes in the cluster.

If sudo privileges are not available then installation can be executed in no-sudo/no-ssh mode. In this case, the admin user has to be same as the service user (which is hdfs by default).

4. Ensure that 70 GB of free disk space is available in /opt if you plan to use default directories for install INSTALL_DIR and logs INSTALL_DIR/talena-logs. However, if you plan to use individual directories for logs and install, then you must edit conf/tl_installer_advanced.conf and ensure that free disk space of 50 GB is available for tl_log_dir and 20 GB is available for tl_install_dir respectively.

5. Ensure that all the nodes are reachable on the network from each other. If you have ICMP ping enabled in your environment, then you can use ping command to check the same.

6. If you have separate disks available for data, then you must prepare the disks to be used with Imanis Data cluster. Create a new file system (mkfs) and mount all the data disks on each node at the same mount point. Please refer to Appendix J for the procedure of formatting and mounting the disks.

7.All the TCP ports should be open between the nodes where Imanis Data cluster is installed. Use the following steps to turn off all the firewall services for the nodes. This step has to be executed on all nodes.

On CentOS/RHEL/OEL 6.x
```
# service iptables stop
# chkconfig iptables off
```

On CentOS/RHEL/OEL 7.x
```
# systemctl stop firewalld
# systemctl disable firewalld
# systemctl status firewalld
```

On Ubuntu
```
# service ufw stop
# service ufw status
# ufw disable
```

8. Download the Imanis Data Self Extractor
The Imanis Data Self Extractor is a script to install Imanis Data software on Linux operating systems. When you run the Imanis Data Self Extractor script, it will extract itself in the current working directory. In this folder, you will see various installer programs, Packages (RPMs), Java Development Kit (JDK) and other files that are used during the installation.

You need to have a Dropbox download link from where you can download the Imanis Data Self Extractor. Please contact Imanis Data Technical Support to receive the Imanis Data Self Extractor. Once you have downloaded the Imanis Data self-extractor script, ensure the following:
   a. The md5sum matches with that of the md5sum in the text file md5sum.txt (available in the Dropbox folder shared with you).
   b. The downloaded self-extract file is executable.

c. If you plan to use the Callhome feature, then you should get a Dropbox access key from the Imanis Data Technical Support Team. This file has will have only the access token for uploading the logs (and other information needed for troubleshooting) to Dropbox account created exclusively for your organization. For security purposes, do not share this file with anyone outside the organization.

2.2 Pre-requisites for Primary cluster

Primary cluster is defined as the existing application cluster on which you are trying to manage and protect data through Imanis Data software. This could be, for example, Cloudera or an Apache Cassandra cluster.

For Hadoop, ImanisData software also supports Non-High Availability cluster and High Availability cluster. It is recommended that administrators verify if Namenode HA is enabled before installing Imanis Data software. Refer to Appendix I for Parameters for Non-High Availability cluster and Parameters for High Availability (HA) cluster for more information.

3   Installing Imanis Data Software

This chapter describes the various modes of installation and the pre-checks that you must conduct before installing Imanis Data software. In addition, the section presents information on setting up Kerberos authentication, call home feature, SSL on Imanis Data cluster, and Nagios with Imanis Data cluster.

3.1   Various Modes of Installation

Depending on the kind of access restrictions you may have for the install user, you need to follow one of the modes of installation.

MODE OF INSTALLATION	DESCRIPTION
SSH-SUDO	You are installing with 'root' user OR an admin user who has sudo privileges. **This is the preferred mode of Installation of Imanis Data Software.**
NOSSH-SUDO	You have root or sudo privileges available for admin user, but this user cannot SSH to other nodes.
NOSSH-NOSUDO	You neither have root or sudo privileges for admin user and this user cannot SSH to other nodes.

NOTE: SSH-NOSUDO mode of installation is NOT supported.

Refer to the information on admin user for more details. Based on the preceding information, you can install Imanis Data software for SSH-SUDO mode of installation (the preferred mode of installation) using the following steps.

Steps for SSH-SUDO Mode of Installation:
1. Log in to the Install Master node as the admin user.
2. Copy the downloaded self-extractor script to the current working directory.

3. Extract the installer using the following command:

```
#./installer_self_extract_4.3.0-RELXXX.sh
```

Where XXX is the release number of the build. For example, REL136.

4. The following will be displayed on the screen:

```
#./installer_self_extract_4.3.0-RELXXX.sh

Imanis Data installation is supported in two modes:
SSH: Root user OR user with sudo privileges can SSH to nodes and execute
commands.

NO-SSH: SSH access is not allowed for root.

Nodes will communicate through Remote Procedure Calls.

Do you want to continue with SSH mode installation? (yes/no) Yes
```

Where XXX is the release number of the build. For example, REL136.

5. Type **Yes** to the question "**Do you want to continue with SSH mode installation?**" in the above confirmation message to continue with the installation. For NOSSH-SUDO or NOSSH-NOSUDO modes of installation, follow the steps listed in Appendix A.

3.2   Running Pre-checks

The following pre-checks must be performed on the Imanis Data cluster nodes. Pre-checks are common for all the modes of installation. The pre-checks ensure that configuration files have been correctly set up, there is sufficient disk space, required third party packages are installed and network connectivity can be established. It is highly recommended to perform the pre-checks and ensure all pre-checks have been passed before installing Imanis Data software.

1. Change the current working directory to the installer directory.
   - If you have chosen SSH mode installation in the preceding step, then do the following:

   ```
   # cd /root/<VERSION>
   ```

   Where /root is the current working directory and VERSION is the version of the software. However, it could be any directory in which you choose to extract the installer.

   - If did not choose SSH mode installation in the preceding step, then do the following:

```
# /root/talena/<VERSION>
```

Where /root is the current working directory and VERSION is the version of the software. However, it could be any directory in which you choose to extract the installer.

2. Modify the configuration file conf/tl_installer.conf in a text editor and specify the following mandatory parameters:
   - tl_physical_nodes=
   - tl_datanode_dirlist=
   - tl_namenode_dir=
   - tl_journalnode_localdir=
   - tl_cluster_name=
   - tl_licence_key_path=
   - tl_tmp_dir=

   In addition to editing the mandatory parameters mentioned above, you can also choose to specify the following optional parameters:
   - HTTPS for ImanisData web UI
   - Kerberos
   - Nagios
   - Callhome
   - Email Notifications

   In case you do not choose to specify the optional parameters, you can use the Cluster Management Utility tool to configure them. For more information, refer to the section Configuring System Using The Cluster Management Utility.

   During installation, you must provide cluster name in the configuration file tl_installer.conf. This value will be used to configure cluster name which will be displayed on the Imanis Data software GUI.

NOTE: In case of NOSSH-SUDO or NOSSH-NOSUDO installation, set the value of parameter `tl_use_ssh=False`.

For more information on mandatory parameters, refer to Appendix C: Parameters for Imanis Data Installation.

3. Change directory to where the installer was extracted and run the following pre-check command:
```
$INSTALL_DIR/$TALENA_VERSION/precheck_talena.sh --assumeyes
```

4. After all the pre-checks failures (if any) are addressed, ensure that you verify all the prerequisites again by running the following command:

```
# ./precheck_talena.sh --assumeno
```

'PASS' is displayed against respective prerequisites upon successful verification.

5. Verify the false positive for IP tables command on Ubuntu 12.x/14.x. As a non-sudo user, it is difficult to positively confirm if IP tables are down or not. Thus, if you hit this false positive, do one the following:
    - Verify if IP tables are disabled as a root or sudo user
    - If IP tables are already disabled, then ignore and continue with the installation

3.3 Installing & Configuring Imanis Data Software

You must run the following script to install and configure Imanis Data software. As mentioned in the preceding paragraph, ensure that all the pre-checks have been successfully passed before running the script.

1. Run the following script to install Imanis Data software:

```
# ./install_talena.sh
```

NOTE: While installer is in progress, it may prompt you when some commands fail to produce the expected output. In this case, you have following choices:

c - continue (assumes that the problem reported is fixed manually and continues as if no errors were reported)

a - abort (aborts installer at this point. You need to fix the issue and re-run installer)

e - exclude (excludes the problematic node and continues with the rest of the nodes in the cluster)

NOTE: In case of NOSSH-NOSUDO mode of installation, there are some steps which need to be performed as a 'root' user. These steps will be prompted to you as the installer encounters them. However, you need NOT take any action at that time. You can continue with the installation process by choosing 'c' for all such actions which require root privilege.

---

2. Accept "End User License Agreement" to proceed with the installation.

3. While installing the Imanis Data software, the installer wizard prompts you to do various tasks. The installer wizard also prompts you to set up Kerberos authentication, Call Home on Imanis Data cluster, SSL on Imanis Data cluster, and also installs Nagios with Imanis Data cluster. For

instructions on setting up Kerberos, Call Home, SSL and Nagios refer to the respective sections in Imanis Data Installation Guide.

Overall the installer takes approximately 30 minutes for the software to install on a three node Imanis Data cluster. Once the installation is complete, smoke tests are initiated automatically to ensure that the system is operating as expected.

NOTE: There are many optional settings in the `conf/tl_installer.conf` in addition to the mandatory parameters such as service placement details, SSH private key for password-less logins that you may want to edit going forward. Refer to Appendix C for more information.

NOTE: In case of NOSSH-NOSUDO mode of Installation, make sure you perform the Post Installation steps as listed in Appendix B.

---

4   Configuring Kerberos in Imanis Data Cluster

This section describes the following points when configuring Kerberos in Imanis Data cluster environment.

- Install time configuration changes for Imanis Data
- Adding new Kerberos realm to Imanis Data cluster
- Adding new Hadoop data repository for a newly added realm
- Configuring external KDC server when adding nodes to the existing Imanis Data cluster

4.1   Install time configuration changes for Imanis Data

You can set up Kerberos authentication in the following ways:

- Setting up a dedicated Kerberos server on the Imanis Data cluster
- Setting up Imanis Data to use an external (existing) Kerberos Server

4.1.1   Setting up Kerberos on the Imanis Data cluster

This section describes the procedure for setting up dedicated Kerberos server on the Imanis Data cluster nodes. When you choose to install Kerberos in the Imanis Data Installer, the installer prompts you to run a separate script as a root user on one of the nodes in the Imanis Data cluster. This script executes the installation and sets up Kerberos on the Imanis Data cluster nodes. At the time of installing Imanis Data, you need to set the appropriate Kerberos related parameters in the `tl_installer.conf` file to enable Kerberos.

1. Edit `tl_installer.conf` file and set the following Kerberos related parameters.

```
tl_kerberos_enable = true/false

(default is false) if set to true, then you need to define the following
parameters:

tl_kdc_realm_name =
```

This is the realm name you must use for the dedicated Kerberos server for Imanis Data Cluster. We assume this name as `TALENA.REALM` in this document throughout.

2. Run the installer.

The installer installs Imanis Data packages along with the packages needed for setting up Kerberos server after running certain pre-checks. Before starting the services, the installer prompts the user to set up Kerberos as follows:

```
Please execute the script: $INSTALL_DIR/bin/talena-kerberos.sh setup internal
on node: XXX with root privileges to setup ImanisData KDC server
Press [Enter] to continue after these steps.
```

Where XXX is the node running the Imanis Data Kerberos server.

3. Access the installer prompt after this script completes execution and press **Enter** to continue with the installation. Ensure that the installer is executed and then at the end of the installation, Imanis Data services should be up and running.

4.1.2 Configuring Imanis Data software using the external KDC Server

In some cases, you may not want to setup a dedicated Kerberos server for Imanis Data cluster. Imanis Data software allows users to use an existing Kerberos setup. The following sections describe the procedure steps for MIT5 Kerberos on a Unix machine and Microsoft Windows AD Kerberos server.

4.1.2.1 For MIT5 Kerberos on a Unix machine

When MIT5 Kerberos running on a Unix machine is used as an external KDC server, the following is assumed:

- The external KDC server is running on node: `primary-node1.mydomain.com`, with the realm name: `PRIMARY.REALM`
- Imanis Data is being installed on `imanis-node1.mydomain.com, imanis-node2.mydomain.com, imanis-node3.mydomain.com`

  You must replace the preceding placeholders with the actual host names and Kerberos Realm name in your own setup. Also ensure that all nodes, including the Kerberos server are resolvable on each Imanis Data node either through DNS or through /etc/hosts.

To install Imanis Data software using MIT5 Kerberos on a Unix machine, do the following:

1. Ensure that external Kerberos server: `primary-node1.mydomain.com` is reachable from each of the Imanis Data nodes.
2. Add the following service principals used by Imanis Data services:

```
kadmin.local -q "add_principal -randkey $SERVICE_USER/imanis-
node1.mydomain.com@PRIMARY.REALM"
kadmin.local -q "add_principal -randkey HTTP/imanis-
node1.mydomain.com@PRIMARY.REALM"
```

> **NOTE**: Replace `$SERVICE_USER` with its value. You can check its value from `/etc/talena/talena-env.sh`

3. Add similar principals for `imanis-node2.mydomain.com` and `imanis-node3.mydomain.com`.
4. Add the following user principals used by Imanis Data software:

```
kadmin.local -q "add_principal -randkey $SERVICE_USER@PRIMARY.REALM"
kadmin.local -q "add_principal -randkey hbase@PRIMARY.REALM"
```

5. Generate a keytab file for all added principals:

```
kadmin.local -p admin/admin -q "xst -norandkey -k /tmp/talena.keytab
$SERVICE_USER@PRIMARY.REALM
hbase@PRIMARY.REALM
$SERVICE_USER/imanis-node1.mydomain.com@PRIMARY.REALM
HTTP/imanis-node1.mydomain.com@PRIMARY.REALM
$SERVICE_USER/imanis-node2.mydomain.com@PRIMARY.REALM
HTTP/imanis-node2.mydomain.com@PRIMARY.REALM
$SERVICE_USER/imanis-node3.mydomain.com@PRIMARY.REALM
HTTP/imanis-node3.mydomain.com@PRIMARY.REALM"
```

The preceding command will create the keytab file at location: `/tmp/talena.keytab`

6. Create new directory called `/tmp/krb-files-copy` on one of the Imanis Data nodes. Assumed as `imanis-node1.mydomain.com`.

7. Copy the `/etc/krb5.conf` file and the generated keytab `/tmp/talena.keytab` to directory `/tmp/krb-files-copy` on `imanis-node1.mydomain.com`.

8. Run the installer with Kerberos enabled and follow the installer prompts. You will get a prompt like the following for setting up Kerberos:

```
Please execute the script: $INSTALL_DIR/bin/talena-kerberos.sh setup external
on one of the cluster nodes with root privileges to configure ImanisData to
use external KDC.
Press [Enter] to continue after these steps.
```

9. Execute the following script on node where you have copied the keytab file and `krb5.conf`. In this case, the node is `imanis-node1.mydomain.com`.

```
# ${INSTALL_DIR}/bin/talena-kerberos.sh setup external
```

10. Follow the prompts and provide the directory path `/tmp/krb-files-copy` as an input for the following prompt.

11. Provide local directory path where `krb5.conf` and keytabs downloaded from KDC `primary-node1.mydomain.com: /tmp/krb-files-copy`.

12. Ensure that the script is successfully executed.

13. On the Installer prompt, press **Enter** to continue the installation and then exit.

14. At the end of installation, Imanis Data services should be up and running.

4.1.2.2   For Microsoft Windows AD Kerberos Server

The Microsoft Windows AD Kerberos server does not allow setting multiple principal names to a single AD user. Hence, you may have to create as many AD users as the number of principals to be created for Imanis Data.

> **NOTE**: It is recommended to do the following before proceeding:
> Replace PRIMARY.REALM below with your Kerberos realm name
> Replace password with your actual passwords

To install Imanis Data software using Microsoft Windows AD Kerberos Server, do the following:

1. Create the users in your Windows AD Kerberos server. The number of AD users to be created is 2 x number of nodes in the Imanis Data cluster + 2. Thus, for a 3-node cluster, you must create 8 users:

```
user1
user2
user3
user4
user5
user6
$SERVICE_USER
hbase
```

2. Replace `$SERVICE_USER` with its value in `/etc/talena/talena-env.sh`. Usually, the value of `$SERVICE_USER` is `hdfs`.

3. Run the following commands on your Microsoft Windows AD Kerberos Server to associate the following service principals to the AD user:

```
setspn -A  $SERVICE_USER/host1@PRIMARY.REALM user1
setspn -A  $SERVICE_USER/host2@PRIMARY.REALM user2
setspn -A  $SERVICE_USER/host3@PRIMARY.REALM user3
setspn -A  HTTP/host1@PRIMARY.REALM user4
setspn -A  HTTP/host2@PRIMARY.REALM user5
setspn -A  HTTP/host3@PRIMARY.REALM user6

ktpass -princ $SERVICE_USER/host1@PRIMARY.REALM -pass password -mapuser user1
-crypto ALL -ptype KRB5_NT_PRINCIPAL -out hdfs_host1.keytab -kvno 0
ktpass -princ $SERVICE_USER /host2@PRIMARY.REALM -pass password -mapuser
user2 -crypto ALL -ptype KRB5_NT_PRINCIPAL -out hdfs_host2.keytab -kvno 0
ktpass -princ  $SERVICE_USER/host3@PRIMARY.REALM -pass password -mapuser
user3 -crypto ALL -ptype KRB5_NT_PRINCIPAL -out hdfs_host3.keytab -kvno 0
ktpass -princ HTTP/host1@PRIMARY.REALM -pass password -mapuser user4 -crypto
ALL -ptype KRB5_NT_PRINCIPAL -out HTTP_host1.keytab -kvno 0
ktpass -princ HTTP/host2@PRIMARY.REALM -pass password -mapuser user5 -
crypto
ALL -ptype KRB5_NT_PRINCIPAL -out HTTP_host2.keytab -kvno 0
ktpass -princ HTTP/host3@PRIMARY.REALM -pass password -mapuser user6 -crypto
ALL -ptype KRB5_NT_PRINCIPAL -out HTTP_host3.keytab -kvno 0
ktpass -princ $SERVICE_USER@PRIMARY.REALM -pass password -mapuser
$SERVICE_USER  -crypto ALL -ptype KRB5_NT_PRINCIPAL -out   hdfs.keytab -kvno
0
ktpass -princ hbase@PRIMARY.REALM -pass password -mapuser hbase -crypto ALL
-ptype KRB5_NT_PRINCIPAL -out hbase.keytab -kvno 0
```

4. Ensure that you replace `$SERVICE_USER` with its value in `/etc/talena/talena-env.sh`. Usually, the value of `$SERVICE_USER` is `hdfs` in the preceding commands.
5. Copy all the keytab files to one of the nodes in the Imanis Data Cluster at some temporary location, for example `/tmp` and merge all keytabs into single keytab. This command will generate `talena.keytab` in your current directory.

```
# ktutil
ktutil: read_kt hdfs..keytab
ktutil: read_kt hbase.keytab
ktutil: read_kt hdfs_host1.keytab
ktutil: read_kt hdfs_host2.keytab
ktutil: read_kt hdfs_host3.keytab
ktutil: read_kt HTTP_host1.keytab
ktutil: read_kt HTTP_host2.keytab
ktutil: read_kt HTTP_host3.keytab
ktutil: write_kt talena.keytab
ktutil: quit
```

6. The `krb5.conf` is a configuration file that has the details about the realm names and servers running Kerberos KDC services. This configuration file is required by Imanis Data nodes to know

where to connect to when requesting for Kerberos tickets. A typical configuration file looks like the following:

```
[logging]
default = FILE:/opt/log/krb5libs.log
kdc = FILE:/opt/log/krb5kdc.log
admin_server = FILE:/opt/log/kadmind.log
[libdefaults]
default_realm = your_realm_name
dns_lookup_realm = false
dns_lookup_kdc = false
ticket_lifetime = 24h
renew_lifetime = 7d
forwardable = true
validate = false
[realms]
your_realm_name = {
kdc = FQDN_of_the_host_running_kdc_server
admin_server = FQDN_of_the_host_running_admin_server
}
```

NOTE: Replace the fields in bold with values for your Kerberos environment.

7. Copy the `krb5.conf` configuration file to each of the Imanis Data nodes at location:
    `$INSTALL_DIR/conf/krb5.conf`.

NOTE: If this file already exists on Imanis Data cluster nodes, then you need to merge the entries with the existing file. Refer to the section Adding a new Kerberos realm to Imanis Data cluster on steps to add entries to an existing `krb5.conf` file.

8. Copy `talena.keytab` to all Imanis Data nodes at the location `$INSTALL_DIR/conf/`. Ensure that the keytab is owned and readable only by `$SERVICE_USER:$SERVICE_USER` user.

9. Verify that the `talena.keytab` is working from each node in the cluster by doing the following:
    a. Ensure that the following command succeeds from all the nodes:

```
# kinit -kVt $INSTALL_DIR/conf/talena.keytab $SERVICE_USER@PRIMARY.REALM
```

    b. Ensure that the following command list the files on your CDH/HDP Hadoop cluster and work if everything is right:

```
# hadoop dfs -ls webhdfs://<primary_active_namenode>:50070/
```

    c. Ensure that the following command succeeds from all the hosts:

```
# kinit -kVt $INSTALL_DIR/conf/talena.keytab hbase@PRIMARY.REALM
```

    d.   Ensure that the following commands succeed from host1, host2, and host3 respectively:

```
# kinit -kVt $INSTALL_DIR/conf/talena.keytab
$SERVICE_USER/host1@PRIMARY.REALM
```

```
# kinit -kVt $INSTALL_DIR/conf/talena.keytab
$SERVICE_USER/host2@PRIMARY.REALM
```

```
# kinit -kVt $INSTALL_DIR/conf/talena.keytab
$SERVICE_USER/host3@PRIMARY.REALM
```

10. Proceed to run steps for enabling Kerberos in Imanis Data.

4.2 Adding a new Kerberos realm to Imanis Data cluster

In order to talk to any Kerberos-enabled primary data repository, Imanis Data cluster must know the Kerberos realm that your primary data source belongs to and the keytab to be used to connect to it.

This script enables you to talk to any Kerberos-enabled primary data repository. The user must provide the following inputs to the script:

    a.   **Information about the Kerberos realm, KDC server hosting the realm, and so on.**
       There are two ways in which the information about new Kerberos realm to be added is given as input to this script.
- By pointing the script to an existing `krb5.conf` file. The conf file will be read and parsed to extract the relevant information. Download the `krb5.conf` file from your primary node to the node on Imanis Data cluster
- In case you don't have access to `krb5.conf` file, you can type the required information manually by answering the prompts.

    b.   **Keytab files to be used**
- Download the keytab file(s) from your primary node to one of the nodes on ImanisData cluster.

The following section displays both the use cases.

4.2.1 User enters the Kerberos configuration parameters through a set of prompts

```
#  $INSTALL_DIR/bin/talena-kerberos.sh addrealm

INFO: Checking nodes reachability
INFO: Checking keytab and krb5.conf

WARN: Imanis services will be restarted as a part of addrealm
      This will take few minutes
      All running jobs will be killed

Do you wish to continue (yes/no) yes ?

INFO: Kerberos configs can be read from krb5.conf OR
      User can interactively provide Kerberos configs
      addrealm need details like KDCServer/Realm/DomainRealms

Do you wish to use krb5.conf to get Kerberos configs (yes/no) yes ? no

Please specify following kerberos configs.

Primary Realm name    : PRIMARY.REALM
Primary KDC hostname : t14vm3.talena.com
Primary Domain Realm list(| seperated)  :
t14vm3.talena.com|t14vm2.talena.com|t14vm1.talena.com
Permitted Encryption(optional) :

NOTE: Please download keytabs from primary to local directory
Provide local directory path where keytabs downloaded from primary : /tmp3

INFO: Successfully read keytab: /tmp3/talena.keytab

INFO: Checking KDC server(t14vm3.talena.com) reachability from
t14vm12.talena.com
INFO: Checking KDC server(t14vm3.talena.com) reachability from
t14vm11.talena.com
INFO: Checking KDC server(t14vm3.talena.com) reachability from
t14vm10.talena.com

INFO:Distributing Kerberos config file to node t14vm12.talena.com
INFO:Distributing Kerberos config file to node t14vm11.talena.com
INFO:Distributing Kerberos config file to node t14vm10.talena.com

INFO: keytabs will be merged with /opt/talena/conf/talena.keytab
INFO:Distributing merged keytab to node t14vm12.talena.com
INFO:Distributing merged keytab to node t14vm11.talena.com
```

```
INFO:Distributing merged keytab to node t14vm10.talena.com

INFO: Updating hadoop conf files
INFO:Distributing updated hadoop conf to node t14vm12.talena.com
INFO:Distributing updated hadoop conf to node t14vm11.talena.com
INFO:Distributing updated hadoop conf to node t14vm10.talena.com

INFO: Restarting Talena services, This will take few minutes. Please wait
....
INFO: Stopping service : UI
INFO: Stopping service : scheduler
INFO: Stopping service : oozie
INFO: Stopping service : hive
INFO: Stopping service : postgres
INFO: Stopping service : tcollector
INFO: Stopping service : hbase
INFO: Stopping service : yarn
INFO: Stopping service : hdfs
INFO: Stopping service : zk

INFO: Starting service : zk
INFO: Starting service : hdfs
INFO: Starting service : yarn
INFO: Starting service : hbase
INFO: Starting service : tcollector
INFO: Starting service : postgres
INFO: Starting service : hive
INFO: Starting service : oozie
INFO: Starting service : scheduler
INFO: Starting service : UI

INFO: addrealm completed for PRIMARY.REALM
```

4.2.2 User points script to a directory where krb5.conf is available

```
# $INSTALL_DIR/bin/talena-kerberos.sh addrealm

INFO: Checking nodes reachability      ←  Ensure all nodes in the cluster are
available for this operation.
INFO: Checking keytab and krb5.conf

WARN: Imanis services will be restarted as a part of addrealm. This will take
few minutes

All running jobs will be killed  ← Perform this operation when no long
running jobs are running.
```

```
Do you wish to continue (yes/no) yes ?

INFO: Kerberos configs can be read from krb5.conf OR User can interactively
provide Kerberos configs addrealm need details like
KDCServer/Realm/DomainRealms

Do you wish to use krb5.conf to get Kerberos configs (yes/no) yes ?

NOTE: Please download krb5.conf and keytabs from primary to local directory
Provide local directory path where krb5.conf and keytabs downloaded from
primary : /tmp3

Primary Realm : PRIMARY.REALM
Primary KDC hostname : t14vm3.talena.com
Domain Realm list : t14vm1.talena.com|t14vm3.talena.com|t14vm2.talena.com

Do you wish to continue (yes/no) yes ?

INFO: Successfully read keytab: /tmp3/talena.keytab

INFO: Checking KDC server(t14vm3.talena.com) reachability from
t14vm12.talena.com
INFO: Checking KDC server(t14vm3.talena.com) reachability from
t14vm11.talena.com
INFO: Checking KDC server(t14vm3.talena.com) reachability from
t14vm10.talena.com

INFO: Distributing Kerberos config file to node t14vm12.talena.com
INFO: Distributing Kerberos config file to node t14vm11.talena.com
INFO: Distributing Kerberos config file to node t14vm10.talena.com

INFO: keytabs will be merged with /opt/talena/conf/talena.keytab
INFO: Distributing merged keytab to node t14vm12.talena.com
INFO: Distributing merged keytab to node t14vm11.talena.com
INFO: Distributing merged keytab to node t14vm10.talena.com

INFO: Updating hadoop conf files
INFO: Distributing updated hadoop conf to node t14vm12.talena.com
INFO: Distributing updated hadoop conf to node t14vm11.talena.com
INFO: Distributing updated hadoop conf to node t14vm10.talena.com

INFO: Restarting Talena services, This will take few minutes. Please wait
....

INFO: Stopping service: UI
INFO: Stopping service: scheduler
INFO: Stopping service: oozie
INFO: Stopping service: hive
```

```
INFO: Stopping service: postgres
INFO: Stopping service: tcollector
INFO: Stopping service: hbase
INFO: Stopping service: yarn
INFO: Stopping service: hdfs
INFO: Stopping service: zk

INFO: Starting service: zk
INFO: Starting service: hdfs
INFO: Starting service: yarn
INFO: Starting service: hbase
INFO: Starting service: tcollector
INFO: Starting service: postgres
INFO: Starting service: hive
INFO: Starting service: oozie
INFO: Starting service: scheduler
INFO: Starting service: UI

INFO: addrealm completed for PRIMARY.REALM
```

This script can also be used to make changes to keytab for an existing data repository, without having to modify the realm information. If the Kerberos realm information is already present with the Imanis Data cluster, then only keytab is processed by the script and changes are incorporated.

4.3 Adding Hadoop, HBase, or Hive data repository

This section describes the steps that must be performed on the primary source cluster to add new Hadoop, HBase, or Hive data repositories.

> **NOTE**: At the end of these steps, service restart is required on both Primary and Imanis Data cluster. Running jobs will be killed. Please plan this activity when no long running jobs are running.

There are some steps that you need to follow while adding a new Hadoop data repository which is Kerberos enabled. This step is needed when you are adding a new data repository with a new set of Kerberos principals.

1. Add the new realm to Imanis Data cluster as mentioned in the section **Adding a new Kerberos realm to Imanis Data cluster**.

2. Make changes to the Primary cluster HDFS and HBase configuration.

3. Let us assume that you plan to use the following on your primary Hadoop cluster:

```
imanis_hdfs@PRIMARY.REALM – Kerberos principal used for HDFS workflows
imanis_hbase@PRIMARY.REALM – Kerberos principal used for HBase workflows
imanis_hive@PRIMARY.REALM – Kerberos principal used for Hive workflows
```

4. These Kerberos principals map to corresponding Unix users. For example,
    imanis_hdfs@PRIMARY.REALM maps to Unix user imanis_hdfs on your primary cluster nodes.

5. Similarly, `imanis_hbase@PRIMARY.REALM` maps to `imanis_hbase` and
    `imanis_hive@PRIMARY.REALM` maps to `imanis_hive` Unix user. All these three users should be
    added to a group, for example, `imanis_supergroup` in all cluster Nodes of your CDH cluster.

    - **Grant imanis_hdfs superuser privileges in primary CDH cluster to the Kerberos principals**
        a. In **Cloudera Manager Admin Console**, click **HDFS** service > **Configuration** tab.
        b. Search for the Superuser Group property and add imanis_supergroup to this property.
        c. Save the changes and restart HDFS service.

    - **Set hbase.superuser to imanis_hbase**
        a. In **Cloudera Manager Web UI**, click **HBase** > **Configuration** tab.
        b. Search for `hbase.superuser` and set it to imanis_hbase.
        c. Deploy the changed configuration and restart HBase service.

    - **On the new CDH, add hbase.superuser to /home/hdfs/imanis/hbase/conf/hbase-site.xml if
      not already present.**

```
<property>
<name>hbase.superuser</name>
<value>imanis_hbase</value>
</property>
```

    - **On the new CDH, add dfs.webhdfs.enabled property to duplicated hadoop hdfs-site.xml file
      /home/hdfs/imanis/hadoop/conf/hdfs-site.xml:**

```
<property>
<name>dfs.webhdfs.enabled</name>
<value>true</value>
</property>
```

4.4 Configuring external KDC server when adding nodes to the existing Imanis Data cluster

This section describes the steps that need to be performed on your Kerberos server as a pre-requisite for the addnode operation when adding more nodes to a Kerberos-enabled Imanis Data cluster.

The following assumptions are made for illustration purposes:
- Your Kerberos server is `primary-node1.mydomain.com` and realm name is `PRIMARY.REALM`
  There is an existing Imanis Data cluster with three nodes `imanis-node1.mydomain.com`, `imanis-node2.mydomain.com` and `imanis-node3.mydomain.com`
- You are now planning to add `imanis-node4.mydomain.com` and `imanis-node5.mydomain.com` to the existing Imanis Data cluster

To add nodes to the external KDC Server, do the following:

1. Add the $SERVICE_USER and HTTP service principals for `imanis-node4.mydomain.com`:

```
kadmin.local -q "add_principal -randkey $SERVICE_USER/ imanis-node4.mydomain.com@PRIMARY.REALM"
kadmin.local -q "add_principal -randkey HTTP/ imanis-node4.mydomain.com@PRIMARY.REALM"
```

2. Add similar principals for `imanis-node5.mydomain.com` by replacing node name in the preceding command.

NOTE: In case of Microsoft Windows Active Directory Kerberos server, you must create two new active directory users for each host which being added and set the service principal names for those users using the setspn command.

3. Generate a keytab file for all added principals.

```
kadmin.local -q "xst -norandkey -k /tmp/talena.keytab
$SERVICE_USER@PRIMARY.REALM
hbase@PRIMARY.REALM
$SERVICE_USER/imanis-node4.mydomain.com@PRIMARY.REALM
HTTP/imanis-node4.mydomain.com@PRIMARY.REALM
$SERVICE_USER/imanis-node5.mydomain.com@PRIMARY.REALM
HTTP/imanis-node5.mydomain.com@PRIMARY.REALM"
```

The preceding command will create the keytab file at location: `/tmp/talena.keytab`

> **NOTE**: In case of Microsoft Windows Active Directory Kerberos server, you must use ktpass command to generate keytabs for each Service Principal Names (SPN) separately and merge it with the existing `talena.keytab` on Imanis Data nodes.

4. Copy the `/etc/krb5.conf` file and the generated keytab file to Install Master node at a temporary location. For example `/tmp/krb5.conf` and `/tmp/talena.keytab` and follow the installer prompts for addnode. During the addnode process, the installer will prompt you with the following prompt:

> **NOTE**: Below steps will NOT be prompted in case of non-sudo mode of operation. talena-postinstall.sh script will take care of calling '$INSTALL_DIR/bin/talena-kerberos.sh addnode'

```
Please execute the script: $INSTALL_DIR/bin/talena-kerberos.sh addnode
on one of the cluster node with root privileges to configure ImanisData nodes
to use external KDC XXX
Press [Enter] to continue after these steps.
```

5. Follow the instructions and wait for the script to complete the execution.

```
# $INSTALL_DIR/bin/talena-kerberos.sh addnode
INFO: Checking nodes reachability
INFO: Following details required to configure ImanisData to use External KDC
keytab: with user and service principals for ImanidData nodes.
INFO: Please follow the steps under section 'Configuring Imanis Data software
using the external KDC Server in ImanisData Installation Guide to generate
keytab for newly added nodes.

Copy generated keytab to local directory

Please press [Enter] to continue :

Provide local directory path where keytab(s) downloaded from KDC
t14vm3.talena.com : /tmp
```

> **NOTE**: Replace `/tmp` with the directory path you have copied the files to.

6. On the Installer prompt, press Enter to continue the installation.

4.4.1 Setting up Call Home on Imanis Data cluster

This feature enables the Imanis Data nodes to periodically upload logs and other Imanis Data cluster metadata to a Dropbox account setup which is exclusively provided to a customer. The data helps Imanis Data support to look at any possible issues and troubleshoot them. You have to ensure that the Imanis Data nodes should have Internet connectivity and be able to connect to dropbox.com to use the Call Home feature.

The following data is collected through the Call Home feature:
1. $LOG_DIR and /tmp/talena-installer-logs from all nodes and /system data from Imanis Data file system (HDFS) which contains aggregated logs of all the jobs run on Imanis Data.
2. Machine specific info like hardware configuration (memory, cpus, and so on) number of disk drives used, amount of free space in each of them.
3. Packages installed on the system.
4. Imanis Data scheduler metadata to help understand the number of Data sources configured, type of data source, policies for backup/archival defined by user etc.
5. Imanis Data catalog.
6. hdfs/hbase configuration files.

Imanis Data does NOT copy any user data during the Call Home process. Also, the data is uploaded incrementally to minimize the amount of data uploaded to Dropbox.

4.4.1.1   Configuring parameters for Call Home

You need to enable the following configuration parameters in the tl_installer.conf file during install time to setup call home.

```
# absolute path of a secret-key file, this key will be used to
# upload logs in dropbox
tl_dropbox_authkey_file=
#-------------------------------------------------------------------------
# specify time in the form of cron expression; logs will be
# uploaded at specified time. example : "0 20 * * *" where logs
# will be uploaded everyday at 8pm
# Here * mean
# MIN HOUR DOM MON DOW
# where
# MIN Minute field 0 to 59
# HOUR Hour field 0 to 23
# DOM Day of Month 1-31
# MON Month field 1-12
# DOW Day Of Week 0-6
tl_callhome_cron_expr=0 2 * * *
#-------------------------------------------------------------------------
Specify day on which full backup of explorer will
```

```
be uploaded to dropbox through callhome.
Value should be one of the mon/tue/wed/thu/fri/sat/sun/none
NOTE : Use 'none' to disable full backup of explorer
#--------------------------------------------------------------------
tl_callhome_full_backup_day=sat
```

This is an optional step. You must specify time in the form of a cron expression; logs will be uploaded to Dropbox at a specified time. For example, "0 2 * * *" means logs will be uploaded everyday at 2 A.M.

As mentioned above, please request for the Dropbox access key file from Imanis Data Technical Support for enabling the Call Home feature.

NOTE: The data uploaded to Dropbox is accessible by you and Imanis Data support only. Nobody can access this data without the access key file. Ensure that you DO NOT share the access key file with anyone and that the key file is NOT readable by others.

4.4.2 Setting up SSL on Imanis Data cluster

The Imanis Data Installer generates both the Imanis Data KeyStore and the SSL Configuration file. You have to make sure that all nodes must have all the relevant SSL certificates. Typically, a SSL certificate encrypts a site's information and creates a more secure connection.

4.4.2.1   Imanis Data KeyStore

Imanis Data software generates Self-Signed (SSL) certificates for each node in the Imanis Data cluster. These self-signed certificates and their corresponding private keys can be found at a place called KeyStore. This is the path of the keystore:

```
${INSTALL_DIR}/conf/${HOSTNAME}.keystore
```

where ${INSTALL_DIR} is the default directory where installation is done for example, /opt/talena and ${HOSTNAME} is the hostname of the machine.

4.4.2.2   SSL Client Configuration

The Imanis Data installer generates a SSL Client configuration file in the following path:

```
${HADOOP_CONFDIR}/ssl-client.xml
```

This configuration file has key values of everything that the client needs to know about the location and password for the Imanis Data KeyStore, location and password for the Imanis Data TrustStore, and so on. Currently, the password used for Imanis Data KeyStore and Imanis Data TrustStore should be same and it must match the entry in the following:

```
${HADOOP_CONFDIR}/ssl-client.xml
```

4.4.3 Installing Nagios with Imanis Data installer

This section provides steps that need to be performed to install the Nagios software on the Imanis Data cluster or to monitor Imanis Data nodes from an external Nagios server.

4.4.3.1   About Nagios

Nagios is a powerful monitoring tool that enables enterprises to quickly identify and resolve IT infrastructure problems before they affect end-users and customers. Nagios can also be used to monitor the health of nodes in Imanis Data cluster. For more details on Nagios, please refer to www.nagios.org.

Imanis Data software supports Nagios Plugins that allow you to monitor hosts, devices, services, protocols, and applications with Nagios. Refer to Appendix G to start using the Nagios plugins that are supported by Imanis Data software.

4.4.3.2   Pre-requisites of installing Nagios

To install Nagios on Imanis Data cluster, refer to information on tl_external_nagios_server_ipaddr=

4.4.3.2.1   Installing Nagios

You can install Nagios with the Imanis Data installer in the following ways:

Option 1: When you want to configure Nagios server on Imanis Data cluster
When you want to monitor the health of the Imanis Data nodes, configure Nagios server on one of the three nodes.

- Set the following parameter in the installer configuration file to True:
  ```
  tl_nagios_enable=True
  ```

  The Imanis Data Installer will automatically configure the following:
  - Configure the Imanis Data installer to install Nagios plugins and NRPE daemon on each node in the Imanis Data cluster
  - Configure the NRPE daemon to accept the remote plugin execution requests from the Nagios server of your choice

Option 2: When you have external Nagios server

You may already have an external Nagios server in your datacenter for monitoring the health of the systems that you manage. In such a case, you may just want to add the Imanis Data cluster nodes to the existing Nagios server and start monitoring them. You can achieve this setup by doing the following:

1. Set the following parameter in the installer configuration file to True.

```
tl_nagios_enable=True
```

2. If Nagios server is already configured in your data centre and you wish to use it for monitoring Imanis Data platform, then you must set the IP address of the Nagios server in the following parameter.

```
tl_external_nagios_server_ipaddr= <ip address of Nagios server>
```

Imanis Data nagios clients will be configured to accept plugin run requests (through check_nrpe) from this IP address only through NRPE daemon.

4.4.3.3   Adding Imanis Data nodes to the external Nagios server

The Imanis Data cluster nodes can be monitored through an external Nagios server.

You have to generate configuration files for each of the Imanis Data nodes by running a script and then copy these configuration files to the external Nagios server.

In the following example, 192.168.1.201 is the external Nagios server and s2node1-centos7,s2node2-centos7,s2node3-centos7 are Imanis Data nodes.

To add Imanis Data nodes to external Nagios server, do the following:

1. Run the following script on any one of the Imanis Data nodes and then press **Enter**.

```
[hdfs@s2vm1-centos7 root] $ /opt/talena/bin/lib/talena-nagios.sh
 genconf
```

2. Copy the following configuration files to the external Nagios server: 192.168.1.201

```
/tmp/30126/s2node1-centos7.cfg
/tmp/30126/s2node2-centos7.cfg
/tmp/30126/s2node3-centos7.cfg
```

3. Restart the external Nagios Server to complete the process of adding Imanis Data nodes to Nagios server.

4.4.4 Setting up Job Notification

The Job Notification feature enables Imanis Data software to send you an email with status of all the jobs run for any given day. By default, this cron job is run daily at 00:01 AM which queries Imanis Data cluster for the list of jobs run since 00:00 hours on that given day.

As a pre-requisite, you must ensure that the Imanis Data nodes have been configured to send emails and the sendmail binary exists on your system.

To enable Job Notification feature in your Imanis Data cluster, you must set the following parameters in tl_installer.conf during installation:

```
#------------------------------------------------------------------------
# comma separated list of email addresses
# tl_daily_job_email : is used for sending daily HTML report of all job runs
#                      during the day
#------------------------------------------------------------------------
tl_daily_job_email=johndoe@example.com
```

An email, as the following, will be automatically sent to the email id mentioned in the above paragraph:

09-21-2016 12:00 HDFS_to_Imanis Status: COMPLETED
Source: HDFS-Repo
Destination: Imanis
Retention at destination: 100
End time: 05-18-2018 12:40

09-21-2016 12:12 HDFS_to_Imanis Status: COMPLETED
Source: Imanis Data
Destination: HDFS-Repo
Retention at destination: 100
End time: 05-18-2018 12:42

09-21-2016 12:57 recovery-example-no-overwrite
Status: FAILED
Source: Imanis Data
Destination: CAS-DC1
Retention at destination: 100
End time: 05-18-2018 12:49

09-21-2016 13:22 Cass_to_Imanis
Status: COMPLETED

```
Source: CAS-DC1
Destination: Imanis
Retention at destination: 100
End time: 05-18-2018 13:07


09-21-2016 13:33 Cass_from_Imanis
Status: RUNNING
Source: Imanis Data
Destination: CAS-DC1
Retention at destination: 100
End time: 05-18-2018 13:06
```

In the preceding example, Green: Jobs Completed, Blue: Jobs Running, and Red: Jobs Failed.
In rare cases when Imanis Data software is not able to generate the daily report due to some issue, the email will consist of appropriate message and also contain an attachment of the log file which would help in troubleshooting the issue.

As a pre-requisite it is expected that the Imanis Data nodes have been configured to send emails. The configuration will not be done automatically by the Imanis Data installer. Thus, you must ensure that your systems are configured to send email and the sendmail binary exists on your system. You can verify if email is configured by running the following command on shell prompt:

```
# echo "Subject: sendmail test" | sendmail -v my@email.com
```

In the preceding paragraph, my@email.com is the e-mail address you want the test email to be sent to. This sendmail command line example will send a blank email with the subject "sendmail test" to my@email.com if the test is successful.

Job Notification emails are sent from the node running UI service. To identify the node that is running UI service, run the following command:

```
# echo $API_SERVER_ADDR
```

5   Verifying Installation

Prior to using Imanis Data software, verify that it has been successfully installed. You can verify the installation on the Imanis Data cluster as well as check to see if the GUI is properly functioning by following steps mentioned in this section.

5.1 Verifying Services

1. Log on to any node in the Imanis Data cluster as $SERVICE_USER (by default this is hdfs user).
2. Run the following command. This command verifies if the installation is successfully done and all the services are running.

```
# $INSTALL_DIR/bin/talena-services.sh status all
```

If the installation is successful, the following will be displayed:

```
[hdfs@talena28 ~]$ $INSTALL_DIR/bin/talena-services.sh status all
Zookeeper daemon(s) are running on [talena27 talena29 talena28]
Namenode(s) are running on [talena27 talena29]
checking if Namenode is in safemode
Datanode(s) are running on [talena27 talena29 talena28]
JournalNode(s) are running on [talena27 talena29 talena28]
Zookeeper failover controller(s) are running on [talena27 talena29]
HDFS daemons are running
ResourceManager(s) are running on [talena28]
NodeManager(s) are running on [talena27 talena29 talena28]
JobHistory Server running on [talena28]
YARN daemons are running
Hbase Master(s) are running on [talena29 talena28]
Hbase RegionServer(s) are running on [talena27 talena29 talena28]
HBase daemons are running
Stats collection is disabled.
Postgres Server running on [talena29]
HiveMetastore is running on [talena29]
HiveServer is running on [talena29]
Oozie daemon is running on [talena28]
Scheduler shared jars are OK
Tomcat server(s) are running on [talena28]
API server(s) are running on [talena28]
UI is running @ http://talena28:8080
```

NOTE: If some daemons are NOT running, then such daemons are displayed in RED. For example:
Datanode(s) are NOT running on [talena-31]

Refer to the section Stop/Start Services to stop or start Imanis Data services.

5.2   Verifying UI

1. Log on to any node in the Imanis Data cluster as $SERVICE_USER (by default this is hdfs user).

NOTE: If you have logged in as a hdfs user before the installation was done, you must log out and login again so that shell environment is set up properly.

2. Run the following command:
   ```
   talena-services.sh status UI
   ```

3. This command displays the URL of the UI node.
   For HTTP, then the following is displayed:
   ```
   # talena-services.sh status UI
   Tomcat server(s) are running on [talena-32]
   API server(s) are running on [talena-32]
   UI is running @ http://talena-32:8080
   ```

4. For HTTPS, then the following is displayed:
   ```
   # talena-services.sh status UI
   Tomcat server(s) are running on [s4vm8-centos7]
   API server(s) are running on [s4vm8-centos7]
   UI is running @ https://s4vm8-centos7:8443
   ```

5. Open Firefox browser (Release 32 and above) on your local computer and type the URL that you received in the address bar. Imanis Data software login screen appears

6. Click the Select a user domain drop-down list, and select the Imanis Data Domain

7. Type admin in the Username and Password field.

8. Type the login credentials, and then click icon. The Imanis Data GUI appears.

NOTE: When you are logging in for the first time, you will be asked to change the default password. This is a mandatory, one-time activity only without which the Imanis Data GUI cannot be accessed.

6 Managing the Imanis Data Cluster

This section provides basic command line options that you can use to manage the Imanis Data cluster. By using these options, you can add more nodes to the existing cluster, upgrade a cluster, or uninstall Imanis Data software. Refer to Appendix C: Parameters for Imanis Data Installation to learn about advanced ways to manage the Imanis Data cluster.

> **NOTE**: If any custom changes have been done in any config files (For example, `hdfs-site.xml` or `hadoop-env.sh etc`) after fresh installation, then the changes have to be done again after running any of the following utilities like upgrade, addnode, cluster_manager, migrate, decommission.

6.1 Getting to know Imanis Data Version

The following section discusses the procedure to determine the version or release number of Imanis Data software.

To know the Imanis Data software version, do the following:

1. Open Firefox browser (Release 32 and above) on your local computer and type the URL that you received in the address bar. Imanis Data software login screen appears.
2. Click the **Select a user domain** drop-down list, and select the type of user. For example, Imanis Data or Domain (see screenshot below):

Username
admin

Password
••••••••

Domain
Imanis Data

Sign In

Imanis Data version4.1   © 2018 Cohesity, Inc.

> **NOTE**: When you are logging in for the first time, you will be asked to change the default password. This is a mandatory, one-time activity only without which the Imanis Data GUI cannot be accessed.

3. Click the **About** option in Main Menu and then view the information regarding Imanis Data software displayed on the dialog box.
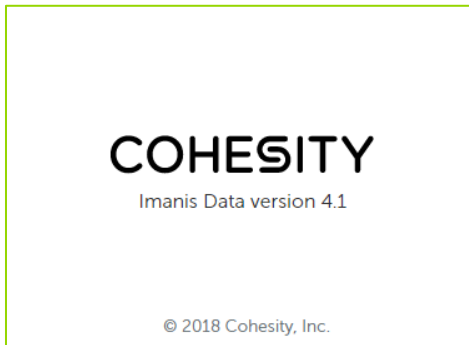
6.2   Viewing Cluster Name from Imanis Data UI

At times, Imanis Data software may be installed on more than one cluster. In this case, it becomes important to identify and know the cluster name from the Imanis Data software UI itself.

In the top right corner, look out for the name of the cluster to identify and know where Imanis Data software is installed. This feature saves the back and forth time that is wasted in finding out the cluster name.

6.3   Adding More Nodes to an Existing Cluster

You can add node(s) to the existing Imanis Data cluster through both the SSH mode and the NO-SSH mode. Before the actual addnode procedure, you must first run pre-checks on the nodes that you want to add to the existing nodes.

6.3.1 SSH Mode

This section describes the steps to run pre-check and add node(s) to an existing cluster in SSH mode.
1. Go to $INSTALL_DIR/$VERSION
   (These variables are set in /etc/talena/talena-env.sh) on the existing node.
2. Run prechecks for the nodes to be added:

```
./precheck_talena.sh --assumeyes --addnode <comma separated list of
nodes to be added>
```

> **NOTE**: The interactive option can be used instead of `--assumeyes` to enable the interactive prechecks execution process.

3. Ensure that all the prechecks for the nodes to be added are successful. Then continue with the procedure of adding the node to an existing cluster by using the following command:

```
./install_talena.sh -a <comma separated list of nodes to be added>
```

6.3.2 NO-SSH Mode

This section describes the steps to run pre-check and add node(s) to an existing cluster in NO-SSH mode.

1. Extract the installer separately on each new node at a location or path which is same on all the existing nodes.
2. Type **No** for NO-SSH Mode in the following prompt:

```
INFO - Upgrade will be performed on nodes: s2vm1-centos7,s2vm3
 centos7,s2vm2-centos7
Do you want to continue [y/n]:
```

3. Go to $INSTALL_DIR/$VERSION
   (These variables are set in /etc/talena/talena-env.sh) on the existing node.
4. Run prechecks for the node(s) to be added:

```
./precheck_talena.sh --assumeyes --addnode <comma separated list of
 nodes to be added>
```

> **NOTE**: The interactive option can be used instead of `--assumeyes` to enable the interactive prechecks execution process.

5. Ensure that all the prechecks for the nodes to be added are successful. Then continue with the procedure of adding the node to an existing cluster by using following command:

```
./install_talena.sh -a <comma separated list of nodes to be added>
```

> **IMPORTANT:** The number of datanode directories must be identical to the number of datanode directories used at the time of installation. In other words, if the tl_datanode_dirlist was set to tl_datanode_dirlist=/mnt/disk1,/mnt/disk2,/mnt/disk3 in tl_installer.conf file then, you must ensure

that three disks are mounted under /mnt as /mnt/disk1, /mnt/disk2 and /mnt/disk3 respectively on the new nodes to be added to the cluster.

6.4   Setting up Imanis Data Licensing

You are required to have a license key during installation. If you do not have the Imanis Data license key, please contact the Imanis Data Technical Support Team at support@imanisdata.com and the Imanis Data Sales Team at sales@imanisdata.com.

IMPORTANT: Ensure that you have the Imanis Data license copied and stored at a path with read access to all users example : `/tmp/licenseKey`. The licence utility expects this file to be readable for service user which is created during the installation if not already present. If the licence file is not accessible to this user during installation, then the licence installation could fail.

6.4.1 Setting up licensing during installation

Ensure that you have placed the license key in the absolute path for example /tmp/licensekey. This path must be mentioned in the configuration file. This is a mandatory step without which the installation process will not proceed.

To set up licensing during installation, do the following:
1. After extracting the Imanis Data installer, modify the following parameter in the configuration file conf/tl_installer.conf.

```
#------------------------------------------------------------------------
----
Absolute Path to the licence key file
#------------------------------------------------------------------------
----
tl_licence_key_path=/tmp/licenseKey
```

2. Run the Imanis Data Installer.

6.4.2 Setting up licensing after installation

Ensure that you have placed the license key in the absolute path for example /tmp/licensekey.

To set up licensing after installation, do the following:

1. Log on to any node in the Imanis Data cluster as $SERVICE_USER (by default this is hdfs user).
2. Run the command cd $INSTALL_DIR/$TALENA_VERSION.
3. Run ./talena-licence.sh --setup --in /tmp/licenseKey.
   where /tmp/licenseKey is the filepath for license key. Ensure that the service user is able to access this filepath.

6.4.3 Upgrading license key

This section describes the steps of upgrading an evaluation license key to production license key to extend the expiration date of the license.

To upgrade license key, do the following:
1. Log on to any node in the Imanis Data cluster as $SERVICE_USER (by default this is hdfs user).
2. Run command cd $INSTALL_DIR/$TALENA_VERSION
3. Run the following command to upgrade license:
   ./talena-licence.sh --upgrade --in /tmp/licenseKey.
4. Run the following command to check license status: /talena-licence.sh --report.

```
---------------------------------------------------------------------
                    IMANIS LICENCE INFORMATION
---------------------------------------------------------------------
Customer Name      : Amex Cluster 1
Expires On         : 13-Sep-2018
Installed On       : 14-Aug-2018
Licence Type       : Evaluation
---------------------------------------------------------------------
```

6.5   Configuring system using the Cluster Management utility

You can configure additional components or change system configuration after installation through the configuration management utility.

Only one component must be configured at one time through any one node of the system.
The cluster management utility is available in $INSTALL_DIR/<VERSION>
where:

INSTALL_DIR is the location where Imanis Data software is installed and <VERSION> is the version of the Imanis Data software.

HELP for the Cluster Management Utility

```
usage : cluster_manager.sh [--createconfig | --reportconfig | --reconfigure
--createconfig : Generates the configuration file from existing configuration
--reportconfig : Reports the configuration values from existing configuration
--reconfigure  : The utility reads the configuration file from the path
installer/conf. Checks if values of parameters have been changed by comparing
with the existing install and confirm the addons to be configured on the
system
```

6.5.1 Configuring Yarn Nodemanager Localdirs component

This section explains the procedure of configuring yarn nodemanager localdirs component to the existing installation.

To configure the Yarn Nodemanager Localdirs component, do the following:

1. Get the existing installation configuration (./cluster_manager.sh --createconfig).
   This command creates and populates files tl_installer.conf and .tl_installer_advanced.conf with existing parameters.

2. Modify the following parameter in conf/tl_installer.conf file.
   ```
   tl_yarn_nodemgr_localdirs=/data2/physical/yarn_nodemgr_localdirs
   ```

3. Run prechecks for cluster_manager.sh utility using command:
   ```
   ./precheck_talena.sh --assumeyes --reconfigure
   ```

4. Reconfigure the component using command:
   ```
   ./cluster_manager.sh --reconfigure
   ```

The reconfigure utility will automatically identify the difference in the existing configuration parameters and the modified configuration parameters. Then it will prompt the user to proceed with the system change.

6.5.2 Configuring Datanode Directories component

This section explains the procedure of configuring datanode directories component to the existing installation.

To configure the Datanode Directories component, do the following:

1. Get the existing installation configuration (./cluster_manager.sh --createconfig). This command creates and populates files tl_installer.conf and .tl_installer_advanced.conf with existing parameters.
2. Modify the following parameter in conf/tl_installer.conf file.

```
tl_datanode_dirlist
```

3. Run prechecks for cluster_manager.sh utility using command:

```
./precheck_talena.sh --assumeyes --reconfigure
```

4. Reconfigure the component using command:

```
./cluster_manager.sh --reconfigure
```

The reconfigure utility will automatically identify the difference in the existing configuration parameters and the modified configuration parameters. Then it will prompt the user to proceed with the system change.

6.5.3 Configuring HTTPS component

This section explains the procedure of configuring HTTPS support component through the Cluster Management utility.

To configure HTTPS component, do the following:
1. Get the existing installation configuration (./cluster_manager.sh --createconfig).
    This command creates and populates files tl_installer.conf and .tl_installer_advanced.conf with existing parameters.
2. Modify the following parameter in conf/tl_installer.conf file.

```
tl_enable_https_webui=true
#-----------------------------------------------------------------------
---
# Absolute Path to the ssl certificate
#-----------------------------------------------------------------------
---
tl_https_certificate_path=
#-----------------------------------------------------------------------
---
# Absolute Path to the private key
#-----------------------------------------------------------------------
---
tl_https_key_path=
```

3. Run prechecks for cluster_manager.sh utility using command:

```
./precheck_talena.sh --assumeyes --reconfigure
```

4. Reconfigure the component using command:

```
./cluster_manager.sh --reconfigure
```

5. Imanis Data supports PEM certificate for tl_https_certificate_path. The users can verify   the usage by using the following command:

```
[root@s4vm9-centos7 certs]# file server.crt server.key
server.crt: PEM certificate
server.key: PEM RSA private key
```

6.5.4 Configuring Nagios component

This section explains the procedure of configuring Nagios component through the Cluster Management utility.

To configure Nagios component, do the following:
1. Get the existing installation configuration (./cluster_manager.sh --createconfig).
2. Modify the following parameter in conf/tl_installer.conf file.

```
#----------------------------------------------------------------------
---
Enable/Disable Nagios on Talena.
In case of External nagios server, provide
tl_external_nagios_server_ipaddr
Allowed values : True/False
#----------------------------------------------------------------------
---

tl_nagios_enable=True

#----------------------------------------------------------------------
---
If you have a Nagios server already configured in your data centre,
and you wish to use the same for monitoring Talena platform, then
provide
IP address of the same. Talena nagios clients will be configured to
accept plugin run requests through check_nrpe) from this IP address
only through NRPE daemon.
#----------------------------------------------------------------------
---

tl_external_nagios_server_ipaddr=
```

3. Run prechecks for cluster_manager.sh utility using command:

```
./precheck_talena.sh --assumeyes --reconfigure
```

4. Reconfigure the component using command:

```
./cluster_manager.sh --reconfigure
```

6.5.5 Configuring Kerberos component

This section explains the procedure of configuring Kerberos component through the Cluster Management utility.

To configure Kerberos component, do the following:
1. Get the existing installation configuration (./cluster_manager.sh --createconfig).
2. Modify the following parameters in conf/tl_installer.conf file.

```
#----------------------------------------------------------------------
---
Enable/Disable Kerberos on Talena.
Set it to True and provide tl_kdc_realm_name
#----------------------------------------------------------------------
---
tl_kerberos_enable=True
#----------------------------------------------------------------------
---
If you would like to use an existing Kerberos realm for Talena, then
provide
the name of the server where KDC is already running. Else leave this
blank
and one of the nodes in the Talena cluster will be setup as KDC server
automatically
#----------------------------------------------------------------------
---
tl_external_kdc_server=
#----------------------------------------------------------------------
---
Realm name.
In case of external KDC server, provide the existing realm name.
In case of internal KDC server, provide the new realm name to be
created.
#----------------------------------------------------------------------
---
tl_kdc_realm_name=
```

3. Run prechecks for cluster_manager.sh utility using command:

```
./precheck_talena.sh --assumeyes --reconfigure
```

4. Reconfigure the component using command:

```
./cluster_manager.sh --reconfigure
```

6.5.6 Configuring Log directory component

This section explains the procedure of configuring Log directory component through the Cluster Management utility.

To configure Log directory component, do the following:
1. Get the existing installation configuration (./cluster_manager.sh --createconfig).
2. Modify the following parameter in conf/tl_installer_advanced.conf file.

```
tl_log_dir
```

3. Run prechecks for cluster_manager.sh utility using command:

```
./precheck_talena.sh --assumeyes --reconfigure
```

4. Reconfigure the component using command:

```
./cluster_manager.sh --reconfigure
```

6.5.7 Configuring Service Email Alert component

This section explains the procedure of configuring Service Email Alert component through the Cluster Management utility.

To configure Service Email Alert component, do the following:
1. Get the existing installation configuration (./cluster_manager.sh --createconfig).
2. Modify the following parameter in conf/tl_installer.conf file.

```
tl_service_alert_email=johndoe@example.com
```

3. Run prechecks for cluster_manager.sh utility using command:

```
./precheck_talena.sh --assumeyes --reconfigure
```

4. Reconfigure the component using command:

```
./cluster_manager.sh --reconfigure
```

6.5.8 Configuring Job Email Alert component

This section explains the procedure of configuring Job Email Alert component through the Cluster Management utility.

To configure Job Email Alert component, do the following:
1. Get the existing installation configuration (./cluster_manager.sh --createconfig).
2. Modify the following parameter in conf/tl_installer.conf file.

```
tl_daily_job_email= johndoe@example.com
```

3. Run prechecks for cluster_manager.sh utility using command:

```
./precheck_talena.sh --assumeyes --reconfigure
```

4. Reconfigure the component using command:

```
./cluster_manager.sh --reconfigure
```

6.5.9 Configuring Tcollector component

This section explains the procedure of configuring the Tcollector component through the Cluster Management utility.

To configure Tcollector component, do the following:
1. Get the existing installation configuration (./cluster_manager.sh --createconfig).
2. Modify the following parameter in conf/tl_installer_advanced.conf file.

```
tl_enable_system_stats_collection= True
```

3. Run prechecks for cluster_manager.sh utility using command:

```
./precheck_talena.sh --assumeyes --reconfigure
```

4. Reconfigure the component using command:

```
./cluster_manager.sh --reconfigure
```

6.5.10 Configuring Recovery Job Alert component

You can configure the number of days after which the recovery job alert notification band will be displayed on the Imanis Data GUI.

While default the value of the parameter is 60 days, typing 0 will indicate the system to disable recovery job alert notifications.

To configure Recovery Job Alert component, do the following:
1. Get the existing installation configuration (./cluster_manager.sh --createconfig).
2. Modify the following parameter in conf/tl_installer_advanced.conf file.

```
tl_recovery_job_alert_days=60
```

3. Run prechecks for cluster_manager.sh utility using command:

```
./precheck_talena.sh --assumeyes --reconfigure
```

4. Reconfigure the component using command:

```
./cluster_manager.sh --reconfigure
```

6.5.11 Configuring System Stats component

You can enable system usage statistics of Imanis Data by configuring this component.

To configure System Stats component, do the following:
1. Get the existing installation configuration (./cluster_manager.sh --createconfig)
2. Modify the following parameter in conf/tl_installer_advanced.conf file.

```
tl_enable_system_stats_collection=True
```

3. Run prechecks for cluster_manager.sh utility using command:

```
./precheck_talena.sh --assumeyes --reconfigure
```

4. Reconfigure the component using command:

```
./cluster_manager.sh --reconfigure
```

6.5.12 Configuring Callhome component

You can enable configure this component to enable Callhome.

To configure Callhome component, do the following:
1. Get the existing installation configuration (./cluster_manager.sh --createconfig)
2. Modify the following parameter in conf/tl_installer.conf file.

```
tl_callhome_enable=
tl_dropbox_authkey_file=
```

3. You can also modify the following optional parameters in conf/tl_installer_advanced.conf file.

```
tl_callhome_cron_expr=0 2 * * *
tl_callhome_alert_email= johndoe@example.com
tl_callhome_explorer_dir=
tl_callhome_full_backup_day=sat
```

4. Run prechecks for cluster_manager.sh utility using command:

```
./precheck_talena.sh --assumeyes --reconfigure
```

5. Reconfigure the component using command:

```
./cluster_manager.sh --reconfigure
```

6.6   Migrating Imanis Data Services

You may want to consider migrating Imanis Data services from one node to another for the following reasons:

- Hardware or Software Maintenance: To support the periodic hardware or software maintenance, you may have to shut down one of the nodes in the Imanis Data cluster for a short period. During this time, you can migrate Imanis Data services running on a given node to other nodes in the cluster so that Imanis Data cluster continues to run without any downtime
- Load balancing of Imanis Data cluster: To achieve load balancing across the cluster, you may want to move certain critical services running on a specific node to other nodes in the cluster
- Scalability of Imanis Data cluster: To make Imanis Data cluster more scalable, you may want to move the critical process like namenode to a dedicated hardware with better specifications after you have completed the initial installation. This task can be executed by migrating the specific service to a new dedicated machine for better performance

However, it is essential to know when to migrate services.
- **Migrate the Imanis Data services if**:

- o The crashed node takes time to restart
  - o The services on the crashed nodes are critical for the functioning of the cluster
- **DO NOT Migrate the Imanis Data service if**:
  - o A node can be quickly rectified and restarted following a crash
  - o The nodes were hosting highly available services, and then migration may not be needed if there is a chance of restarting the crashed node in a reasonable time

We can categorize the services into different types for handling the migration:

1. **Services which have HA with auto failover configured: Namenode, HMaster**
   When a node hosting namenode or HBase master crashes, cluster functionality is not affected because of the HA configuration with auto failover feature. This ability ensures that another instance of the namenode and/or hmaster running on another node takes over as active node and serves the requests. However, these services become prone to SPOF (Single Point of Failure) after this point as only one instance of this type is running. Hence, it is important to bring up the failed namenode/hmaster on an existing node or on a new node so as to make these services highly available again.

2. **Services that are run as an ensemble of servers: Zookeeper, Journal Nodes**
   In the Imanis Data cluster, a minimum 3 instances of Zookeeper and Journal Nodes are needed. Due to this redundancy, these services also provide HA. If you lose one instance of these services, it does NOT affect the functionality of the cluster. However, after losing one instance, it is recommended to migrate the particular service on another node or on a new node as soon as possible so the high availability of this service is maintained for future.

3. **Services which do not have HA or redundancy: Resource Manager, Oozie, web UI**
   For these services, we do not have HA configured yet. Hence, if there is a node (hosting any of these services) crash, then the services become temporarily unavailable affecting the cluster functionality. Hence it is recommended that we migrate these services quickly to another node or new node in the cluster.

4. **Slave services: datanode, node manager and hbase region server**
   These services also provide HA because usually you will have multiple instances of these processes running on the cluster (minimum of 3). Depending on the number of replication factor that is used for the data on the Imanis Data cluster, 1 slave node failure can be tolerated. For example, for HDFS backup, the default replication factor is 2. So even if one slave node fails, the cluster functionality will not be affected.

NOTE: When you start services migration, all the Imanis Data services will be stopped. Thus, any backup or recovery workflows running will be aborted.

6.6.1   Performing Migration of Imanis Data Services

To migrate Imanis Data services from one node to another, you will be prompted to login as an admin user. Once the admin user is entered, you will have to select Password or Private Key authentication method followed by the respective input. The migration process will start once these inputs are entered.

You need to use the Imanis Data installer tool with a different set of arguments to perform the migration.

NOTE: The cluster management utility is available in $INSTALL_DIR/$TALENA_VERSION where: INSTALL_DIR is the location where Imanis Data software is installed and <TALENA_VERSION> is the version of the Imanis Data software.

The following use cases describe how to use the CLI (command line interface) to migrate services in a Imanis Data cluster.

Scenario 1: A '5-node' Imanis Data cluster with the following services are placed and are running on the cluster after successful installation.

node1	namenode1, web UI daemons, oozie, slaves
node2	namenode2, Hmaster1, slaves
node3	Hmaster2, Resource Manager, Job History Server, slaves
node4	slaves
node5	slaves

Example 1: Node1 crashes due to hardware failure and it is not likely to come up for a few days. In such a case, the user should consider migrating the services hosted on node1 (namenode1, web UI, oozie) to another node or a new node in the cluster.

NOTE: Currently migration of slave services (data node, region server, node manager) is NOT supported. You can only add a new node with slave services configured into the Imanis Data cluster.

1. Change directory $INSTALL_DIR/$TALENA_VERSION where $INSTALL_DIR is installation directory and $TALENA_VERSION is the version of Imanis Data and migrate all services from node1 to other nodes in the cluster.

```
$INSTALL_DIR/$TALENA_VERSION/install_talena.sh --migrate all --source
node1
```

In this case, the destination nodes are chosen automatically by the migration tool based on the resources available in the cluster.

2. Type appropriate inputs for the following system prompts:

```
Migration mode is set to offline.

All Talena services will be stopped (if running) before starting
migration.
This will stop currently running jobs.
Press 'c' continue or 'a' to abort (c) : c
--------------------------------------------------------------
Enter the admin user used for installing Talena : root
Select authentication mode:
1. Password
2. Private key
Enter your choice (1): 1
Enter root password:
Re-enter root password:
--------------------------------------------------------------
```

3. Type Y to continue and No to discontinue services migration.

```
Do you want to continue [y/n] (y) : (y)
-----------------------------------------------------------
INFO - Stopping Talena services, please wait....
```

- Change directory to location where the installer was extracted and migrate all services from node1 to node2, node3 only

```
$INSTALL_DIR/$TALENA_VERSION/install_talena.sh --migrate all --
source   node1 --target node2,node3
```

- Change directory to location where the installer was extracted and add a new node called node6 in the cluster and migrate services to this new node.

```
$INSTALL_DIR/$TALENA_VERSION/install_talena.sh --addnode node6 --
service none
```

```
$INSTALL_DIR/$TALENA_VERSION/install_talena.sh --migrate all --
source   node1 --target node6
```

- Change directory to location where the installer was extracted and add a new node called node6 with slave nodes on them and then migrate services from failed node to the new node

```
$INSTALL_DIR/$TALENA_VERSION/install_talena.sh --add node  node6
--service slaves
```

```
$INSTALL_DIR/$TALENA_VERSION/install_talena.sh --migrate all --
source   node1 --target node6
```

- Change directory to location where the installer was extracted and migrate Oozie and UI services from node1 to other nodes in the cluster

```
$INSTALL_DIR/$TALENA_VERSION/install_talena.sh --migrate ui,
oozie --source   node1
```

NOTE: When you run migrate -all (migrate all services from a node); all the services except slave services (datanode, nodemanager, hbaseregion server) are migrated. To remove such a node from Imanis Data cluster, you must first explicitly run decommission to gracefully remove the data from that node. For more information, refer to the section on [Decommissioning Nodes from an Existing Cluster](#).

Example 2: Node5 has crashed. You want to replace node5 with a new node configured with slave nodes running on it and decommission node1 (as it will no longer come up).

Change directory to where the installer was extracted and run the following script:

```
$INSTALL_DIR/$TALENA_VERSION --addnode node6 --service slaves
```

```
$INSTALL_DIR/$TALENA_VERSION --decommission node1
```

Example 3: Preventing a dead node from joining the Imanis Data cluster
If a dead node restarts after some time, then depending on whether you have made any changes to the cluster or not, the behavior will be different.

Scenario 2: Changes are made to the cluster
A few changes are made to the cluster while the node was down. The changes are defined as:
- If some (or all) services were migrated from the dead node or
- If a new node was added to the cluster or
- If some nodes were decommissioned from the cluster

In this case, when the dead node restarts, it will be stopped from starting any Imanis Data services. The dead node will not be allowed to join the cluster. In this case, the user is expected to add this node back to the cluster using add node option.

Change directory to where the installer was extracted and run the following script:

```
$INSTALL_DIR/$TALENA_VERSION/install_talena.sh --addnode deadnode
```

In the above command, 'deadnode' is name of the node.

Scenario 3: No changes were made to the cluster

In this scenario, one can add the node back to the cluster by running the following command:

```
$INSTALL_DIR/$TALENA_VERSION/bin/talena-services.sh start all
```

6.7   Decommissioning Nodes from an Existing Cluster

You can decommission nodes from the existing Imanis Data cluster by running the installation script with the –d option. This operation is supported for slave nodes only. Slave nodes are nodes which host the Hadoop services: data node, region server, and node manager.

If a node is decommissioned, other services running on the node are NOT migrated automatically. For example, if a node is running Oozie and UI services along with the slave services (datanode, regionserver and nodemanager), then decommissioning process will only move the slave services from the node. To migrate Oozie and UI services from the node, you must run the installer with the –m (migrate) option.

The process of decommissioning a slave node does NOT happen instantaneously as the process is mainly dependent on the size of the data available on the node that you are trying to decommission.

WARNING: When a particular node is decommissioned, large amounts of data is moved between the other nodes in the Imanis Data cluster causing network latency. Thus, it is advisable that decommission should be performed only when it is absolutely necessary.

NOTE: The cluster management utility is available in $INSTALL_DIR/$TALENA_VERSION
where: INSTALL_DIR is the location where Imanis Data software is installed and <TALENA_VERSION> is the version of the Imanis Data software.

To decommission nodes, do the following:
1. Log on as a service user to the one of the Imanis Data cluster nodes.
2. Run the following command to view a list of slave nodes:

```
echo $SLAVE_NODES
```

3. A list of slave nodes will be displayed. For example:

```
node1 node2 node3
```

4. Identify the nodes that you want to decommission and run the following command from the directory where the installer was extracted:

```
/install_talena.sh -d node1, node2
```

To decommission nodes if Kerberos is enabled:
If Kerberos is enabled on the Imanis Data cluster, it is recommended to delete Kerberos principals after decommissioning or removing a node from the Imanis Data cluster. If using an external KDC Server, delete Kerberos principals from the KDC Server after decommissioning or removing a node from the Imanis Data cluster.

```
hdfs/<hostname>@<realm_name>
HTTP/<hostname>@<realm_name>
```

In the preceding example, replace the word <hostname> with the name of the node that you are trying to decommission from the Imanis Data cluster.

6.7.1 Verifying Decommissioning of Nodes

The process of decommissioning a slave node does not happen instantaneously as it is mainly dependent on the size of the data available on the node that you are trying to decommission.

To verify decommissioning of nodes, do the following:
1. Log on as a service user.
2. Run the following command to check the status all the decommission process.
3. `${INSTALL_DIR}/bin/talena-decommission.sh status`

4. The following appears on the page:

```
Nodes where decommission was initiated [node1 node2]
Data node decommission status
Hostname: node1
Decommission Status : Decommission in progress
Hostname: node2
Decommission Status : Decommissioned
Region server decommission completed on nodes: [ node1]
Enabling Region Load Balancer
```

```
Nodes where decommission was initiated [node1 node2 ]
Data node decommission status
Hostname: node1
Decommission Status : Decommissioned
```

```
Hostname: node2
Decommission Status : Decommissioned
Region server decommission completed on nodes: [node1 node2]
Enabling Region Load Balancer : Done.
```

When the decommissioning process is complete the following message is displayed:

IMPORTANT: Do not run any other installer commands such as addnode or upgrade until the decommissioning process is complete.

6.8   Upgrading Imanis Data cluster

This section describes the procedure for upgrading Imanis Data software from an earlier release to the latest one.

IMPORTANT:
Upgrade is supported from 4.1 and 4.2 releases to 4.3 release. For customers using Cohesity Storage: Upgrades from older releases are not supported and will require a fresh installation of the cluster.
For customers not using Cohesity Storage: Upgrades from older releases will require an upgrade first to 4.1/ 4.2 version and then to 4.3 release.

6.8.1   SSH Mode

This section describes the procedure to upgrade Imanis Data software through the SSH Mode.

To upgrade Imanis Data software, do the following:
1.   Stop the services.
     Before the upgrade, ensure that no jobs are running on the cluster and stop the services using the following command:
```
${INSTALL_DIR}/bin/talena-services.sh stop all
```

     If you have Imanis Data Web UI session open, ensure that you save the changes and logout.
2.   Create a directory named "Imanis Data_4.3.0" and cd to this particular directory.
3.   Download the Imanis Data installer ./installer_self_extract_4.3.0-RELXXX.sh into this directory. Where XXX is the release number of the build. For example, REL136.
4.   Unpack the installer self extract

The installer self extract must be extracted at a location that was not used by installer self extract of earlier version or release. This step ensures that the new version of the software does not overwrite the previous version of the installer copy.

When you run the self extract script, the system will prompt you for SSH/NO-SSH. If you have done the installation using SSH, then select SSH or else select NO-SSH.

```
# ./installer_self_extract_4.3.0-RELXXX.sh
```

Where XXX is the release number of the build. For example, REL136.

The following is displayed on the screen. Type Yes to continue the installation with SSH mode.

```
Talena installation is supported in two modes.
SSH: tadmin user can ssh to nodes and execute commands.
NO-SSH: SSH access is not allowed for tadmin.
Nodes will communicate through Remote Procedure Calls.

Do you want to continue with SSH mode installation ? (yes/no) : yes
```

5. Run the upgrade script from the installer directory and then run the following command to start the upgrade process.

```
cd 4.3.0<buildnumber>
./install_talena.sh -u
```

6. The installer will prompt you for admin user to be used. The last admin user used is printed in the (). For example, (root). If this is correct, then press Enter to continue. The following is displayed on the screen:

```
Setting up Logs

Log file name /tmp/talena-installer-logs/talena_installer.log_2018-03-
05-06-09-19

Enter the admin user (root):
```

Verify that the admin user is present on all nodes and then continue with the installation procedure.

7. Select the authentication mode that you want to use for connecting to Imanis Data nodes:

```
Select authentication mode:
1. Password
2. Private key

Enter your choice (1): 1

Enter root password:
Re-enter root password:
```

```
INFO - ------------------------------------------------------------
--------
```

8. Press Enter. This is not applicable to clusters where install master (the node from where you are running the installer is one of the nodes) in the cluster itself:

```
Enter hostname of one of the nodes from Talena cluster.
Please make sure that provided node is reachable from current
node(s2vm1-centos7) :
```

9. Verify that the nodes listed correctly and then press Enter:

```
INFO - Upgrade will be performed on nodes: s2vm1-centos7,s2vm3-
centos7,s2vm2-centos7
Do you want to continue [y/n] (y):
```

10. Verify if services have started by running the following command:

```
${INSTALL_DIR}/bin/talena-services.sh status all
```

All the services should be displayed as up and running.

11. Open Firefox browser on your local computer and type the URL that you received in the address bar. Imanis Data software login screen appears.

12. Click the **Select a user domain** drop-down list, and select the type of user. For example, Imanis Data or Domain (see screenshot below):

13. Type the login credentials, and then click ➡ icon.
14. Press Ctrl+F5 in the browser after you have upgraded Imanis Data software and then proceed to do your tasks in the software. This will clear cache and stale data for the portal.

6.8.2   NO-SSH Mode

This section describes the procedure to upgrade Imanis Data software through the NO-SSH Mode.

To upgrade Imanis Data software, do the following:
1. Stop the services. Before you start the upgrade process, ensure that jobs are NOT running on the cluster. You can stop the services using the following command:
```
${INSTALL_DIR}/bin/talena-services.sh stop all
```

   If you have Imanis Data Web UI session open, ensure that you save the changes and logout.

2. Download and extract the installer in the path from where the installation is performed. You must extract the installer on all nodes. The path where the installer needs to be extracted can be found by running the following command as a service user:
```
${INSTALL_DIR}/bin/talena-services.sh stop all
```

   For example, if your install directory is /home/installation/talena, then new installer has to be extracted at /home/installation

3. Run the following script:
```
./installer_self_extract_4.3.0-RELXXX
```

   Where XXX is the release number of the build. For example, REL136.

4. Run the upgrade script from the installer directory and then run the following command to start the upgrade process.
```
cd talena/4.3.0-RELXXX
./install_talena.sh -u
```

5. The installer will prompt you for admin user to be used. The last admin user used is printed in the (). For example, (root). If this is correct, then press enter to continue. The following is displayed on the screen:
```
Setting up Logs
```

```
Log file name /tmp/talena-installer-logs/talena_installer.log_2018-03-
05-06-09-19

Enter the admin user (root):
```

Verify that the admin user is present on all nodes and then continue the installation procedure.

6. Verify that the nodes listed correctly and then press Enter:

```
INFO - Upgrade will be performed on nodes: s2vm1-centos7,s2vm3-
centos7,s2vm2-centos7
Do you want to continue [y/n]:
```

7. Type No for NO-SSH Mode and Type Yes for SSH Mode.
   Verify if services have started by running the following command:

```
${INSTALL_DIR}/bin/talena-services.sh status all
```

All the services should be displayed as up and running.

8. Open Firefox browser on your local computer and type the URL that you received in the address bar. Imanis Data software login screen appears.

9. Click the **Select a user domain** drop-down list and select the type of user. For example, Imanis Data or Domain (see screenshot below):

10. Type the login credentials, and then click icon.

11. Press Ctrl+F5 in the browser after you have upgraded Imanis Data software and then proceed to do your tasks in the software.

> **NOTE**: In case of NO-SSH mode (precheck/fresh installation/addnode/upgrade) the tar extraction must be done with admin user only.

6.8.3 Post Upgrade

All existing Hadoop Data Repositories must be re-discovered.

6.9   Patching Imanis Data Software

This section describes the procedure of installing a patch for Imanis Data software.

To install a patch, do the following:
1. Download the patch PATCH-00.tar.gz
2. Copy the file into a temporary location. For example, /tmp
3. Run the patch application utility as SERVICE user (hdfs)

```
$INSTALL_DIR/bin/talena-patch-apply.sh -f /tmp/PATCH-00.tar.gz
```

4. Type the patch number to be applied when the system prompts.

> **NOTE**: If the following prompt is displayed on the screen during the upgrade process,
>
> ==
> **INFO - Starting hive services.**
> **ERROR - Failed to execute: $ INSTALL_DIR/bin/talena-services.sh -d -s start hive on node: s2vm1-centos7**
>
> **Failed to start hive service.**
> **Reason : Check installer log for command output.**
> **Command executed : $ INSTALL_DIR/bin/talena-services.sh -d -s start hive**
>
> **Correct the problem manually and press (c) to continue installation.**
> **To abort installation and exit installer press (a)**
> **Enter your choice (c/a) c :**
> ==
>
> Then, please run the following command on another terminal:

> **# talena-services.sh status hive**
> **HiveMetastore is running on [s2vm3-centos7]**
> **HiveServer is running on [s2vm3-centos7]**
>
> If the status command shows hive as running, then this is a false alarm. On some systems, it takes more time to come up than expected. You can safely continue with the upgrade process by pressing Enter.
> If the status shows either HiveMetastore or Hiveserver as NOT running, then please contact the Imanis Data Technical Support for further investigation.

> **NOTE**: Prior to applying the patch, the system copies files on all the nodes. This procedure may take considerable time depending on the size of the files in the PATCH.

The following information is displayed:

```
----------------------------------------------------------------------
------
          Patches available for Application / Rollback
----------------------------------------------------------------------
------
The patch number is         : PATCH-40
Patch dependency            :
Applied (Yes/No)            : No
Description :
PATCH fixes issues with LDAP user search
Please select patch number from above options : PATCH-40
```

You only have to enter the patch number. For example, 40 as highlighted in the preceding note.

The applicator performs various steps which may take considerable time. Your patience is appreciated.

5. Verify if the patch is applied to the system once the patch is successfully applied:

```
[hdfs@s2vm3-centos7 bin]$ $INSTALL_DIR/bin/talena-patch-apply.sh -i
Warning : -i option specified ignoring other options
-----------------------------------------------------------------------
-------
          Patches available for Application / Rollback
-----------------------------------------------------------------------
-------
The patch number is     : PATCH-40
Patch dependency  :
Applied (Yes/No)  : Yes
Description       :
```

6.10 Viewing Imanis Data Job Stats

You can use the imanis-job-stats utility to get all the information for jobs run on Imanis Data cluster. Imanis-job-stats utility will create multiple csv files as per as per the application type and job type and will contain information of Job name, Nominal Time, Status, Start Time, End Time for every job instance of a job.

If -outDir is not specified, CSV files will be available under $LOG_DIR. Full path for CSV files will be printed by the utility on the page.

To view Imanis Data Job Stats, do the following:

1. Open the command prompt.
2. Type the following command:

```
$ INSTALL_DIR/tools/imanis-job-stats.sh
```

3. Type the following command to open the help content for the Imanis Job Stats utility. Once the help content is displayed on the screen, use the options listed in the help content to get the information in stats that you need.

```
$ INSTALL_DIR/tools/imanis-job-stats.sh -h
```

4. The following help content is displayed on the screen:

```
-all               Get stats of all job instances from all data
                   repositories with startDate a month back and endDate
                   as today for all jobTypes and applicationTypes

-appType <arg>     Application type for job.Valid app types are {HDFS,
                   HIVE, HBASE, GLACIER, S3, AZURE, SYS, CASSANDRA,
```

```
                        MONGODB, COUCHBASE}

-dataSource <arg>     Specifies Data Source.

-endDate <arg>        Get stats of job instances with endDate in
yyyy_MM_dd
                      format.

-h                    Shows the help text.

-help                 Shows the help text.

-jobType <arg>        Type of JobType.Valid jobtypes are

{BACKUP,RESTORE,INCR_RESTORE,BLOCK_RESTORE,ARCHIVE,GA
                      RCHIVE,DIRECT_REPLICATION}

-outDir <arg>         Directory under which CSV files will be written.

-startDate <arg>      Get stats of job instances with startDate in
                      yyyy_MM_dd format.
```

NOTE: If startDate and endDate are not specified, then by default stats from last 30 days along with today will be displayed. For example, if today is April 24 2018 then stats from March 24 2018 to April 24 2018 will be displayed.

6.11 Uninstalling Imanis Data software

You can clean up cluster nodes by erasing all of the Imanis Data packages and configuration files from all Imanis Data nodes by running the installer with -c option. Use this option when you have to uninstall Imanis Data software. However, for a fresh installation, you do not need to do these steps as the Imanis Data Installer will clean up for the earlier packages and configuration files.

To uninstall Imanis Data software, do the following:
1. Log on as an admin user to the Install Master.
2. Run the following option:

```
./install_talena.sh -c
```

3. Imanis Data software will prompt you to provide a list of servers to uninstall. Follow the system prompts to complete the un-installation process.

> **IMPORTANT**: Cleanup removes the Imanis Data packages (RPMs starting with Tx) and INSTALL_DIR, `/opt/talena directories`. Ensure that you do not have any important data in these directories. For example, the log files from previous runs under `/opt/talena/talena-logs` directory.

6.12 Stop/Start Services

After you install Imanis Data software, all services are automatically initiated and smoke tests are executed. However, with the following script you can stop/start various services and run smoke tests as and when needed. Use stop/start services in case you want to shut down the Imanis Data cluster.
`${INSTALL_DIR}/bin/talena-services.sh`

> **IMPORTANT**: If one or more nodes get rebooted, log in to any one of the nodes and run this command manually to start services: `${INSTALL_DIR}/bin/talena-services.sh start all`

How to use this script?
You can use this script in two different ways:

```
${INSTALL_DIR}/bin/talena-services.sh {action} {services}
```

OR

```
${INSTALL_DIR}/bin/talena-services.sh {smoke} {tests}
```

You must be logged in as a service user to run this script.

Where:
- action should be one of {configure|start|stop|restart|status|smoke|get-active|-v}
- services should be one or more of {zk|hdfs|yarn|hbase|tcollector|postgres|hive|oozie|scheduler|UI|all}

- smoke tests should be one or more of {zk|hdfs|yarn|hbase|postgres|hive|dedupe|ec|backupRestore|oozie|scheduler| monitor|UI|callhome|namenodeHA|hbaseHA|all}

- get-active option should be one or more of {namenode|hmaster}

If there are more than one services or tests, then separate the services or tests by a comma.

NOTE:

yarn will start/stop ResourceManager, NodeManagers, JobHistoryServer

hbase will start/stop HMaster, HRegionServers

hdfs will start/stop NameNode, DataNodes, JournalNodes

WARNING: The action **Setup** must be used with caution as the data on the Imanis Data cluster may get erased permanently. Use the action Setup if you want to reformat the Imanis Data cluster without going through the installation process again. Also, if you use the action **Stop** on any node in the cluster, Imanis Data software will cease to work.

7   Appendix A: Alternate Modes of Installation

This section describes the alternate modes of installation. For more information, refer to the section Various Modes of Installation.

Steps for NOSSH-SUDO Mode of Installation

1. Copy the installer to each node (under the same path) where you want to extract and install.
2. If admin user is a non-root user and has sudo access, then NOPASSWD should be set for admin user in /etc/sudoers file on all nodes. For example, if Imanis Data is your admin user then set:

```
talena ALL=(ALL) NOPASSWD: ALL
```

3. Disable Firewall manually on all Imanis Data nodes. For more information, refer to the section Pre-requisites for Imanis Data cluster.
4. Change directory (cd) to a location where you have copied the installer on each node.
5. Extract the installer on each node:

```
# ./installer_self_extract_4.3.0-RELXXX.sh
```

6. Where XXX is the release number of the build. For example, REL136.
   The following will be displayed on the screen:

```
  # ./installer_self_extract_4.3.0-RELXXX.sh

Talena installation is supported in two modes:
SSH: Root user can SSH to nodes and execute commands.
NO-SSH: SSH access is not allowed for root.

Nodes will communicate through Remote Procedure Calls.

Do you want to continue with SSH mode installation? (yes/no) No
```

   Where XXX is the release number of the build. For example, REL136.

7. Type **NO** as shown in the above screenshot to continue.
8. Go to any one node and proceed with Running Pre-checks and continue with the installation and verification of the installation thereafter.

Steps for NOSSH-NOSUDO Mode of Installation

1. In this mode of installation, your admin user and service user both have to be hdfs. Login as hdfs user and follow the steps below:
2. Copy the installer to each node under the same path. For example, /home/hdfs

3. Disable Firewall manually on all Imanis Data nodes. For more information, refer to the section Pre-requisites for Imanis Data cluster.
4. Change directory (cd) to /home/hdfs on each node
5. Extract the installer on each node:

```
# ./installer_self_extract_4.3.0-RELXXX.sh
```

Where XXX is the release number of the build. For example, REL136.

The following will be displayed on the screen:

```
# ./installer_self_extract_4.3.0-RELXXX.sh

Talena installation is supported in two modes:
SSH: Root user can SSH to nodes and execute commands.
NO-SSH: SSH access is not allowed for root.

Nodes will communicate through Remote Procedure Calls.

Do you want to continue with SSH mode installation? (yes/no)NO
```

Where XXX is the release number of the build. For example, REL136.

6. Type **NO** for as shown in the above screenshot to continue.
7. Go to any one node and proceed with Running Pre-checks and continue with the installation.
8. Further, follow the instructions in Appendix B for post-installation steps.

8   Appendix B: Post Install Steps for NOSSH-NOSUDO mode of installation

This section describes the post install steps that you must do for the NOSSH - NOSUDO mode of installation.

8.1   Running Post Install script

If you are running the installer in NOSSH mode and without sudo privileges, then run the following script as a root on each node separately. For the convenience of the user, a script is provided under:

```
# $INSTALL_DIR/bin/talena-postinstall.sh
```

8.2   Setting Up and Starting Imanis Data services

Once the post install steps are executed on all nodes, login as hdfs user on any one of the Imanis Data cluster nodes and run execute the following commands to setup and start Imanis Data services

```
# $INSTALL_DIR/bin/talena-services.sh -f setup all
```

To ensure that the installation is done properly, you can run set of smoke tests on the Imanis Data cluster.

```
# $INSTALL_DIR/bin/talena-services.sh smoke all
```

Follow the steps in the section Verifying the Installation to verify if Imanis Data software is successfully installed.

9 Appendix C: Parameters for Imanis Data Installation

The parameters for Imanis Data software installation are divided into the following sections: Mandatory Parameters and Optional Parameters.

Mandatory Parameters

You must provide the following parameters prior to starting the software installation process.

PARAMETER	DESCRIPTION
`tl_physical_nodes=`	▪ List of physical nodes. ▪ Provide a comma separated list of hostnames. Hostnames should be DNS resolvable names. For example: tl_physical_nodes=node1.enterprise.com,node2.enterprise.com,node3.enterprise.com
`tl_datanode_dirlist=`	▪ Directory where datanodes keep the actual HDFS data. Ensure that paths provided for tl_datanode_dirlist should have read/write permissions for services user.
`tl_namenode_dir=`	▪ Directory where namenode keeps its metadata. Please make sure that, path provided for tl_namenode_dir should have read/write permissions for services user. Directories will be created as: <<provided dir path>>/talena-namenode. For example: ▪ tl_namenode_dir=/mnt/disk1/physical/nn,/mnt/disk2/physical/nn
`tl_journalnode_localdir=`	▪ This is absolute path on the JournalNode machines where edits and other local state is maintained by journal nodes. Please make sure that, path provided for tl_journalnode_localdir should have read/write permissions for services user. Directories will be created as: <<provided dir path>>/talena-journalnode. For example: ▪ tl_journalnode_localdir=/mnt/disk1/physical/jn,/mnt/disk2/physical/jn
`tl_cluster_name=`	▪ This parameter will be used to set cluster name which represents a unique identifier for the installed cluster.

PARAMETER	DESCRIPTION
`tl_licence_key_path=`	▪ This parameter will be used to set the absolute Path to the licence key file.
`tl_tmp_dir=`	▪ This parameter will be used to set temporary directory

Optional Parameters

At times you may want to use the optional commands to use the Imanis Data cluster in a better way.

PARAMETER	DESCRIPTION
`tl_private_key=`	▪ Private key file: Path to the private key file to be used as authentication for admin user. All cluster hosts must accept the same private key.
`tl_slave_nodes=`	▪ Manual placement: The installer determines automatic placement for these services.
`tl_exclusive_services=`	▪ Services which require dedicated machine to run possible values namenode1, hbasemaster, resourcemanager, backuphbasemaster, namenode2, oozie, ui. ▪ Note that these should be single instance per cluster kind of service only, that is, services like zookeeper, journal node, datanode, nodemanager, regionserver are invalid choices for this setting.
`tl_external_kdc_server=`	▪ Use this optional command when you use the external KDC server. ▪ Set tl_external_kdc_server=true ▪ When using an existing Kerberos realm for Imanis Data, provide the name of the server where KDC is already running.

PARAMETER	DESCRIPTION
`tl_callhome_cron_expr=0 2 * * *`	▪ Run a cron job at 2 am every day to upload the logs and other meta data required for supporting issues and troubleshooting. Specify time in the form of cron expression; logs will be uploaded at specified time. ▪ For example: "0 20 * * *" where logs will be uploaded everyday at 8pm. Here * means: MIN HOUR DOM MON DOW where MIN Minute field 0 to 59 HOUR Hour field 0 to 23 DOM Day of Month 1-31 MON Month field 1-12 DOW Day Of Week 0-6
`tl_external_nagios_server_ipaddr=`	▪ If you have a Nagios server already configured in your data center and you want to use it for monitoring Imanis Data platform, then provide the IP address of the same. Imanis Data Nagios clients will be configured to accept plugin run requests (through check_nrpe) from this IP address only through NRPE daemon.
`tl_dropbox_authkey_file=`	▪ Absolute path of the file containing appkey and secret-key. This key will be used to upload logs in Dropbox account created exclusively for your company. You need to get this file from Imanis Data support team before beginning the installation.
`tl_enable_https_webui=`	▪ Enable HTTPS support by setting the parameter to True.

PARAMETER	DESCRIPTION
`tl_https_certificate_path=`	▪ Absolute Path to the ssl certificate
`tl_https_key_path=`	▪ Absolute Path to the private key
`tl_installer_advanced.conf`	▪ Configuration file for advanced settings

Following are the supported file formats for tl_https_certificate_path:

```
[root@s4vm9-centos7 certs]# file server.crt server.key
server.crt: PEM certificate
server.key: PEM RSA private key
```

10 Appendix D: Installation Pre-checks

This appendix describes the pre-checks performed by the pre-check option. This step must be performed successfully before installing Imanis Data software.

You can run the installer with –pre-check option to verify if the cluster nodes meet all of the hardware and software requirements for installing and running Imanis Data software. Pre-checks are a set of tests which ONLY verifies if the system meets all prerequisites and does NOT install Imanis Data software.

Currently, the following checks are done:

PRECHECK NAME	DESCRIPTION
Minimum nodes check	- Checks minimum required nodes availability for the Imanis Data cluster installation. - Minimum of 3 nodes are required.
Network ports range check	- Checks availability of all TCP ports required for the Imanis Data cluster installation. 'netstat -a' command should show required ports as free.
Admin user sudo access check	- Checks admin user details. - The admin user's entry should present in '/etc/sudoers' file on all cluster nodes '.bashrc' file should exist in the admin user's home directory 'id' command should display same effective group name for admin user on all cluster nodes
Service user existence check	- Checks service user details. '.bashrc' file should exist in the service user's home directory. 'id' command should display same effective group name for service user on all cluster nodes.
Bash shell check	Checks Bash shell is set as user's shell

PRECHECK NAME	DESCRIPTION
Hostname validation	Checks if hostnames provided in tl_installer.conf matches with the output of "hostname and hostname --fqdn" command
Private key check for password protection	Checks private key is not passphrase protected
Admin can act as service user check	Checks admin can act as service user. Admin user should be able to act as service user
Free disk space check for installer	- Checks free disk space availability for packages under ${INSTALL_DIR}. 'df –kh {0}' command should display minimum free space 5 GB and above.
UID and GID check for RPyC daemon	Checks if uid and gid of RPyC daemon matches with uid and gid of admin user
Install directory symbolic link check	Verify if install directory is NOT a symbolic link
Shared directories across nodes check	Checks if any installation directories are not shared across Imanis Data cluster nodes.
Mounts exec check	Checks if required mounts are mounted with exec option
Network reachability check	Checks cluster nodes network reachability from Install Master and reachability among all cluster nodes.
Operating System Version check	- OS Type check on all cluster nodes. - The expected OS type is CentOS 6.2, Red Hat 6.2 or above.
Internet connectivity check	Checks Internet connectivity
Packages existence check	Checks version mismatch for prerequisite packages.

PRECHECK NAME	DESCRIPTION
Dropbox uploader tool check	▪ Checks if dropbox_uploader.sh is available
Postgres version check	▪ Verify the supported version of postgres is installed
Path check	▪ Checks system commands required for the Imanis Data installer. ▪ Make sure that all the above system utilities should be present on all cluster nodes.
Crontab access check for service user	▪ Checks if admin user has crontab access on all cluster nodes
Firewall status check	▪ Checks if IPtables are disabled. ▪ IPtables should be disabled on all cluster nodes.
Ulimit check	▪ Checks if ulimits are set to expected values for service user
DNS lookup check	▪ Perform DNS lookup for all Imanis Data cluster nodes
Loopback address check	▪ Perform loopback address check on all Imanis Data cluster nodes.
IPV4 check	▪ Checks if given host addresses are in IPV4 format.
Free disk space check for temp	▪ Checks free disk space in /tmp. ▪ 'df -kh /tmp' command should display minimum free space 4 GB and above.
Free disk space check for logs	▪ Checks free disk space in log directory: /opt/talena/talena-logs ▪ 'df -kh /opt/talena/talena-logs' command should display minimum free space 50 GB and above.

PRECHECK NAME	DESCRIPTION
Free disk space for yarn localdir	▪ Free disk space check for Imanis Data local directory
Network ports availability check	▪ Checks availability of all TCP ports required for Imanis Data cluster installation
Network ports reservation check	▪ Checks reservation of all TCP ports required for Imanis Data cluster installation
SeLinux status check	▪ Checks if SeLinux is disabled. Command 'sestatus' should show status as 'disabled' ▪ Refer to Appendix L for disabling SeLinux.
Directory availability check	▪ Checks if following services have atleast one data directory exists on given nodes
Datanode directory access check for service user	▪ Checks if datanodes have at least one data directory specified on each node
Timezone check	▪ Checks if all cluster nodes have same timezone.
Clock synchronization check	▪ Checks if clocks are synchronized across cluster nodes.
Mailx check	▪ Send mail using mailx command and ask user to verify if mail has received
Swap space check	▪ Check swap space requirement for Imanis Data cluster installation
Service placement (Role assignment) check	▪ Checks if Imanis Data services can be placed (role assignment) on the given cluster nodes.
Memory check for Yarn containers	▪ Checks if Imanis Data nodes have enough memory to create yarn containers

PRECHECK NAME	DESCRIPTION
Free disk space check for namenode	▪ Checks free disk space in namenode directory
Free disk space for callhome	▪ Checks free disk space in callhome direcotry

Pre-checks generate a consolidated report for various tests as PASS or FAIL. Depending on the kind of failure, you need to take specific actions as mentioned in the report. In case of failures, for example, the following is displayed:

```
Free disk space for log directory check (/opt/talena/talena-logs): FAIL
(s2vm4,s2vm3,s2vm2)
Description: Check free disk space in log directory /opt/talena/talena-logs
Expected: 'df -kh /opt/talena/talena-logs' command should display minimum
free space 50 GB and above
Action: Make some free space in /opt/talena/talena-logs by deleting unwanted
data
```

In the above example, the "Free disk space for log directory" check has failed. The system displays the description of the check, what is expected, and if there is any action required from you to fix it. You must then fix the failure and rerun the pre-check.

However, some failures do NOT demand any action from you. For example:

```
Talena RPMs check: FAIL(s2vm4,s2vm3,s2vm2)
Description: Check if Talena RPM's exists on a system.
Expected: Talena RPMs should not be present on any cluster node
Action: NO action needed from user. Installer will cleanup stale RPMs if
found, during actual installation
```

11 Appendix E: Installation Checklist

Prior to starting the installation process, it is recommended that the administrator completes the information column. This information can be used as a quick reference list to speed up the installation process.

Imanis Data Cluster

DETAILS	INFORMATION
IP address or host name of the Imanis Data nodes	
List of Data Directories on the Imanis Data cluster	
Imanis Data License File	

Hadoop Cluster

DETAILS	INFORMATION
IP address and port number of HIVE metastore on primary cluster	
IP address of Namenode on primary cluster	
Login credentials for primary Hadoop cluster	
Location of HDFS and HIVE configuration directories	

12 Appendix F: Command Line Options List

Use the following command line options to learn advanced ways of managing the Imanis Data cluster.

TERM	DEFINITION	NOTES
-h	--help	Print this help.
-c	--cleanup	Cleanup the Imanis Data cluster.
-u	--upgrade	Upgrades Imanis Data RPMs and config files.
-a	--addnode newnode1, newnode2, newnode3	Adds slave nodes.
-p	--prechecks	Performs pre-checks for the Imanis Data Installer without actual installation (dry run).
-n	--shun	Skips smoke tests, resource checking while placement and autosizing of heap memory, primary namenode checks.
-d	--decommission node1,node2	Performs decommissioning of nodes. Supported only for slave nodes.
-m	--migrate service1,service2,[all] --source node1,node2 [--target node3,node4] [-l/--live]	Migrates the specified services from source to target node(s). Valid service names: zookeeper, namenode, hbasemaster, oozie, ui, resourcemanager, journalnode, talena-nagios-server

13 Appendix G: Nagios Plugins

This appendix lists the Nagios plugins that are supported by Imanis Data software.

TERM	DEFINITION	NOTES
check_ping	Uses the ping command to probe the specified host for Packet Loss (percentage) and Round Trip Average (milliseconds)	If RTA is 100 ms or if Packet Loss is 20%, then a 'WARNING' alert is generated
		If RTA is 500 ms or Packet Loss is 60%, then a 'CRITICAL' alert is generated
check_SSH	Tries to connect to an SSH server at a specified host.	If the node is not reachable through SSH, then a 'CRITICAL' alert is generated
check_load	Checks the current system load average for the past 1,5 and 15 minutes	If the load average for the last 1, 5, 15 minutes is above 15,10, and 5 respectively, then a 'WARNING' alert is generated
		If the load average for the last 1, 5, 15 minutes is above 30, 25, and 20 respectively, then a 'CRITICAL' alert is generated
check_memory	Checks the amount memory consumption of the system	WARNING: If the memory consumption goes above 80%, then a 'WARNING' alert is generated
		If the memory consumption goes above 90%, then a 'CRITICAL' alert is generated
check_cpu	Checks the CPU consumption of the system	If the CPU consumption goes above 80%, then a 'WARNING' is generated

TERM	DEFINITION	NOTES
		If the CPU consumption goes above 90%, then a 'CRITICAL' alert is generated
check_cpu	Checks the CPU consumption of the system	If the CPU consumption goes above 80%, then a 'WARNING' is generated
		If the CPU consumption goes above 90%, then a 'CRITICAL' alert is generated
check_talena_jobs	Checks if there are any failed jobs on the Imanis Data cluster	If a job has failed once since 00:00 hours on any given day, then a 'WARNING' alert is generated
		If a job has failed (more than once) since 00:00 hours on any given day, then 'CRITICAL' alert is generated
check_talena_disk	Checks the usage of each mounted file system on a given node. NOTE: Only checks the ext3 | ext4 file systems.	If the free space left is less than 10% of the capacity, then a 'WARNING' alert is generated
		If the free space left is less than 5% of the capacity, then 'CRITICAL' alert is generated
check_talena_services	Checks the status of Imanis Data services on the cluster.	If any of the service is not running, then a 'CRITICAL' alert is generated

TERM	DEFINITION	NOTES
SMART checks	Checks the health of data disks used by Imanis Data and reports if any disk is likely to fail.	If the disk(s) are about to fail, 'CRITICAL' alert is generated. Immediate action is required. If the health of the disk(s) could be determined, 'WARNING' alert is generated.
Imanis Data Callhome	Checks if Callhome feature of Imanis Data is working fine or not.	If Imanis Data callhome is not running, 'CRITICAL' alert is generated. If the last run of Imanis Data explorer data upload fails or if waiting for Imanis Data callhome to respond times out, 'WARNING' alert is generated.

14 Appendix H: Quick Troubleshooting

This appendix includes troubleshooting information that you can use should a problem occur while installing Imanis Data software. The troubleshooting section is divided into the following sections:

1. Issues that result in the Imanis Data installer exiting automatically: In this case, the Imanis Data installer displays the error message on the console and exits automatically. It is recommended that you fix the error and restart the installation process.

2. Issues in which the Imanis Data installer prompts user for corrective action and/or removes only the affected node and continues with the installation: In this case, the Imanis Data installer prompts you to continue or abort the installation process.

3. Issues that are observed with Imanis Data cluster after the installation is complete

4. Issues that are observed during upgrading Imanis Data cluster

5. Issues that are observed during adding and/or upgrading Imanis Data license

6. Region servers dying due to out of memory error

1. Issues that result in the Imanis Data installer exiting automatically

ERROR MESSAGE	POSSIBLE PROBLEM	POSSIBLE SOLUTION
Failed to extract installer on <node>	Insufficient disk space directory	Make at least 4 GB disk space available on target nodes
	Insufficient permission to target	Check and rectify directory ${INSTALL_DIR} permissions
	Corrupted file - installer.tar.gz	
Failed to extract python on <node>	Insufficient disk space	Make at least 4 GB disk space available on target nodes
	Insufficient permission to target directory	Check and rectify directory ${INSTALL_DIR} permissions
	Corrupted file - python.tar.gz	

ERROR MESSAGE	POSSIBLE PROBLEM	POSSIBLE SOLUTION
Exception while "Checking nodes reachability from Install Master"	Network connectivity	• Check for packet loss • Check for latency

2. Issues in which the Imanis Data installer prompts the user for a corrective action, and/or removes only the affected node, and continues with the installation process.

ERROR MESSAGE	POSSIBLE PROBLEM	POSSIBLE SOLUTION
Error while stopping IPTables on <node>	Failed to unload module	Try stopping the iptables by running '/sbin/service iptables stop' as root and check for the error message
OS type check failure	OS type other than Centos and Red Hat will cause OS type check failure	OnlyRHEL6.x and CentOS 6.x are supported
RPMs <name> already installed on <node>	Installation fails when the Imanis Data RPMs are already installed on the cluster node	Remove the Imanis Data RPMs by executing "rpm -e <rpm-name>" or run cleanup
Failed to disable se-linux on <node> OR Error determining SELinux mode	SeLinux may be actively used in the system	Check se-linux mode In '/etc/sysconfig/selinux' and if it is 'enabled' then disable the same using command "setenforce 0" Refer to Appendix L for disbaling SeLinux

ERROR MESSAGE	POSSIBLE PROBLEM	POSSIBLE SOLUTION
Failed to install <package> on <node>	The package may already be installed	Run rpm -qa | grep<package>to see if the package is already installed. If it is installed, remove the package
Time difference on all nodes is greater than …	Inconsistent timezone on cluster nodes.	Install NTP to synchronize clock times between cluster nodes. Ensure that the difference between clock times on each of the Imanis Data nodes is not more than 30 seconds.
Nodes <node list> not reachable from <node>	Network problem	Check if you can manually ping each of the nodes listed in conf/tl_installer.conf file from the install master and from each node itself
Failed to copy <file> to <node>	Insufficient disk space	Check if there is enough disk space using df command
	Insufficient permission to target directory	Check the ownership of the folder under which file creation failed
Error while connecting to server <nodename> using RPyC	rpyc client is not running on node	Check if rpyc_classic.py is running in ps output.
	rpyc port '18812' already in use	Check if the port is already in use using lsof command
	IPTables are enabled	Check if there is firewall configured which is blocking certain ports

ERROR MESSAGE	POSSIBLE PROBLEM	POSSIBLE SOLUTION
Failed to install RPMs <RPM list> on <node>	Imanis Data RPMs already exist	Run rpm -qa	grep<RPMname> to check if the rpm exists. If it exists, it needs to be removed using command rpm -e <RPMName>
	/opt has no space left	Run command df -kh /opt to check free space available on /opt. Increase the free disk space to at least 4 GB
	/tmp has no free space	Run command df -kh /tmpto check free space available on /tmp. Increase the free disk space to at least 4 GB
Failed to install JDK on <node>	/opt has no space left	Run command df -kh /opt to check free space available on /opt. Increase the free disk space to at least 4 GB

3. Issues that are observed with the Imanis Data cluster after the installation is complete

Issue: If more than one node in the Imanis Data cluster is restarted at the same time, then some services fail to come up.

Solution: This issue is caused as some services like hbase, for example, are dependent on HDFS and Zookeeper to function properly. Depending on the node where the services are hosted, if more than one node is restarted around the same time, it is possible that the dependent services like zookeeper and/or hdfs are also coming up and are not functional at a point when hbase is coming up. As a result, hbase will not be started as the dependencies are not met.

In such cases, a manual start services may be needed:
- Ensure that all nodes are booted completely and then login to any one node as service user (usually hdfs) and run the following command:

```
${INSTALL_DIR}/bin/talena-services.sh start all
```

This command will only start those services which are not up and running and will not affect the services which are already running.

Issue: If the Nagios server is down

Solution: First you must verify the Nagios status by using the following command. The command displays running the Nagios NRPE daemon and the IP address where the Nagios server is running.

```
${INSTALL_DIR}/bin/talena-services.sh status nagios
```

If the Nagios server is NOT running, then you can NOT start/stop Nagios using the following script:

```
${INSTALL_DIR}/bin/talena-services.sh start|stop nagios
```

You need to use the following script as root user:

```
${INSTALL_DIR}/bin/lib/talena-nagios.sh {start|stop} nagios
```

Issue: If the host running the Nagios lxc container is restarted

Solution: /etc/init.d/talenad attempts to restart the Nagios server if you reboot or restart the host machine. You do NOT have to do anything explicitly to start the Nagios server.

If there is any problem starting the Nagios server, you can view the log file to determine the issue:

```
cat /opt/talena/talena-logs/nagios-container.log
```

4. Issues that are observed during upgrading Imanis Data cluster
In case of SUDO-NO-SSH upgradation, the install directory is owned by hdfs user which is set during the course of installation.

However, if a non-hdfs user (with sudo) tries to extract the Imanis Data installer, the extraction fails as permission is denied. Thus to be able to upgrade in SUDO-NO-SSH mode, refer to the following procedure.

To upgrade Imanis Data cluster with SUDO-NO-SSH, do the following:
1. The installer must be extracted in the path from where the installation is performed. This particular path can be found by running the following command as a service user:

```
echo ${INSTALL_DIR}
```

If your install directory is /home/installation/talena, then new installer has to be extracted at /home/installation.

Extract the Imanis Data installer in the path from where the installation is performed using the following command:

```
# sudo ./installer_self_extract_x.y.z.sh –n
```

Where x.y.z is the complete Imanis Data release number.

Change the ownership of installer_self_extract_x.y.z.sh to the admin user using the following command:

```
# sudo chown –R <ADMIN_USER>: talena/x.y.z
```

Where x.y.z is the complete Imanis Data release number.

5. Issues that are observed during adding and/or upgrading Imanis Data license

This section includes troubleshooting information that you can use should a problem occur while adding and/or upgrading Imanis Data licence.

Issue: Failed to setup licence.

Solution: Manually restoring to the last valid licence file.

Steps to manually restore to the last valid licence file:

1. Switch to the service user, for example, hdfs.
2. Change directory to $INSTALL_DIR/conf.
3. Search for files .imanis-licence and .imanis-licence.json with timestamp suffix. For example,

   ls -al .imanis-licence*
   .imanis-licence_2018-09-04-12-55-17
   .imanis-licence.json_2018-09-04-12-55-17

4. Select last working copy of .imanis-licence file and restore it.
5. Select last working copy of .imanis-licence.json file and restore it.
6. Run : talena-scp.sh $INSTALL_DIR/conf/.imanis-licence $INSTALL_DIR/conf/.imanis-licence
7. Run : talena-scp.sh $INSTALL_DIR/conf/.imanis-licence.json $INSTALL_DIR/conf/.imanis-licence.json

NOTE: To find out the last working copy of the files, compare the checksum of the licence file used during Imanis Data installation and/or upgrade. The md5sum of the used file should match with .imanis-licence_<timestamp>.

6. Region servers dying due to out of memory error

In rare cases even after increasing the region server memory, the region server may continue to die due to out of memory error. In this case, it is recommended to increase the Hbase heapsize of region server.

To increase the heap size of region server, do the following:

1. Edit the following:

```
$HBASE_HOME/conf/hbase-env.sh
```

2. Set the Hbase Heapsize to 16GB:

```
export HBASE_HEAPSIZE="16384"
```

7. Replacing a failed datanode disk

In unlikely scenarios when datanote disks fail, use the following steps to replace the failed disks with new disks:

To replace a failed datanode disk, do the following:

1. Stop services
2. Unmount the file system on the failed disk.
3. Replace the disk drive.
4. Format and mount the disk drive at the same mount point where it was mounted earlier.
5. As $SERVICE_USER, run the following command:

```
# mkdir -p   <mount_point>/talena-datanode/
```

```
# chown -R  $SERVICE_USER:   <mount_point>/talena-datanode/
```

6. Start services.
7. After HDFS is up and running, verify that `<mount_point>/talena-datanode/` is being used. Generally it must have a directory called `'current'` and a file called `'in_use.lock'`

15 Appendix I: Parameters for Non-HA and HA Cluster

Use the following information to set the parameters if your primary cluster is Hadoop cluster (Non-High Availability).

PARAMETER	EXPECTED FORMAT	DESCRIPTION
fs.defaultFS/fs.default.name	hdfs://string or hdfs://string:number	Mandatory
dfs.webhdfs.enabled	Boolean value Must be set to true in hdfs-site.xml	Mandatory If this property is missing, HDFS will NOT be detected. Please verify configuration for this property. If missing, this property must be added.
dfs.namenode.http.address	If present, it is expected in format host:port. If not, port number is assumed to be 50070 (default value).	Mandatory

Use the following information to set the parameters for High Availability cluster.

In addition to the parameters listed above, the following parameters are all also expected. A setup is discovered as HA setup if dfs.nameservices value matches with the name part of fs.defaultFS parameter. Please consult your Hadoop cluster administrators and refer to the Hadoop distribution documentation to ensure that the following parameters are present in the configuration file.

PARAMETER	EXPECTED FORMAT	DESCRIPTION
dfs.nameservices	string	Mandatory
dfs.client.failover.proxy.provider.nameservice	name of a class	Mandatory
dfs.ha.namenodes.nameservice	comma separated strings	Mandatory

PARAMETER	EXPECTED FORMAT	DESCRIPTION
dfs.namenode.rpc.address.nameservice.namenode[12]	host:port	Mandatory
dfs.namenode.servicerpc-address.nameservice.namenode[12]	host:port	Mandatory

16 Appendix J: Formatting & Mounting New Hard Drives

In case you have multiple disks that you want to use for Imanis Data data repository then you must follow the steps below.

1. First list out the disks with the following command:

```
# fdisk –l
```

An example of the output is presented below:

Disk /dev/sdb: 2147 MB, 2147483648 bytes, 4194304 sectors

Units = sectors of 1 * 512 = 512 bytes

Sector size (logical/physical): 512 bytes / 512 bytes

I/O size (minimum/optimal): 512 bytes / 512 bytes

2. Now type the following to find out what disk devices were actually mounted:

```
# df –h | grep ^/dev
```

Results of the command below:

/dev/mapper/centos-root 8.5G 2.9G 5.7G 34% /

/dev/sda1 497M 164M 334M 33% /boot

From the results and by comparing the two outputs you can see that /dev/sdb was not mounted. To add this disk to the system, continue with the following procedures. You can repeat these steps if you have multiple disks that were not setup.

3. Create a File System on the new disk:

```
# mkfs.ext4 /dev/sdb
```

Follow the steps above for all the disks you found previously that were not in use.

4. Now add the disks to the /etc/fstab to allow them to mount automatically. First get the ID of the newly formatted disk.

```
# blkid –o full –s UUID
```

You will see results similar to the following:

/dev/sdb: UUID="2ded38f5-b060- 4e98-827a- 2c55910ec053"

5. Now add the disks to /etc/fstab:

```
# vi /etc/fstab
```

6. Add the line highlighted below using the UUID information you previously received above.
   # Disk 1 directory on /dev/sdb
   UUID=2ded38f5-b060- 4e98-827a- 2c55910ec053 /mnt/disk1 ext4 errors=remount-ro 0 1

7. Create a mount point directory:

```
vi # mkdir /mnt/disk1
```

8. Now mount up the newly formatted drive(s) that were added to the /etc/fstab:

```
# vi # mount –a
```

9. Make sure the new drives were mounted:

```
# vi # df –h
```

17 Appendix K: Creating & Adding a Swap File

This appendix section describes the procedure of creating and adding a new swap file.

To create and add a swap file, do the following:

1. Determine the size of the new swap file in megabytes and multiply it by 1024 to determine the number of blocks. For example, the block size of a 64 MB swap file is 65536.
2. At a shell prompt as root, type the following command with count being equal to the desired block size:

```
dd if=/dev/zero of=/swapfile bs=1024 count=65536
```

3. Set up the swap file with the following command:

```
mkswap /swapfile
```

4. To enable the swap file immediately but not automatically at boot time:

```
swapon /swapfile
```

5. To enable it at boot time, edit /etc/fstab to include the following entry:

```
/swapfile swap swap defaults 0 0
```

(The next time the system boots, it enables the new swap file.)

6. After adding the new swap file and enabling it, verify if it is enabled by viewing the output of the following command:

```
cat /proc/swaps or free.
```

18 Appendix L: Disabling SeLinux

This appendix section describes the procedure of disabling Selinux in CentOS, RedHat, and Ubuntu.

18.1 Disabling SeLinux in CentOS

Security-Enhanced Linux (SELinux) allows you to set access control through policies. You must disable SELinux on each host before you deploy CDH on your cluster.

To disable SELinux in CentOS, perform the following steps on each host:

1. Check the SELinux state.

```
getenforce
```

2. If the output is either permissive or disabled, you can skip this task and go on to disable the firewall. If the output is enforcing, continue to the next step.
3. Open the /etc/selinux/config file (in some systems, the /etc/sysconfig/selinux file).
4. Change the line SELINUX=enforcing to SELINUX=permissive.
5. Save and close the file.
6. Restart your system or run the following command to disable SELinux immediately:

```
setenforce 0
```

18.2 Disabling SELinux in Ubuntu

Security-Enhanced Linux (SELinux) allows you to set access control through policies. You must disable SELinux on each host before you deploy CDH on your cluster.

To disable SELinux in Ubuntu, perform the following steps on each host:

1. Stop the SELinux service:

```
sudo /etc/init.d/apparmor stop
```

2. Disable the SELinux service:

```
sudo update-rc.d -f apparmor remove
```

Your Feedback

Was this document helpful? [Send us your feedback](#)!

ABOUT COHESITY

Cohesity makes your data work for you by consolidating secondary storage silos onto a hyperconverged, web-scale data platform that spans both private and public clouds. Enterprise customers begin by radically streamlining their backup and data protection, then converge file and object services, test/dev instances, and analytic functions to provide a global data store. Cohesity counts many Global 1000 companies and federal agencies among its rapidly growing customer base and was named to Forbes' "Next Billion-Dollar Startups 2017," LinkedIn's "Startups: The 50 Industry Disruptors You Need to Know Now," and CRN's "2017 Emerging Vendors in Storage" lists.

For more information, visit our [website](#) and [blog](#), follow us on [Twitter](#) and [LinkedIn](#) and like us on [Facebook](#).