



IBM Analytics, Solutions



IBM Customer Insight for Communications Service Providers Deployment Guide

Fixpack version 1.0.4.1.0

CONTENTS

List of Figures.....	viii
1 Solution overview.....	9
1.1 Features.....	9
1.2 Users and benefits	9
1.3 Intended audience.....	9
1.4 Base architecture components	10
1.5 Framework	10
1.5.1 IBM Predictive Customer Intelligence (PCI)	10
1.5.2 IBM Open Platform	11
1.5.3 IBM Big Insights 4.1	11
1.5.4 IBM Analytics Accelerator Framework Platform	11
1.5.5 IBM Streams	11
Design and architecture	13
2 Planning for deployment	14
2.1 Roles and responsibilities.....	14
2.2 Hardware requirements.....	15
2.2.1 Partitioning Recommendations for Slave Nodes	16
2.2.2 Redundancy (RAID) recommendations	16
2.3 Software requirements	16
2.4 Media packaging	17
2.5 Deployment servers	17

2.6	Deployment process.....	19
2.7	Deployment dependencies	19
2.8	Deployment sequence.....	20
2.9	Post deployment verification.....	22
2.10	Deployment worksheets	22
2.11	Provisioning data.....	22
3	Deploying the solution.....	23
3.1	Before deploying products and components.....	23
3.1.1	Installing the IBM PCI optional components.....	23
3.1.2	Installing the IBM Big Insights - BigSQL value added service.....	24
3.2	Downloading the Customer Insight for CSP solution	26
3.3	Deploying the Customer Insight for CSP solution	26
3.3.1	Prerequisite steps	26
3.3.2	Modifying the permissions file.....	26
3.3.3	Running the Solution Installer to deploy the Customer Insight solution .	27
3.4	Configuring BigSQL access in a master/client Analytics Platform installation.....	29
3.4.1	Running install_telsol for a master / client installation	29
3.4.2	Deploying CEA dataset.....	29
3.4.3	Running a BigSQL query or update statement on the client	30
3.4.4	Connecting Cognos Reports to BigSQL	30
3.5	Deploying the Customer Insight datasets	31
3.5.1	Validating the dataset deployment.....	32
3.5.2	Verifying dataset scheduling.....	32
3.6	Provisioning data.....	33
3.6.1	Types of provisioning data.....	33

3.6.2	When to provision	34
3.6.3	Provisioning data before or after installing Telco Solutions.....	35
3.7	Creating the Telco Database	36
3.7.1	Verifying the Telco database installation	36
3.8	Deploying the Database Loader	37
3.8.1	Updating the Database Loader NPS and Churn configuration files	37
3.8.2	Updating execution permission on the Linux scripts	44
3.8.3	Base encoding the passwords in the configuration files.....	44
3.8.4	Manually running the Churn Database Loader job	45
3.8.5	Checking the Churn Database Loader job	45
3.8.6	Manually running the NPS Database Loader job	46
3.8.7	Checking the NPS Database Loader job.....	46
3.8.8	Configuring the Database Loader cron job.....	46
3.8.9	Verifying that the Database Loader cron job is set up correctly	47
3.9	Configuring SPSS components	47
3.9.1	Configuring the SPSS Modeler Server connection to the Analytic Server.....	48
3.9.2	Configuring SPSS Modeler Client connection to the Analytic Server.....	48
3.9.3	Configuring the SPSS Modeler Server connection to DB2	49
3.9.4	Configuring the SPSS Modeler Client connection to DB2.....	52
3.10	Deploying the SPSS Modeler	53
3.10.1	Before you begin.....	53
3.10.2	Deploying the Analytic Server datasources.....	53
3.10.3	Deploying models in SPSS Collaboration and Deployment Services	54
3.10.4	Scheduling SPSS Job Triggers	54
3.10.5	Validating the installation	57
3.11	Deploying the visualization component.....	58
3.11.1	Before you begin.....	58
3.11.2	Deploying the reports.....	60
3.11.3	Deploying the dashboards	62

	3.11.4	Verifying deployment of reports and dashboards	64
4		Configuring the Customer Insight for CSP solution	65
	4.1	Managing resources assigned to dataset processing	65
		4.1.1 Configuring the Yarn queues	65
		4.1.2 Updating solution datasets to run on a specific queue	65
		4.1.3 Queue management summary commands	65
		4.1.4 Queue file locations	66
		4.1.5 Verifying a query has run on the correct queue.....	66
5		Troubleshooting installation	67
	5.1	Problems and solutions during installation.....	67
		5.1.1 Port already in use when running setup.sh	67
		5.1.2 Deployment of the solution fails.....	67
		5.1.3 Telco database fails to create.....	67
		5.1.4 Cannot locate the Database Loader solution content	67
		5.1.5 Churn/NPS Database Loader fails to run due to no records	67
		5.1.6 Churn/NPS Models do not execute	68
		5.1.7 Database Loader Cron Jobs are not visible as boss	68
		5.1.8 Cron Jobs not running	68
		5.1.9 Cron Job removed or does not exist.....	68
		5.1.10 Collaboration and Deployment Services Windows client cannot connect to the Collaboration and Deployment Services repository.....	68
		5.1.11 No DB2 datasource is available to select in Modeler Client	70
		5.1.12 Analytic Server connection fails in Modeler Client.....	70
		5.1.13 The SPSS model fails to run.....	70
		5.1.14 Jobs are queued in Collaboration and Deployment Services.....	70
		5.1.15 Provisioning a CSV fails	70
		5.1.16 How to verify that all RPMs are installed	70
		5.1.17 How to verify cron jobs are set up correctly.....	71
		5.1.18 What to do if a dataset fails to run	72
6		Appendices.....	73

6.1	Sample Deployment Sequence Worksheet	73
6.2	Sample Data Required Worksheet	75
7	Glossary	79

LIST OF FIGURES

Figure 1 Architecture.....	13
Figure 2 Typical Customer Insight for CSP Deployment	18
Figure 3 Deployment dependencies	20
Figure 4 Customer Insight for CSP deployment sequence	21
Figure 5 Sample PCI optional component installation	23
Figure 6 Service Selection	25
Figure 7 SPSS Dataset and Model Activity Flow Sample	55

1 Solution overview

IBM® Customer Insight for Communications Service Providers (CSPs) is a prepackaged and self-contained software solution that integrates the functionality of many IBM software products. The solution is distributed as a collection of RPMs and contains the following core elements:

- Dashboards and reports to drive improved marketing and customer care performance
- Advanced and predictive analytics of customer activity across locations, devices, applications, and interests.

1.1 Features

IBM Customer Insight for CSP comprises of a set of analytics jobs (scripts, hive queries) that are deployed on top of the IBM Analytics Accelerator Framework (AAF) platform to generate datasets for reporting and visualization purposes.

Dataset generation is automated as part of the installation and runs at varying intervals over the underlying datasets. SPSS models that are included in the solution are run against the tables in Hadoop and produce small datasets. The analytics outputs are tables in Hadoop or DB2.

IBM Customer Insight for CSP contains a set of dashboards and Cognos reports. Dashboards and reports are run against the datasets in the Analytics Accelerator Framework platform or DB2. Reports are generated for Churn and Net Promoter Score (NPS).

1.2 Users and benefits

The solution provides benefits to a wide range of Telco users as described in the following table.

If you are a	IBM® Customer Insight can help you
Business analyst	Design dashboards and reports
Data scientist	Design and build models
Customer service representative	Understand customer behaviour
Solution architect	Design Telco applications
Marketing manager	Understand your customer segments, profiles, behaviour

1.3 Intended audience

This document is intended for people who are installing, administering, and maintaining the solution.

This document assumes that users have prior knowledge of or proficiency with the prerequisite software. Training for these base products is outside the scope of this document. If you require training for these products, ask your systems integrator or IBM representative where you can obtain information about base component training opportunities.

Note: Internal training can be provided and can be obtained by contacting the Lab services trainer for Telco Products.

1.4 Base architecture components

IBM Customer Insight for CSP provides and operates with the following base architecture components:

- IBM Predictive Customer Intelligence
 - Cognos Analytics
 - SPSS
 - DB2
- IBM Open Platform (IOP)
- IBM Big Insights 4.1
- IBM Analytics Accelerator Framework Platform
- IBM Streams

1.5 Framework

Figure 1 describes the IBM Customer Insight for CSP architecture, design and framework.

1.5.1 IBM Predictive Customer Intelligence (PCI)

The Customer Insight for CSP solution requires IBM® Predictive Customer Intelligence. PCI gives you the information and insight that you need to provide proactive service to your customers. PCI version 1.1.1 includes a number of component products:

- IBM WebSphere Application Server, installed on the predictive analytics node (pcipanode)
- SPSS Products, installed on the predictive analytics node (pcipanode)
- Cognos Analytics, installed on the Cognos analytics node (pcibinode)
- DB2 Enterprise Server, installed on the data node (pcidbnode)
- IBM Message Queue, installed on the Integration bus node (pciibnode)
- IBM Integration Bus
- Client interfaces

When you install PCI, you have access to all the above components, however not all the components listed above are used by IBM Customer Insight for CSP. The following PCI component products are used in the Customer Insights for CSP solution:

- SPSS Modeler Server for churn prediction and customer profiling models

- SPSS Collaboration and Deployment Services for deploying models
- Cognos Analytics for Churn, NPS, and visualization of reports and dashboards
- DB2 for Telco database and provisioning

Note PCI version 1.1.1 includes Cognos 11.0.2 which is required to support the IBM Customer Insight for CSP dashboards.

Important: Before you install the Customer Insight for CSP solution, ensure that PCI is installed and operational. In section 3.1.1, prerequisite tasks are required to prepare the PCI deployment for operation with the Customer Insight for CSP solution.

1.5.2 IBM Open Platform

IBM Open Platform V4.1 provides a set of open source tools used for data sets and analysis. IBM Customer Insight for CSP uses the following tools from IBM Open Platform:

- Ambari is an Apache Hadoop open source component and part of the IBM Open Platform. Ambari is a system for provisioning, managing, and monitoring Apache Hadoop clusters.
- Hadoop, Hive, Knox and Parquet for data set storage and encoding.
- Sqoop to transfer data, for analysis by SPSS jobs, from Hadoop to DB2.

1.5.3 IBM Big Insights 4.1

IBM Big Insights 4.1.0 is a collection of powerful value-add services that can be installed on top of the IBM Open Platform with Apache Hadoop.

The value-add services in Big Insights include: IBM Big SQL, IBM Big Sheets, IBM Big R, and IBM Text Analytics.

IBM Customer Insight for CSP uses Big SQL. Big SQL is an IBM DB2-style interface to Hadoop. Big SQL is used by the CEA reports to access data for the reports.

1.5.4 IBM Analytics Accelerator Framework Platform

The IBM Analytics Accelerator Framework (AAF) platform consists of the base data sets and services, upon which Customer Insight for CSP bases its analysis. IBM AAF Version 1.0.4 is used. The underlying platform version is Analytics Platform 3.1.0.1.

AAF is the foundation for the Customer Insight solution. The platform must be up and running prior to installing Customer Insight for CSP.

IBM AAF is layered on top of the Telecom Analytics supporting programs. Telecom Analytics is the 'first chargeable component' included in each IBM Now Factory solution. It is a bundle of supporting programs that must be licensed in order to deliver the IBM Analytics Accelerator Framework (AAF).

Telecom Analytics is available only in the context of providing a licensing vehicle for the programs required by IBM Now Factory products. It cannot be applied outside of IBM Now Factory Product suite.

1.5.5 IBM Streams

IBM® InfoSphere® Streams is a software platform that enables the development and execution of applications that process information in data streams. InfoSphere Streams

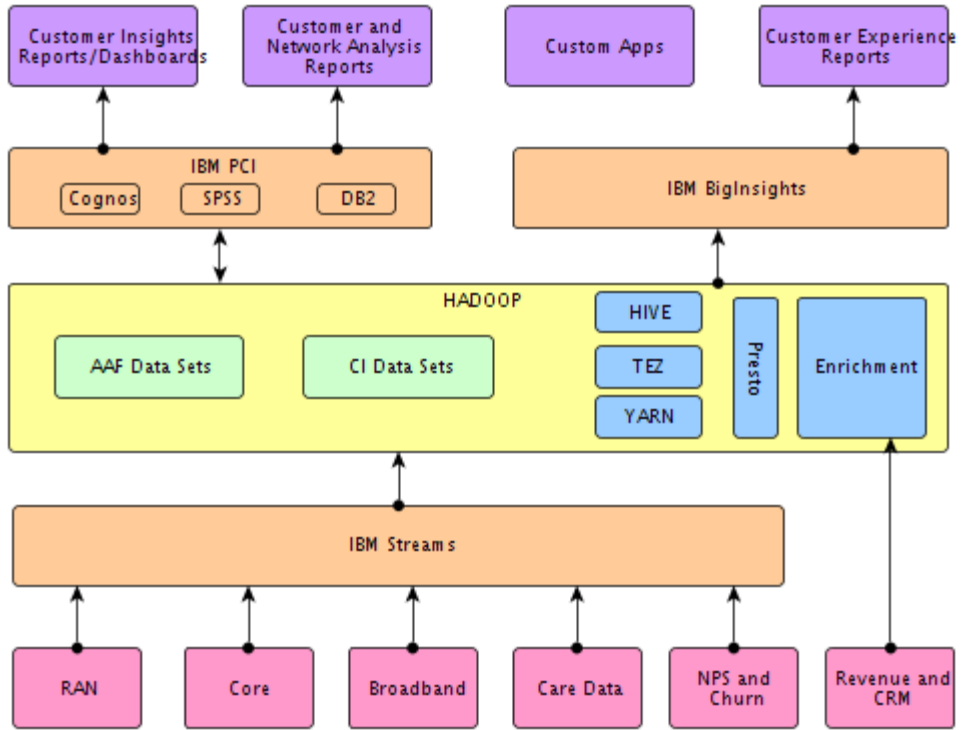
enables continuous and fast analysis of massive volumes of moving data to help improve the speed of business insight and decision making.

InfoSphere Streams consists of many components such as streams processing applications, domains, instances, and resources. IBM Analytics Accelerator Framework (AAF) platform runs on IBM Streams.

Design and architecture

Figure 1 describes the solution design and architecture.

Figure 1 Architecture



2 Planning for deployment

Planning for deployment is critical to the success of a Customer Insight for CSP implementation.

Successful planning includes analyzing how you are going to use the solution, obtaining the required hardware and software for deployment, and preparing the deployment infrastructure.

Use this information to understand and effectively deploy the solution in your environment to suit your business needs.

2.1 Roles and responsibilities

The solution implementation requires coordination of the deployment across multiple roles.

Depending on your implementation, solution architects, designers, analysts, developers, and IBM service team members can be key contributors in the deployment of your solution.

System administrators and their counterparts are responsible for deploying and maintaining the implementation. Installers and administrators are expected to have technical skill in the following areas:

- Using the Red Hat Linux operating system.
- Using open source database technology and tools (Hadoop, Hive).
- Working with virtual machines (VMs) and configuring VM connectivity.
- Providing system administration for the component products that comprise the solution.
- Background in the deployment and management of software in a Linux server environment.
- Background in the physical installation of server equipment at a data center.
- Data analytics platforms (structured/unstructured Data), Business Intelligence reporting tools (for example, Cognos) and Hadoop or other HDFS type systems.
- Using SQL to obtain data from standard database systems.
- Background in Hive or other query language for querying big data file systems.
- Able to create scripts to process data using standard scripting methods (Python, Perl, Bash, XML)
- Strong understanding of the fundamental operation of enterprise scale networks, routing, firewalls

Telecommunications broad knowledge

- Networks, Network Interfaces, Customer Operations
- Customer Care Support and Operations
- Marketing (campaign management)
- Customer Experience Management (CEM)

Domain knowledge

- UMTS (2G/3G) and LTE (4G) network architecture
- Mobile data interfaces and protocols (e.g. Gn, S11- S1-U, S6a)
- Fixed Broadband networks, architecture and protocols
- Hands on experience with any Hadoop deployment.
- Hands on experience with any blade-based computer platform

2.2 Hardware requirements

Review the hardware requirements and ensure that your environment meets the minimum standards before you attempt to install the solution.

The hardware requirements for your installation depend on how you plan to deploy the solution. Servers are required to house the products that deploy the solution. When planning your implementation, ensure that you have adequate server capacity to host the software and to deploy the solution.

The detailed system requirements information is available through the *Analytics Platform Advanced Configuration Guide*. The document recommends the following base configuration for all nodes in your cluster:

- Root partition: OS and core program files 150GB-500GB
Note: It is possible to create smaller root partition (20GB-40GB) in case of separate partitions for `/var` and `/tmp` partitions. In this case `var` and `tmp` partitions should equally split the rest of the recommended disk space and mounted on `/var` and `/tmp` mount points in `/etc/fstab`.
- Swap: Size 1X-2X system memory, but not more than 64-128GB. Swapping memory does not necessarily indicate abnormal operation. However, if the system swaps memory excessively to disk this can be an indication of problems, which should not happen normally.
- Optional boot partition: 100MB-200MB, linux boot image files - only required when grub cannot read root partition (bigger than 2TB, big inode size (>128))

For a 3-4TB hard drive, the following are sample recommendations.

#	Purpose	Size	Mounted
1	Boot	100MB	/boot
2	Swap	128GB	no mount point
3	Root	300GB	/
4	var	1.3-1.8TB	/var

5	tmp	1.3-1.8TB	/tmp
---	-----	-----------	------

2.2.1 Partitioning Recommendations for Slave Nodes

- Hadoop Slave node partitions:
 - Hadoop should have its own partitions for Hadoop files and logs.
 - Drives should be partitioned using ext3, ext4, or XFS, in that order of preference. HDFS on ext3 has been publicly tested on the Yahoo cluster, which makes it the safest choice for the underlying file system. The ext4 file system may have potential data loss issues with default options because of the "delayed writes" feature. XFS reportedly also has some data loss issues upon power failure. Do not use LVM; it adds latency and causes a bottleneck.
- On slave nodes only, all Hadoop partitions should be mounted individually from drives as `"/mnt/d[0-n]"`.

Hadoop Slave Node Partitioning configuration example

- `/root` - 150-500GB (ample room for existing files, future log file growth, and OS upgrades) in case of bigger disks the rest of the disk may be assigned to a separate partition and mounted under `/mnt/` as Hadoop storage.
- `/mnt/d0/` - [full disk] first partition for Hadoop to use for local storage
- `/mnt/d1/` - second partition for Hadoop to use
- `/mnt/d2/` -

2.2.2 Redundancy (RAID) recommendations

- Master nodes - Configured for reliability (RAID 1, RAID 10, dual Ethernet cards, dual power supplies, etc.)
- Slave nodes - RAID is not necessary, as failure on these nodes is managed automatically by the cluster. All data is stored across at least three different hosts, and therefore redundancy is built in. Slave nodes should be built for speed and low cost.

2.3 Software requirements

Review the software requirements and ensure that your environment meets the minimum standards before you attempt to install the solution. The detailed system requirements information is available through the Software Product Compatibility Reports web site (click [here](#) for the PCI report).

The following table summarizes the software requirements.

Software required	Software version
IBM Predictive Customer Intelligence	1.1.1

IBM Open Platform (IOP)	4.1
IBM Big Insights	4.1
IBM Analytics Accelerator Framework for CSP	1.0.4
IBM Streams	4.1.1

2.4 Media packaging

The solution is an integrated, multiple-product solution that may include software installed on virtual machine images and other associated software and documentation. Your purchase agreement determines how the solution is delivered to you.

The components and/or virtual images that comprise the solution are packaged on physical media. They are delivered on a hard drive that you can use for deployment.

The fixpack release is available from [Fix Central](#). Fixpack version 1.4.0.1 is available [here](#).

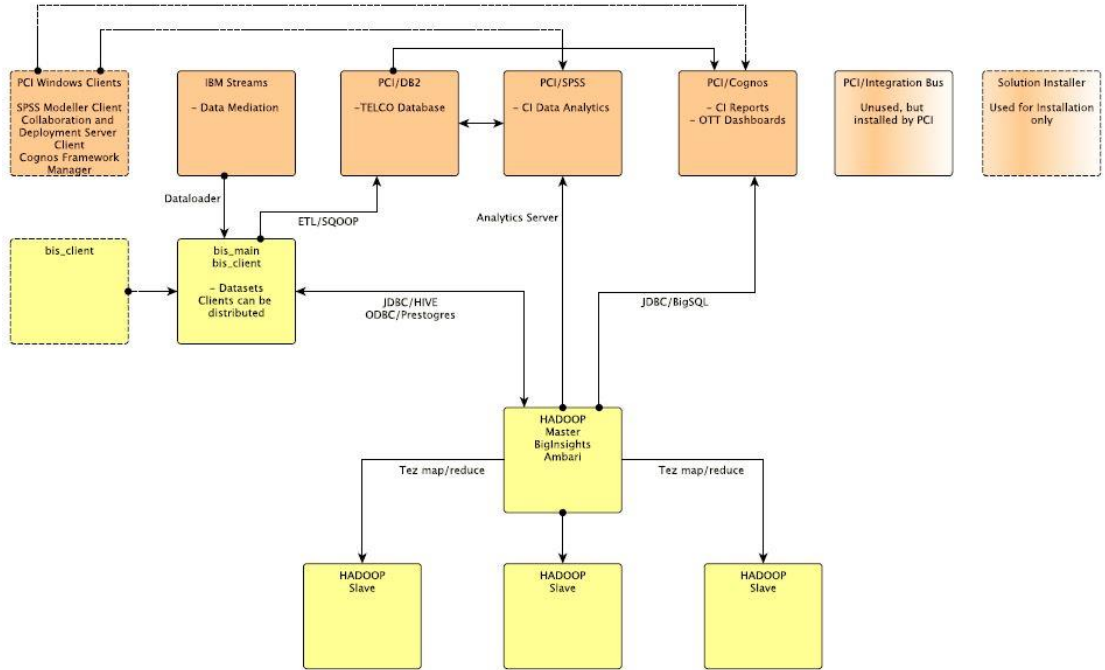
2.5 Deployment servers

IBM Customer Insight for CSP works with a typical deployment of Predictive Customer Intelligence 1.1.1.

Figure 2 displays a typical IBM Customer Insight for CSP deployment configured as a three node Hadoop environment. The functional components are deployed to physical/virtual nodes. The recommended node names are shown in parenthesis, and the nodes should be mapped to IP addresses in the `/etc/hosts` files in the environment.

Note The PCI components may be deployed as individual VMs on a single physical node. Also, the IBM Streams component may be distributed over a number of additional nodes; this has not been reflected in the diagram.

Figure 2 Typical Customer Insight for CSP Deployment



Note: PCI can be hosted on a single node for development and test purposes, or as four VMs on a single physical server.

- PCI/Cognos node (pcibinode) is the PCI business intelligence node where Cognos Analytics is installed. Cognos content should be deployed to the business intelligence node.
- PCI/DB2 node (pcidbnode) where DB2 is installed. The database content should be deployed to the DB2 node.
- PCI/SPSS node (pcipanode) where data analytics and SPSS models should be deployed.
- PCI/Windows Client
The SPSS models are also required by the SPSS Client products installed on the Windows node. As the Solution Installer works on Linux machines only, the SPSS models will need to be manually copied from the PCI/SPSS node to the PCI Windows Client machine.
- PCI/Integration Bus node is the integration manager node. This node is not used by Customer Insight for CSP but is installed by PCI.

There are two ways to install PCI - manual installation or using the Solution Installer. The Solution Installer is typically used in deployments.

The Hadoop Master node (aafnode) is where AAF, IBM Open Platform and IBM Big Insights is installed. The solution datasets should be deployed to the aafnode node.

2.6 Deployment process

IBM Customer Insight provides installation instructions and various checklists that can be used to guide you through the deployment process. It is important to understand the entire deployment process before attempting to implement the solution.

Because some products in the solution have dependencies on other products or components, the software package must be deployed and configured in a specific sequence.

Review the following steps, which describe the general process for deploying the solution:

- Perform the preliminary steps needed to prepare the deployment environment provided in the Planning for deployment section of this document.
- Gather the network configuration information that will be requested during deployment.
- Perform the following steps for each server.
 - Deploy the software to the host machine.
 - Perform the post-deployment verification checks.
 - If necessary, perform the post-deployment configuration steps required.

Additional detail about these steps is provided in the Deploying the solution section of the document. The following deployment aids are available to assist in the deployment process.

- Sample Deployment Sequence Worksheet that you can use to review the installation flow. Refer to section 6.1.
- Sample Data Required Worksheet used to gather and record the individual parameters that are needed. Refer to section 6.2.
- Detailed instructions for installing in Section 3: Deploying the solution.

2.7 Deployment dependencies

The component products in the solution work together to provide customer insight to data. Some components have dependencies on other products in the solution.

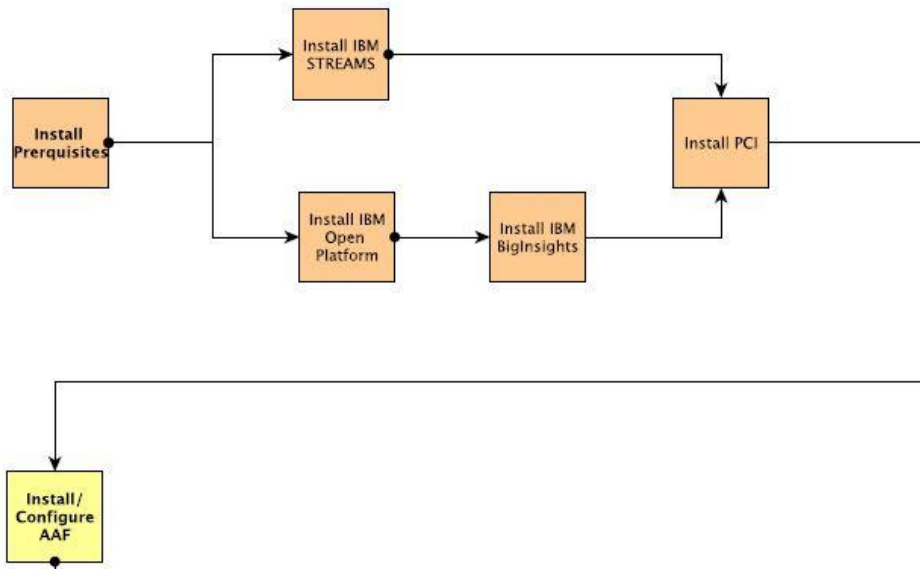
The dependencies between components affects deployment and maintenance of the servers included in the solution. Deployment must be performed in a specific sequence.

Figure 3 Deployment dependencies displays the dependencies for Telecom Analytics (TA) and Analytics Accelerator Framework (AAF).

Software required	Installation instructions
IBM Predictive Customer Intelligence 1.1.1	Read here

<p>IBM Open Platform (IOP) 4.1 and IBM Big Insights 4.1</p> <p>The Big Insights components required are:</p> <ul style="list-style-type: none"> • All of the contents of the BigInsights Analyst module.(Big SQL, BigSheets, and the BigInsights® Home services) • All of the contents of the BigInsights Data Scientist module. • All of the contents of the BigInsights Enterprise Management module 	<p>Read here</p> <p>The BI components required are available from here</p>
<p>IBM Analytics Accelerator Framework Platform 1.0.4</p>	<p>Read here (Internal wiki page link to IBM Now Factory Analytics documentation)</p>
<p>IBM Streams</p>	<p>Read here</p>

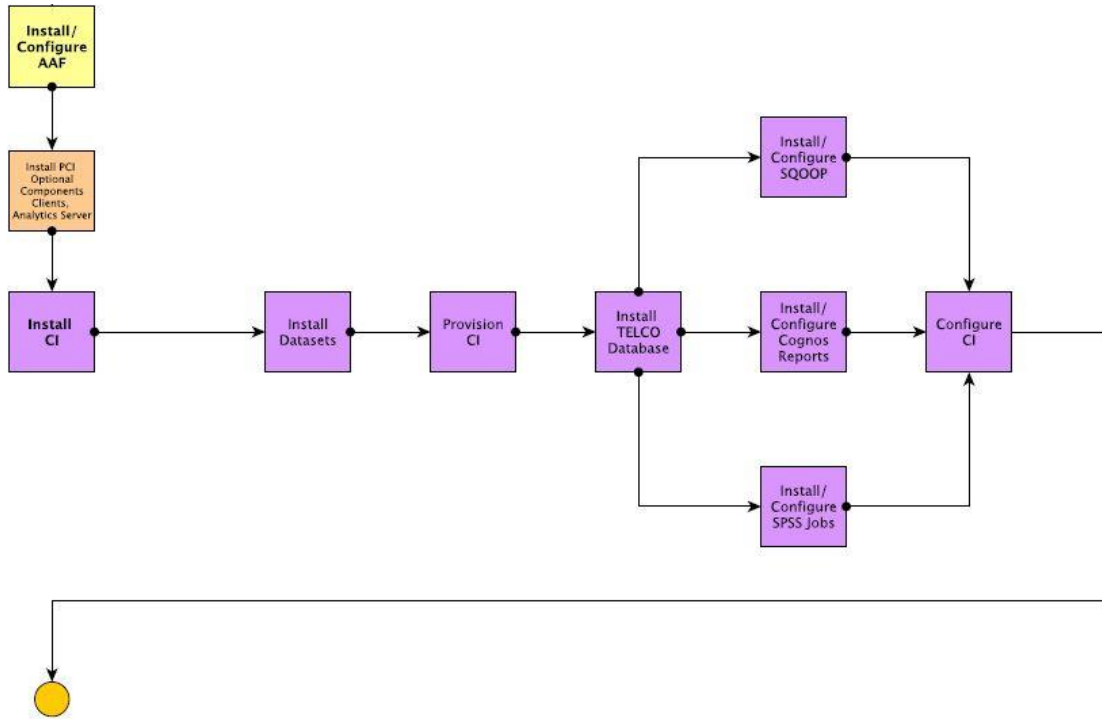
Figure 3 Deployment dependencies



2.8 Deployment sequence

Because of dependencies that exist between some of the products included in the solution, the software images for the products must be deployed and configured in a specific sequence.

Figure 4 Customer Insight for CSP deployment sequence



The following deployment sequence is typically required:

- First deploy the Analytics Accelerator Framework platform.

AAF dependencies are outlined in Figure 3 Deployment dependencies.

The installation process is documented in the *CNA 9.1 Installation and Configuration Guide*, available from the IBM Now Factory [documentation page](#).

Refer to the [IBM Open Platform installation documentation](#) for pre-requisite steps.

Note: The AAF deployment installs a basic set of standard IBM Open Platform services.

The Big Insights BigSQL service, available in the Big Insights value-added package, must also be installed for use with the Customer Insight for CSP solution.

The Big Insights BigSQL installation must be completed after the base IBM Open Platform is installed. For information on deploying the Big Insights BigSQL service, read the [Big Insights deployment](#) documentation and notes in section 3.1.2.

- Then provision data into the AAF platform as documented in the *CNA Mediation Operations and Configuration Guide* available from the IBM Now Factory [documentation page](#).

At this point, you have successfully installed and configured AAF and are ready to begin the Customer Insight for CSP deployment steps.

- Deploy and configure the optional PCI SPSS components, after completing the Predictive Customer Intelligence 1.1.1 installation, as described in section 3.1.1.

- Download the Customer Insight for CSP solution from fixcentral as described in section 3.2.
- Install and use the Solution Installer to deploy the Customer Insight for CSP solution, as described in 3.3.
- For a master client Analytics Platform installation, configure BigSQL access, as described in section 3.4.
- Deploy the datasets, provision the Customer Insight solution, install the Telco database, and complete the dataset generation configuration as described in sections 3.5 to section 3.10.
- Deploy the Customer Insight for CSP content (SPSS Models, Cognos Dashboards, Cognos Reports) as described in section 3.11.
- Perform any post installation configuration as described in section 4.

2.9 Post deployment verification

After deployment, it is important to verify that the software is deployed as expected.

Perform the checks documented at the end of each section to determine whether deployment was completed successfully.

2.10 Deployment worksheets

Use the deployment worksheet to record information that you must supply when you deploy the solution. Refer to section 6.1 and 6.2.

2.11 Provisioning data

Provisioning is the process of pre-loading supplemental data into the database in order to support the loading, mediation and analysis of the network fact data subsequently. For more information see Section 3.6 Provisioning data.

3 Deploying the solution

Deploying the solution includes the following tasks:

- Complete the preliminary steps that are needed to obtain the appropriate hardware and software and then prepare the deployment environment.
- Deploy the products and components in the solution, in the required sequence.
- Verify that the products and components in the solution have been deployed correctly.

Perform post-deployment configuration steps that are needed to customize your implementation.

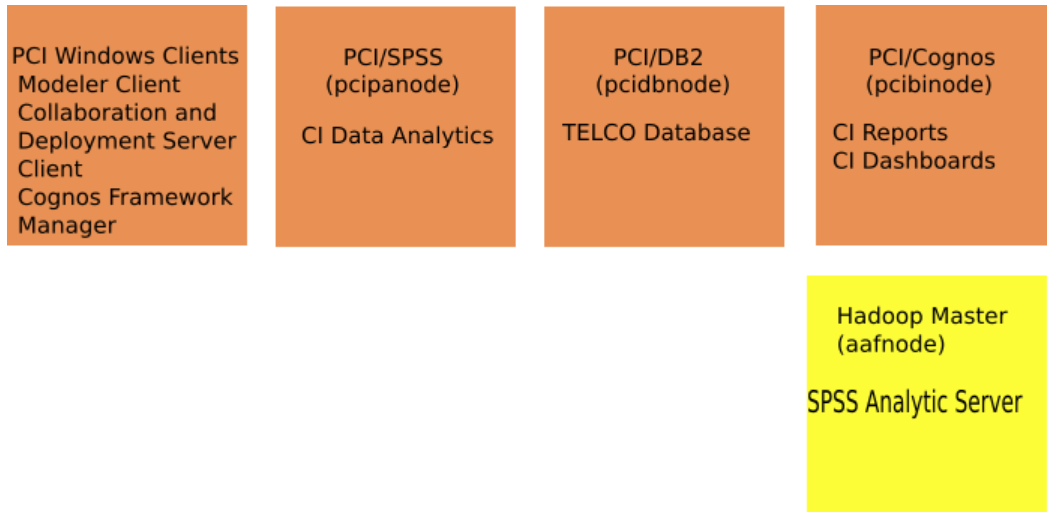
3.1 Before deploying products and components

You must first create the infrastructure needed to deploy the solution. Ensure that the base architecture products and components as outlined in section 2.7 have been installed and are operational.

3.1.1 Installing the IBM PCI optional components

The required components of IBM Predictive Customer Intelligence, listed in 1.5, must be installed. Figure 5 highlights where the PCI components are installed in a typical configuration on both the Windows Client and Linux machines.

Figure 5 Sample PCI optional component installation



The following optional components must also be installed:

- SPSS Analytic Server
- SPSS Collaboration and Deployment Services Client
- SPSS Modeler Client

Complete the following steps to install the optional components.

On the PCI Windows Client machine

SPSS Collaboration and Deployment Services Client	Installation instructions
SPSS Modeler Client	Installation instructions

On the PCI Linux machines

<p>SPSS Modeler Server on the PCI/SPSS node (pcipanode) at: /usr/IBM/SPSS/ModelerServer/17.1/</p> <p>Note: The data folder in the above path is required in order to run the SPSS models</p>	Installation instructions
SPSS Collaboration and Deployment Service Repository on the PCI/SPSS node (pcipanode)	Installation instructions
SPSS Analytic Server on the Hadoop Master node (aafnode)	Installation instructions

3.1.2 Installing the IBM Big Insights - BigSQL value added service

The IBM Open Platform must be installed as outlined in section 2.7. When using the Ambari dashboard to install Big SQL ensure the services outlined Figure 6 are selected.

Figure 6 Service Selection

<input type="checkbox"/> Service	Version	Description
<input checked="" type="checkbox"/> HDFS	2.7.1	Apache Hadoop Distributed File System
<input checked="" type="checkbox"/> YARN + MapReduce2	2.7.1	Apache Hadoop NextGen MapReduce (YARN)
<input checked="" type="checkbox"/> Hive	1.2.1	Data warehouse system for ad-hoc queries & analysis of large datasets and table & storage management service
<input checked="" type="checkbox"/> HBase	1.1.1	Non-relational distributed database and centralized service for configuration management & synchronization
<input checked="" type="checkbox"/> Pig	0.15.0	Scripting platform for analyzing large datasets
<input checked="" type="checkbox"/> Sqoop	1.4.6	Tool for transferring bulk data between Apache Hadoop and structured data stores such as relational databases
<input checked="" type="checkbox"/> Oozie	4.2.0	System for workflow coordination and execution of Apache Hadoop jobs. This also includes the installation of the optional Oozie Web Console which relies on and will install the ExtJS Library .
<input checked="" type="checkbox"/> ZooKeeper	3.4.6	Centralized service which provides highly reliable distributed coordination
<input type="checkbox"/> Flume	1.5.2	A distributed service for collecting, aggregating, and moving large amounts of streaming data into HDFS
<input checked="" type="checkbox"/> Ambari Metrics	0.1.0	A system for metrics collection that provides storage and retrieval capability for metrics collected from the cluster
<input checked="" type="checkbox"/> Kafka	0.8.2.1	A high-throughput distributed messaging system
<input checked="" type="checkbox"/> Knox	0.6.0	Provides a single point of authentication and access for Apache Hadoop services in a cluster
<input checked="" type="checkbox"/> Slider	0.80.0	A framework for deploying, managing and monitoring existing distributed applications on YARN
<input checked="" type="checkbox"/> Solr	5.1.0	Solr is the popular, blazing fast open source enterprise search platform from the Apache Lucene project
<input type="checkbox"/> Spark	1.5.1	Apache Spark is a fast and general engine for large-scale data processing

Note: It is recommended that you install the Big SQL service with at least two nodes in the cluster to see the best performance with at least one node designated as the Big SQL master.

To deploy the Big Insights – BigSQL service, complete the following steps:

1. Complete the preparation steps for BigSQL

```
# bigsql-precheck.sh
```

2. Confirm you have Hive connectivity from the node where BigSQL will be installed.

```
# su hive
```

```
hive> show tables;
```

3. Ensure that home is not mounted by displaying information on mount points

```
# mount|grep suid
```

```
gvfs-fuse-daemon on /root/.gvfs type fuse.gvfs-fuse-daemon
(rw,nosuid,nodev)
```

4. Check if you can ssh as root into each node without being prompted for a password. If not, as root, run:

```
# ssh-keygen
```

```
# ssh-copy-id root@nodeaddress
```

Verify that you can now passwordlessly `ssh` into that node.

5. Open a browser and access the Ambari server dashboard.

`http://<server-name>:8080`

Note: Ambari is installed as part of the IBM Open Platform installation during the AAF deployment process on the Hadoop Master node (aafnode).

6. In the Ambari web interface, click Actions > Add Service and select the Big Insights - Big SQL service.

To check that Big Insights is installed successfully, complete the [steps](#) to validate your installation. For more information, read the documentation on how to [install the Big Insights – BigSQL service](#).

3.2 Downloading the Customer Insight for CSP solution

Download the solution fixpack from Fix Central to a Red Hat node from where you plan to install the solution. The Solution Installer node, in Figure 2, can be used as a launchpad node for deployment. Download

3.3 Deploying the Customer Insight for CSP solution

The Solution Installer is used to deploy the solution. PCI also has a Solution Installer that is used to deploy and install the PCI product components.

3.3.1 Prerequisite steps

Prior to running the Solution Installer, the following prerequisite steps must be completed:

1. Install and configure the base architecture components required by the solution. (Refer to section 1.4 outlining the components and section 2.7 showing the installation dependencies).
2. Download the Customer Insight for CSP solution to the node selected in Section 3.2. The node is used to run the Solution Installer and deploy the content to the PCI and Hadoop Master node (aafnode).
3. Modify the sudoers file for the user who runs the installation as described in 3.3.2. A root user is required to run the installation.
4. Understand the deployment environment. There are a combination of nodes that the Solution Installer installs to. The installer must determine where each product component will reside. A sample deployment is provided in section 3.3.3.

Note Firewall configurations required to allow the Solution Installer to launch and deploy the solution are handled in section 3.3.3 when deploying the solution content using the Solution Installer.

3.3.2 Modifying the permissions file

Run the steps on the PCI DB2 node, PCI Cognos Analytics node and Hadoop Master node.

A root user or a user with sudo permissions on each node is required to run the installation.

To install with sudo user permissions the user must be added to the sudoers file.

1. Log in as root user.
2. Enter the following command to open the sudoers file for editing:
`visudo -f /etc/sudoers`
3. Locate the following line:
`Defaults requiretty`
4. Comment out the `Defaults requiretty` by typing the hash symbol (#) in front of `Defaults requiretty` to comment out the line. The line will appear as `#Defaults requiretty`
5. If you run the installer as a user with sudo user permissions, go to the end of the file, and add the following line for your user:
`username ALL=(ALL) NOPASSWD: ALL`
6. Save and close the file.
7. Repeat these steps on each computer on which you install a Customer Insight for Communication Service Providers component.

3.3.3 Running the Solution Installer to deploy the Customer Insight solution

Deploy the solution as follows:

1. Log on to the node where the Customer Insight for CSP product package is downloaded.
2. Decompress the solution package. Extract the package to location:
`/opt/IBM/IS_CSP_Customer_Insight_1.0.4/SolutionInstaller.`
3. If the Solution Installer (including the IBM Predictive Customer Intelligence Solution Installer) has been run on the node previously, run the following command on each of the PCI and Hadoop Master nodes:

```
./cleanupClient.sh.
```

Note: Client in the above command refers to the nodes you are deploying to: the PCI / Cognos business intelligence node (pcibinode), the PCI DB2 node (pcidbnode) and the Hadoop Master node (aafnode).

Run `./cleanup.sh` on the deployment node from which the Solution Installer is launched.

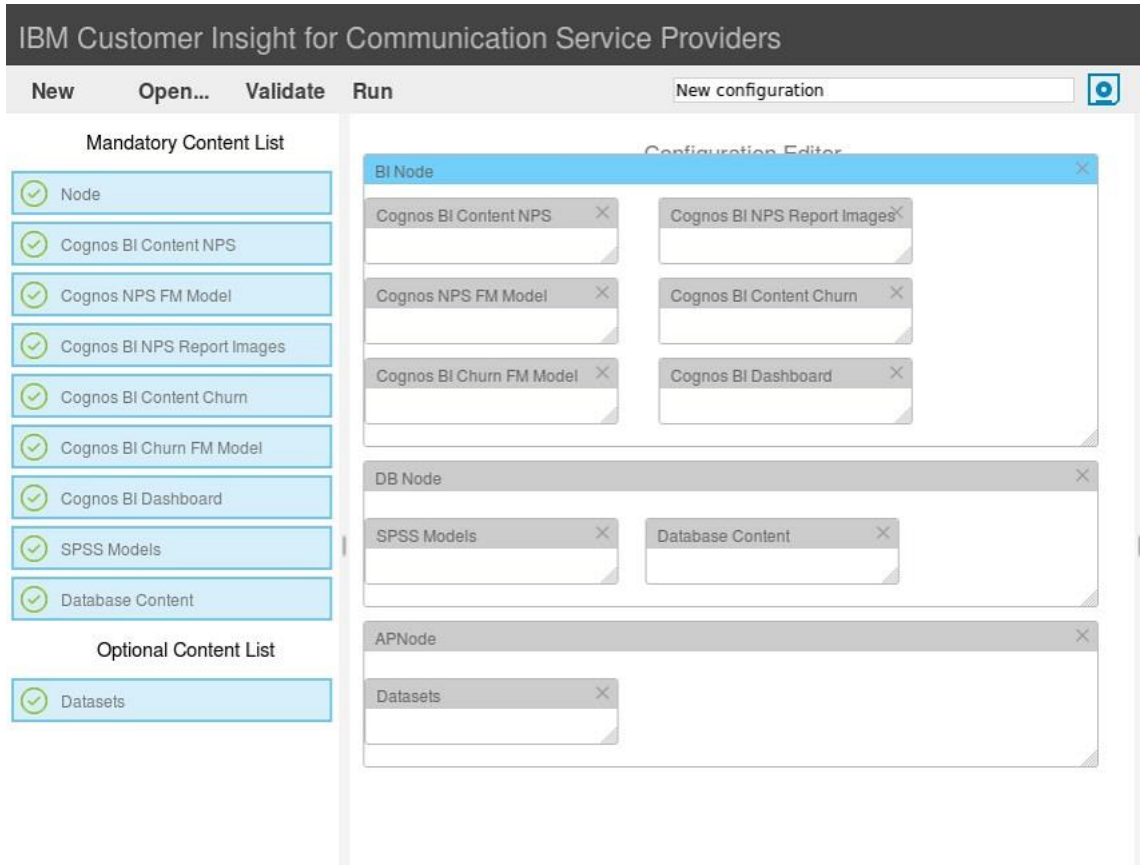
4. Navigate to the Solution Installer directory in the following location:
`/opt/IBM/IS_CSP_Customer_Insight_1.0.4/SolutionInstaller.`
5. Open the ports that are required by the Solution Installer by running the following command:
`./firewall.sh`
6. Run the setup command to start the installation process:

```
sh setup.sh username first_name last_name email password
```

This command creates a user with the details supplied (for example admin / admin) for accessing the Solution Installer web server. The web server is started and the URL for the Solution Installer displays in the command line window.

7. A browser window should open automatically. If it doesn't, copy and paste the URL provided in the output of the command from step 6 into a web browser and bookmark it.
8. Read the license agreement.
Accept the license agreement to begin the Solution Installer deployment otherwise the installation process will end.
9. Create the required nodes for the deployment process, enter valid credentials and drag and drop the solution content to each node.

Sample deployment:



Note: The node names provided in the screen above are an example only. The values are freeform text fields that can be specified by deployment teams. In the example:

- **BI node** is where Cognos Analytics is installed. All Cognos content should be deployed to the business intelligence node.
- **DB node** is where DB2 is installed.

The database content and SPSS models should be deployed to the DB2 node.

Note: The SPSS models are not used on the DB2 node but the Solution Installer works on Linux machines only. The SPSS models are required by

the SPSS Client products installed on the Windows node. The SPSS models will need to be manually copied from the DB2 node to the Windows Client machine.

- **AP node** is the Hadoop Master node where AAF, IBM Open Platform, and IBM Big Insights is installed. Database content is installed to this location also.

By following the sample deployment, the installer will have a reference point to where each component is deployed to during the installation and configuration process.

10. Finally once each component is assigned to a node and all green check marks are displayed in the left panel select the Run button in the toolbar.
11. Once the deployment has completed successfully a success pop up message is displayed.

3.4 Configuring BigSQL access in a master/client Analytics Platform installation

On a master/client installation of Analytics Platform, BigSQL access is not available on the client node. The CI and AAF `install_telsol` script requires BigSQL access to complete operations such as synchronizing tables and views to BigSQL schemas on the master node.

To facilitate this operation, the Customer Insight for CSP solution contains a jar file and SQL scripts that enable the installer to complete the required operations on the client. One additional AAF deployment step is required as outlined in section section 3.4.2.

The Customer Experience reports require a defined function that must be manually installed on the master node.

3.4.1 Running `install_telsol` for a master / client installation

While running the `install_telsol` script on the client node, on a client master configuration or on the master node of a single node installation, the BigSQL connection information is requested. When the credentials are entered, the installer runs the installation steps on the master node.

3.4.2 Deploying CEA dataset

After running `install_telsol` on a client master Analytics Platform environment, install the `cea-cognos-reports-udf-*.jar` manually on the master node. To install the jar on the master node:

1. Move the jar from the `tmp` directory on the client node to the `tmp` directory on the master node.
2. Connect to `bigsql` and install the jar using the following commands on the master node.

```
db2 connect to bigsql
```

```
db2 "call sqlj.install_jar('file:/tmp/<JAR_FILE>', cea-cognos-reports-udf) "
```

3. After running the command, restart bigsql.

Note: Replace <JAR_FILE> with the full name, including version, of the cea-cognos-reports-udf jar.

Updating the BigSQL connection

If the BigSQL connection information changes or was entered incorrectly then update the connection settings.

1. Run the following command entering the required information when prompted.

```
[root@dubxpcvm3126 master-config]# ./configureBigSQLConnection.sh
BigSQL Connection Configuration
[Enter the BigSQL Host, followed by [ENTER]: hostname.mul.ie.ibm.com
[Enter the BigSQL Port, followed by [ENTER]: 32051
[Enter the BigSQL Username, followed by [ENTER]:bigsql
Enter the BigSQL Password, followed by [ENTER]:
```

2. Enter the BigSQL hostname, port (example: 32051), username (example BigSQL) and password.

If the connection fails on installation exit the installation resolve the issue and then re-run the `install_telsol` script uninstalling and installing each installed dataset.

3.4.3 Running a BigSQL query or update statement on the client

The BigSQL Linux script is located in `/opt/tnf/apps/dataset-common/scripts` on the client node of a master client configuration or on the master node of a single node installation. The following summarises a sample running of the script:

```
./runBigSQLStmt.sh query stmt "select * from tnf.dim_cell"
./runBigSQLStmt.sh query file "/opt/tnf/apps/sampleFile.txt"
./runBigSQLStmt.sh update stmt "CALL
SYSHADOOP.HCAT_SYNC_OBJECTS('TNF', 'CGR_APPLICATION', 't',
'REPLACE', 'CONTINUE');"

```

Note: If the wrong type of update is specified for a query, then the code will fall back and execute a query after the failed update and vice versa.

The query logs are located at: `/opt/tnf/apps/dataset-common/log/tools.log`

3.4.4 Connecting Cognos Reports to BigSQL

Ensure Cognos Reports always connect to the Master node to read the data required by the reports. The client node does not support read access for Cognos reports.

Check the query logs located `/opt/tnf/apps/dataset-common/log/tools.log` and request the master node administrator to log on to the master node and provide the `bigsql-sched.log` file. The file is generally located in the `/var/ibm/bigsql/logs` directory. Ensure the `bigsql` user has `hdfs` access to read files.

3.5 Deploying the Customer Insight datasets

Use the following procedure to install your solution datasets.

1. As the root or sudo user, start the installer from the directory where the installation package is uncompressed:

```
$ sudo ./analytics-platform/install_telsol.sh
```

When installing, a check is performed for an existing use case. For example:

```
-- Installing Customer Behaviours use case...
Checking for existing installation of Customer Behaviours use
case...
```

```
WARNING: Previous installation of Customer Behaviours use
case found: customer-behaviours -
```

```
Remove? (y/n):
```

Most use cases can run using default configuration values.

If non-default configuration values are required for any use case, the values can be updated after installation and before the use case is run.

The Net Promoter Score use case is an exception. The installer requires that you enter values for the following NPS configuration settings as there are no valid default values.

Name	Value
NPS_BUCKET_SIZE	Set this to as high as possible based on the size of your cluster (reducers), for example, 30.
OPERATOR_MCC_MNC	Example: '1234','3198'

For some use cases, data must be loaded into the database before the use case is run. Data should be in the form of CSV files. Before attempting to provision your data, the installer pauses to ensure that these files are available in the correct directory.

```
-- Loading provisioned data...
```

```
Provisioning data to be loaded must be stored as CSV files in
/opt/tnf/apps/bis-main-var/bisprovisioning-tool/csv_files/ -
if you have any data to provision, please copy it to that
directory now and press any key to resume...
```

If the provisioned data CSV files have not already been copied to that directory, do it now and press any key to resume the installer which will attempt to load them. It will determine which files should be loaded into which tables

based on the filenames and ask for confirmation that the correct files are being loaded to the correct tables.

Table to load data to	Data file
cb_category	csv_files/cb_category_20150320.csv
cea_unacceptable_trend_config	csv_files/cea_unacceptable_trend_config_00000000.csv
nps_cell_provisioning_table	csv_files/nps_cell_provisioning_table_20150331.csv
nps_crm_provisioning_table	csv_files/nps_crm_provisioning_table_20150331.csv
nps_device_provisioning_table	csv_files/nps_device_provisioning_table_20150331.csv
nps_survey_table	csv_files/nps_survey_table_20150429.csv
performance_assign_config	csv_files/performance_assign_config_00000000.csv

Load data to tables according to the list? (y/n):

Review this information and provide the appropriate response. Upon successful installation, the following message is displayed:

```
-- IBM Telecom Solutions installed completed successfully!
```

3.5.1 Validating the dataset deployment

To view the progress of the dataset deployment, check the `install_telsol.log` in the current directory.

3.5.2 Verifying dataset scheduling

Data processing involves the recurring execution of a series scheduled tasks. After a default installation, datasets are run on the following scheduled basis:

Dataset ¹	Schedule
customer-profile-data-setup	00.30, every Monday morning (weekly tables)
ott-applications	02.30, every Monday morning (weekly tables)
customer-profile	02.30, every Monday morning (weekly tables)
user-profile	04.30, every Monday morning (weekly tables)

¹ **Note:** The table contains a sample of datasets and associated intervals. Not all deployed datasets are included in this table.

customer-behaviours	*.45, every hour (hourly tables) 03.15, every day (daily weighted interest tables) 03.30, every day (daily other tables) 00.30, every Monday morning (weekly tables)
---------------------	---

Datasets are run on an hourly, daily, or weekly basis using cron jobs.

To verify dataset scheduling, and view or modify cronjobs, do the following:

1. Log on as `boss` user to the Hadoop Master node (aafnode).
 - Note:**The `boss` user has already been created when AAF is installed.
2. To edit the cronjobs, run the command:

```
crontab -e
```

To verify the full list of cron jobs, run the command:

```
crontab -l
```

The customer-profile-data-setup dataset triggers the customer profile models.

Note:

Run the customer-profile-data-setup dataset after the churn and NPS datasets are run. For more information, see Scheduling SPSS Job Triggers.

3.6 Provisioning data

Provisioning data is loaded into a set of Hive tables which are used specifically for storing provisioning data. A general description of provisioning is provided in this section.

For more information, refer to the provisioning guides available from the [IBM Now Factory documentation page](#).

3.6.1 Types of provisioning data

The following are the types of provisioning data.

1. Subscriber CRM dimensions (for mobile subscribers)
2. Network interface data

Constant dimensions (constant, operator-independent values relating to the interface (e.g., network protocol constants) - the default values provided should be used and should not be modified unless there is a system upgrade)

Modifiable dimensions (operator-independent values relating to the interface - the default values provided may be used but can be modified if required)

Operator-specific dimensions (values for these dimensions MUST be manually specified for each deployment as they are specific to the operator, e.g., cell, APN, GGSN)

3. Usecase-specific dictionary data:

Constant values

Modifiable values (e.g., mapping of application/domain names to behaviour categories)

Operator-specific data (e.g., cell data, device data)

Note: A default set of provisioning files is shipped in `/opt/tnf/apps/bis-main-var/bis-provisioning-tool/csv_files` covering all the above categories. For modifiable and operator-specific data, only sample files are provided. The files **MUST** be replaced with new versions appropriate to the deployment.

Provisioning data is stored in a set of Hive tables according to the following naming conventions:

Data	Naming convention
Subscriber CRM data	DIM_SUBSCRIBER_MOBILE - mobile subscribers
Interface dimensions	One table per dimension "dim_<XXX>", For example: XXX=apn
Usecase-specific dictionary data	No naming convention applies

3.6.2 When to provision

Data is provisioned in two ways: manually or automatically.

Manual provisioning

Manual provisioning should be performed at the following times:

- Before installing the Telco Solutions usecases
- During the installation of Telco Solutions usecases (the installation script prompts the user to copy any non-default provisioning data files to the source directory - it will then load all provisioning files in that directory)
- After the installation of Telco Solutions usecases but before loading any network data

Note: It is essential that manual provisioning of all subscriber CRM and network interface provisioning data is carried out **BEFORE** network fact data loading commences as, during the process of loading and mediation carried out by Streams, network data is correlated with entities in the provisioned data in order to enrich the output.

Dictionary data for each usecase must be loaded before that particular usecase is run.

Auto-provisioning

Automatic provisioning occurs during network fact data loading. Provisioning data values are automatically extracted from the network fact data during the loading of that network data and no manual intervention is needed.

It is only supported for certain tables such as `dim_apn`, `dim_collector` and `msisdn_imsi` (`msisdn_imsi` is only auto-provisioned from Gn/LTE TDR and aggregated control plane tables - not from voice/SMS tables).

3.6.3 Provisioning data before or after installing Telco Solutions

If provisioning data before or after installing Telco Solutions, and not using the Telco Solutions install script, two options are available:

1. Load all data in one go using the `load-all.sh` command
2. Load each provisioning data file individually using the `load.sh` command

Note: If provisioning is performed after you install Telco Solution, disable cron jobs to ensure they do not run until the data is provisioned. Once data is provisioned, cron jobs should be enabled again.

Loading all at once

To load all tables at once:

1. Copy all provisioning files to `/opt/tnf/apps/bis-main-var/bis-provisioning-tool/csv_files`.

The directory must contain just one file per table to be loaded of the form `<table_name>_<date in form YYYYMMDD>.csv`, e.g., `cb_category_20170320.csv`

2. Go to the folder:

```
cd /opt/tnf/apps/bis-main-var/bis-provisioning-tool
```

3. Run the load command

```
./load-all.sh
```

Log output is written to the console and also to log files in the `./log` directory.

Note: `load-all.sh` will return a successful return code even if there are some errors - you must check the console or the log files for ERROR messages

4. Verify that the table is populated by querying the database.

Each line in the csv file (excluding the header line) should result in a single record being created in the corresponding table (unless there were invalid lines). To verify that data was loaded correctly, check that the number of records in the table is equal to the number of data lines in the file. Individual lines can then be compared against individual records

Loading each table individually

To load each table individually:

1. Copy the corresponding csv file to `/opt/tnf/apps/bis-main-var/bis-provisioning-tool/csv_files`

2. Go to the folder:

```
cd /opt/tnf/apps/bis-main-var/bis-provisioning-tool
```

3. Run the load command:

```
./load.sh -f csv_files/<file name>.csv -t <table name>
```

Note: The "tnf" qualifier is not required.

Log output is written to the console and also to log files in the ./log directory

4. Verify that the table is populated by querying the database.

Each line in the csv file (excluding the header line) should result in a record being created in the corresponding table (unless there were invalid lines). To verify that data was loaded correctly, check that the number of records in the table is equal to the number of data lines in the file. Individual lines can then be compared against individual records

Note: Run manual provisioning on the Hadoop master node (aafnode) where the Analytics Platform server is installed.

3.7 Creating the Telco Database

During deployment of the Customer Insight for CSP solution to the nodes, the Solution Installer automatically creates the Telco database.

Prior to running the Solution Installer, ensure that DB2 is started on the PCI DB2 node. Hive should be started on the Hadoop Master node.

The Telco database is required by the Database Loader in order to migrate data from Hive to DB2. The Churn and NPS models run on DB2 and create predictive outputs required by the Cognos Churn and NPS reports.

3.7.1 Verifying the Telco database installation

To verify if the Telco database has been created, connect to the database and list the tables.

1. Log on to the PCI DB2 node (pcidbnode) with a user ID that has access to the IBM DB2@ database. The DB2 username and password are set when PCI was installed. For example: db2inst1.
2. Connect to the database:
db2 connect to TELCO;

```
Database Connection Information
Database server      = DB2/LINUX08664 10.5.5
SQL authorization ID = DB2INST1
Local database alias = TELCO
```

3. List the tables in the schema NPS:
db2 list tables for schema NPS

```

Table/View          Schema      Type      Creation time
-----
FACTOR_IMPORTANCE  NPS         T         2016-07-18-11.34.01.240026
FACT_SCORE         NPS         T         2016-07-18-11.34.01.286355
MASTER_CUSTOMER   NPS         T         2016-07-18-11.34.01.333844
MASTER_DATA_ACTIVITY_LOCATION NPS         T         2016-07-18-11.34.01.378746
MASTER_DEVICE     NPS         T         2016-07-18-11.34.01.423606
PREDICTED_NETWORK_NPS_SCORE NPS         T         2016-07-18-11.34.01.508327
PREDICTED_SUBSCRIBER_NPS_SCORE NPS         T         2016-07-18-11.34.01.465209
STAGING_NPS_SCORE_TABLE NPS         T         2016-07-18-11.34.01.122448

8 record(s) selected.
    
```

4. List the tables in the schema BBCI
 db2 list tables for schema BBCI

```

Table/View          Schema      Type      Creation time
-----
CGR_DEVICE         BBCI        T         2016-07-18-11.34.02.123683
CHURN_DATA         BBCI        T         2016-07-18-11.34.02.068338
CHURN_MANAGEMENT   BBCI        T         2016-07-18-11.34.02.419369
CHURN_MANAGEMENT_HISTORY BBCI        T         2016-07-18-11.34.02.471134
CHURN_TREND        BBCI        T         2016-07-18-11.34.02.210732
SUBSCRIBER_BILLING BBCI        T         2016-07-18-11.34.02.259749
SUBSCRIBER_CARE    BBCI        T         2016-07-18-11.34.02.365935
SUBSCRIBER_CRM     BBCI        T         2016-07-18-11.34.02.166749
SUBSCRIBER_LEVEL_CX_SCORE_WEEK> BBCI        T         2016-07-18-11.34.02.312822
TOP_CHURN_FACTORS  BBCI        T         2016-07-18-11.34.02.525788

10 record(s) selected.
    
```

Once tables are returned from the list tables command, then the Telco database has been successfully created.

3.8 Deploying the Database Loader

The Database Loader is automatically installed by the Solution Installer during the deployment of the solution. The Database Loader is installed to the /opt/tnf/apps/telco-dbexport directory on either the master Analytics Platform node (in a master AP installation) or the client Analytics Platform node (in a master client AP installation).

3.8.1 Updating the Database Loader NPS and Churn configuration files

The NPS (nps_config.properties) and Churn Database Loader configuration files (churn_config.properties) are located in folder /opt/tnf/apps/telco-dbexport/scripts/conf. The configuration files are located on the Hadoop node (aafnode).

Table 1 and Table 2 contains recommended settings for the configuration files. Update the files as described in the tables in order for the job to extract data from the Hadoop instance and load the data into DB2.

Table 1: Recommended NPS configuration settings

Option	Value	Description
HiveHost	localhost:10000/default	Hive host for accessing

		the hive instance.
HiveUser	<username>	Hive user with access to run hive queries.
HivePasswd	<password>	Hive password for the associated hive user.
HiveDatabase	tnf	Hive schema. Do not change.
HiveTableName	nps_score_table	Hive table name to be exported do not change.
DB2SchemaName	NPS	Do not change.
DB2Host	localhost:50000	DB2 host and port of the PCI DB2 instance.
DB2DBName	TELCO	Do not change.
DB2User	db2inst1	DB2 user with access to connect, read and write to DB2.
DB2Passwd	<password>	DB2 password for the DB2 user.

DB2StringLen	255	Do not change.
DB2TableName	staging_nps_score_table	Do not change.
DB2FactTableName	FACTOR_IMPORTANCE	Do not change.
SqoopHost	localhost:12000/sqoop	Sqoop host.
HiveDataDirectory	/apps/hive/warehouse/tnf.db	The hive data directory is determined by running desc formatted <table name> and checking the location values. The Hive data directory is the value after the port number.
HiveDeltaTableName	Nps_score_table_latest	Do not change.
SqoopRecordPerStatement	100	Do not change.
NumberOfMapJobs	5	Number of map reduce jobs used by Sqoop.
HivePrimaryKey	Job_execution_timestamp	Do not change.
SPSSEndPoint	http://localhost:9080/ process/ services/ProcessManagement	The SPSS endpoint for the PCI SPSS

		Repository. Only the Port and Host should be updated.
SPSSID	569a069593724d5600001513e1c502db7fd	Job ID of the NPS SPSS job to be triggered after a successful export. The value can be determined by right-clicking the job in Deployment Manager and selecting properties
SPSSUser	admin	SPSS Collaboration and Deployment Services user id. User to log in to the Collaboration and Deployment Services repository.
SPSSPasswd	<password>	Password for the SPSS user.
SqoopPath	/usr/iop/4.1.0.0/sqoop/bin	Sqoop path is the location of the Sqoop-

		export file on the Hadoop node. The value should not change.
--	--	--

Table 2: Recommended Churn configuration settings

The following table displays the recommended Churn configuration settings.

Option	Value	Description
HiveHost	localhost:10000/default	Hive host for accessing the hive instance.
HiveUser	boss	Hive user with access to run hive queries.
HivePasswd	<password>	Hive password for the associated hive user.
HiveDatabase	tnf	Hive schema. Do not change.
HiveTableName	subscriber_crm, subscriber_billing, subscriber_care, churn_data,subscriber_level_cx_score_weekly, cgr_device	Hive table name to be exported do not change.
DB2SchemaName	BBCI	Do not change.
DB2Host	localhost:50000	DB2 host and port of

		the PCI DB2 instance.
DB2DBName	TELCO	Do not change.
DB2User	db2inst1	DB2 user with access to connect, read and write to DB2.
DB2Passwd	<password>	DB2 password for the DB2 user.
DB2StringLen	255	Do not change.
SqoopHost	localhost:12000/sqoop	Sqoop host.
HiveDataDirectory	/apps/hive/warehouse/tnf.db	The hive data directory is determined by running desc formatted <table name> and checking the location values. The Hive data directory is the value after the port number.
SqoopRecordPerStatement	100	Do not change.

NumberOfMapJobs	5	Number of map reduce jobs used by Sqoop.
SPSSEndPoint	http://localhost:9080/ process/ services/ProcessManagement	The SPSS endpoint for the PCI SPSS Repository. Only the Port and Host should be updated.
SPSSID	569a069593724d56000001513e1c502db857	Job ID of the Churn SPSS job to be triggered after a successful export. The value can be determined by right-clicking the job in Deployment Manager and selecting properties
SPSSUser	admin	SPSS Collaboration and Deployment Services user id. User to log in to the Collaboration and Deployment

		Services repository.
SPSSPasswd	<password>	Password for the SPSS user.
SqoopPath	/usr/iop/4.1.0.0/sqoop/bin	Sqoop path is the location of the Sqoop-export file on the Hadoop node. The value should not change.
HadoopPort	9000	Hadoop port.
HadoopHost	localhost	Hadoop host.

3.8.2 Updating execution permission on the Linux scripts

Complete the following steps on the Hadoop node (aafnode):

1. Navigate to the script directory of the Database Loader folder.

```
cd /opt/tnf/apps/telco-dbexport/scripts/
```

2. Run the following commands to update permissions:

```
chmod 755 script/sqoop.sh.
```

```
chmod 755 runDBExport.sh.
```

```
chmod 755 runDBExportForCron.sh.
```

3.8.3 Base encoding the passwords in the configuration files

The Customer Insight for CSP solution uses base64 encoding so that no credentials are visible in plain text in files.

Run the following command to base64 encode the passwords in each of the <config_file_name>.properties files (where config_file_name is either churn_config or nps_config):

```
/opt/tnf/apps/telco-dbexport/scripts/encodePasswordProperties.sh -f  
conf/<config_file_name>.properties
```

Plain text passwords in the configuration file will be encoded. If a password is updated, change the password property in the configuration file to a plain text equivalent, and re-encode the updated property.

A password can be encoded by passing the property name with the `-p` flag to the `encodePasswordProperties` script.

For example:

```
/opt/tnf/apps/telco-dbexport/scripts/encodePasswordProperties.sh -f  
conf/<config_file_name>.properties -p <SPSSPasswd>
```

Note: The password encode script must be run on a configuration file prior to running `runDBExport.sh`, or the decoded plain-text passwords will be invalid and the Sqoop job will fail.

3.8.4 Manually running the Churn Database Loader job

The first time the Database Loader is run for churn it should be run manually. The reason is when the data is loaded in DB2 the Churn model is automatically triggered. The first time the model is triggered the model will fail to run successfully because the Churn model has not been trained. Ensure the model is trained after the first execution. Review the IBM SPSS Modeler documentation for more information on training a model.

Then rerun the Database Loader for Churn to ensure a successful execution.

To run the Churn Database Loader job.

1. Navigate to the main Database Loader Folder.

```
cd /opt/tnf/apps/telco-dbexport/scripts/
```

2. Run `./runDBExport.sh churn`

3.8.5 Checking the Churn Database Loader job

Validate the job has run successfully.

1. Monitor the output of the `runDBExport.sh` command.
2. Alternatively wait until the command finishes executing and check the `churn_dbexport.log` located in the main `/opt/tnf/apps/telco-dbexport/scripts` folder.

Verify that there are no errors.

Note: The export may fail due to the SPSS Job not being triggered. The SPSS Job will only trigger successfully when the Solutions SPSS models are installed and configured.

3. Log on to the PCI DB2 node with a user ID that has access to the IBM DB2® database. For example: `db2inst1`.
4. Connect to the database:

```
db2 connect to TELCO;
```

5. Select from one of the Churn tables to ensure that one table populated with data. For example run the command:

```
Select count(*) from BBCI.cgr_device;
```

3.8.6 Manually running the NPS Database Loader job

The first time the Database Loader is run for NPS it should be run manually. The main reason is to determine that the end to end solution is running correctly and that the model is being executed successfully. NPS should not require manual intervention to train the model.

To run the NPS Database Loader job.

1. Navigate to the main Database Loader folder
`cd /opt/tnf/apps/telco-dbexport/scripts/`
2. Run `./runDBExport.sh nps`

3.8.7 Checking the NPS Database Loader job

Validate the job has run successfully.

1. Monitor the output of the `runDBExport.sh` command at `/opt/tnf/apps/telco-dbexport/log`.
2. Alternatively wait until the command finishes executing and check the `nps_dbexport.log` located in the main `/opt/tnf/apps/telco-dbexport/scripts/` folder.

Verify that there are no errors.

Note: The export may fail due to the SPSS Job not being triggered. The SPSS Job will only trigger successfully when the Solutions SPSS models are installed and configured.

3. Log on to the PCI DB2 node (`pcidbnode`) with a user ID that has access to the IBM DB2® database. For example, `db2inst1`.
4. Connect to the database:

```
db2 connect to TELCO;
```

5. Select from one of the NPS tables to ensure that a single table is populated with data. For example run:

```
Select count(*) from NPS.staging_nps_score_table;
```

3.8.8 Configuring the Database Loader cron job

To configure the Database Loader cron job:

1. Log on to the Hadoop node (`aafnode`).
2. Switch to the boss user by running the command: `su boss`
3. List the cron tabs by running the command: `crontab -l`

Cron job execution schedules are in the following format:

```
* * * * *
```

[Minute] [Hour] [Day of the Month] [Month of the Year]
 [Day of the Week]

Where:

- Minute ranges from 0 to 59
 - Hour ranges from 0-23
 - Day of Month ranges from 1-31
 - Month of the Year ranges from 1-12 or JAN-DEC
 - Day of the Week ranges from 1-7 where 1 stands for Monday or SUN-SAT
4. Record the execution time for the Churn (churn-dataset) and NPS (net-promoter-score) datasets.
 5. Run the tab `crontab -e`

Tip: Take care not to change or remove existing cron jobs.
 6. Type `i` to insert into the crontab file.

A sample configuration for a Churn cron job that runs every Monday at 3am and logs the cron job output to `churn_db_export_cron.log` is as follows:

```
* 3 * * MON cd /opt/tnf/apps/telco-dbexport/scripts/ &&
./runDBExportForCron.sh churn > /tmp/churn_db_export_cron.log
2>&1
```

A sample configuration for a NPS cron job that runs every Monday at 3am and logs the cron job output to `nps_db_export_cron.log` is as follows:

```
* 3 * * MON cd /opt/tnf/apps/telco-dbexport/scripts/ &&
./runDBExportForCron.sh nps > /tmp/nps_db_export_cron.log
2>&1
```

3.8.9 Verifying that the Database Loader cron job is set up correctly

To verify that the cron job is set up correctly:

1. Log on to the Hadoop node (aafnode) after the time that the cron job is configured to run.
2. Ensure that the Churn and NPS export logs are created in the `/opt/tnf/apps/telco-dbexport/log` directory and that there are no errors in the log.

In the sample provided, the log for churn is `churn_db_export_cron.log` and it is located in the `tmp` directory on the Hadoop node.

3.9 Configuring SPSS components

To enable SPSS Modeler Client and SPSS Modeler Server to work with the SPSS Analytic Server and the DB2 Node, some configuration updates are required.

3.9.1 Configuring the SPSS Modeler Server connection to the Analytic Server

1. Open the SPSS Modeler Server `options.cfg` file at the following location:

`/usr/IBM/SPSS/ModelerServer/17.1/config/options.cfg`

Note: SPSS Modeler Server is installed on the PCI/SPSS node (pcipanode) and information on installing is described in section 3.1.1.

2. Update the Analytic Server settings by adding the following two lines.

`as_url, http://{AS_SERVER}:{PORT}/admin/{TENANT}`

`as_prompt_for_password, {Y|N}`

- AS_SERVER - The IP address of the Analytic Server.
- PORT - The Analytic Server port number.
- admin / (TENANT) The tenant that the SPSS Modeler Server installation is a member of.
- Prompt for Password (as_prompt_for_password) Specify N if the SPSS Modeler Server is configured with the same authentication system for users and passwords as the system that is used on Analytic Server; for example, when you use Kerberos authentication, otherwise, Y.

3. Restart the Modeler Server by running the following commands:

```
cd /usr/IBM/SPSS/ModelerServer/17.1
```

```
./modelersrv.sh stop
```

```
./modelersrv.sh start
```

Validate the connection

To check the connection, you must complete the next step. SPSS Modeler Client will not connect to the Analytic Server if the `options.cfg` file is not set up correctly.

3.9.2 Configuring SPSS Modeler Client connection to the Analytic Server

1. Open the SPSS Modeler Client.

Note: Installation instructions for SPSS Modeler Client are referenced in Section 3.1.1.

2. Select **Tools > Server Login**

3. Enter server login details as specified in the SPSS Modeler Server `options.cfg` file at the following location:

`/usr/IBM/SPSS/ModelerServer/17.1/config/options.cfg`

Validate the connection

To check the connection, ensure the `options.cfg` file is set up correctly.

To further validate the connection, open SPSS Modeler Client and complete the following steps to test an Analytic Server input datasource in an SPSS stream is available for selection.

1. Open the SPSS Modeler Client.

2. Select **File > New Stream**
 If prompted, connect to the SPSS Modeler Server.
 If you are not prompted after creating the new stream, click the **Server** button on the bottom left corner of the Modeler Client user interface and create the connection to Modeler Server.
3. Select the **Sources** tab, and drag an Analytic Server node onto the white blank stream.
4. Right-click the stream and select **Edit**.
5. Click **Select** beside the Datasource field.
6. If prompted, enter your Analytic Server login details.

The Analytic Server data sources should be available for selection.

3.9.3 Configuring the SPSS Modeler Server connection to DB2

Complete the following steps to ensure the Churn and NPS SPSS Models run on DB2.

1. Stop the IBM SPSS Modeler Server. Go to `/usr/IBM/SPSS/ModelerServer/17.1` and at the UNIX command prompt type:

```
./modelersrv.sh stop
```

Note: SPSS Modeler Server is installed on the PCI/SPSS node (pcipanode) and information on installing is described in section 3.1.1.

2. Navigate to the folder `/root/SDAP711`
3. Run the `setodbcpath.sh` script to update the ODBC path in the scripts.
4. Edit the `odbc.sh` script to add the definition for ODBCINI to the bottom of the script. For example:

```
ODBCINI=/root/SDAP711/odbc.ini; export ODBCINI
```

ODBCINI must point to the full file path of the `odbc.ini` file for IBM SPSS Modeler. The `odbc.ini` file lists the ODBC data sources that you want to connect to. A default `odbc.ini` file is installed with the drivers.

5. Update the `odbc.ini` file, add the data source and specify the driver in the [ODBC Data Sources] section as follows:
`TELCO=IBM DB2 ODBC Driver`
6. In the `odbc.ini` file, create a Telco data source connection.

```
[TELCO]
Driver=/opt/ibm/db2/V10.5/lib64/libdb2o.so
DriverUnicodeType=1
Description=IBM DB2 ODBC Driver
ApplicationUsingThreads=1
AuthenticationMethod=0
BulkBinaryThreshold=32
BulkCharacterThreshold=-1
BulkLoadBatchSize=1024
CharsetFor65535=0 #Database applies to DB2 UDB only
```

```

Database=TELCO
DefaultIsolationLevel=1
DynamicSections=200
EnableBulkLoad=0
EncryptionMethod=0
FailoverGranularity=0
FailoverMode=0
FailoverPreconnect=0
GrantAuthid=PUBLIC
GrantExecute=1
GSSClient=native
HostNameInCertificate=
IpAddress=IP_Address_of_DB_server
KeyPassword=
KeyStore=
KeyStorePassword=
LoadBalanceTimeout=0
LoadBalancing=0
LogonID=db2inst1
MaxPoolSize=100
MinPoolSize=0
Password=password
PackageCollection=NULLID
PackageNamePrefix=DD
PackageOwner=
Pooling=0
ProgramID=
QueryTimeout=0
ReportCodePageConversionErrors=0
TcpPort=50000
TrustStore=
TrustStorePassword=
UseCurrentSchema=0
ValidateServerCertificate=1
WithHold=1
XMLDescribeType=-10

```

Note: You must use the driver library `libdb2o.so` with IBM SPSS Modeler. Ensure that you set `DriverUnicodeType=1` to avoid buffer overflow errors when you connect to the database.

7. If you are using the 64-bit version of IBM SPSS Modeler Server, define and export `LD_LIBRARY_PATH_64` in the `odbc.sh` script:

```

if [ "$LD_LIBRARY_PATH_64" = "" ]; then
LD_LIBRARY_PATH_64=<library_path> else
LD_LIBRARY_PATH_64=<library_path>:$LD_LIBRARY_PATH_64 fi
export LD_LIBRARY_PATH_64

```

Where `<library_path>` is the same as for the `LD_LIBRARY_PATH` definition in the script that was initialized with the installation path. For example, `/opt/spss/odbc/lib`.

Tip: Copy the if and export statements for LD_LIBRARY_PATH in the odbc.sh file, and append them to the end of the file. Then, replace the LD_LIBRARY_PATH strings in the newly appended if and export statements with LD_LIBRARY_PATH_64.

Here is an example of the odbc.sh file for a 64-bit IBM SPSS Modeler Server installation:

```
if [ "$LD_LIBRARY_PATH" = "" ];
then
  LD_LIBRARY_PATH=/opt/spss/odbc/lib
else
  LD_LIBRARY_PATH=/opt/spss/odbc/lib:$LD_LIBRARY_PATH
fi
export LD_LIBRARY_PATH
if [ "$LD_LIBRARY_PATH_64" = "" ];
then
  LD_LIBRARY_PATH_64=/opt/spss/odbc/lib
else
  LD_LIBRARY_PATH_64=/opt/spss/odbc/lib:$LD_LIBRARY_PATH_64
fi
export LD_LIBRARY_PATH_64
ODBCINI=/opt/spss/odbc/odbc.ini; export ODBCINI
```

Ensure that you export LD_LIBRARY_PATH_64, and define it with the if loop.

Here is an example with the <library_path> variable specified.

```
if [ "$LD_LIBRARY_PATH_64" = "" ];
then
  LD_LIBRARY_PATH_64=/root/SDAP711/lib
else
  LD_LIBRARY_PATH_64=/root/SDAP711/lib:$LD_LIBRARY_PATH_64
fi
export LD_LIBRARY_PATH_64
```

8. Configure IBM SPSS Modeler Server to use the driver.

- (a) Go to /usr/IBM/SPSS/ModelerServer/17.1 and edit modelersrv.sh. Add the following line immediately below the line that defines SCLEMDNAME:

```
. <odbc.sh_path>
```

Where odbc.sh_path is the full path to the odbc.sh file.

For example: . /opt/spss/odbc/odbc.sh

Ensure that you leave a space between the first period and the file path.

- (b) Save modelersrv.sh

9. Configure the IBM SPSS Modeler Server to use the ODBC wrapper named libspssodbc_datadirect.so.

- (a) Go to the /usr/IBM/SPSS/ModelerServer/17.1/bin directory.

- (b) Remove the existing libspssodbc.so soft link by using the following command:

```
rm -fr libspssodbc.so
```

(c) Link the new wrapper to `libspssodbc.so` by using the following command:

```
ln -s libspssodbc_datadirect_utf16.so libspssodbc.so
```

10. Copy `db2cli.ini.sample` from `/opt/ibm/db2/V10.5/cfg` to `/home/db2inst1/sqllib/cfg` and rename `db2cli.ini`.

Configure the `db2cli.ini` file to add the sections for the Telco database:

```
[TELCO] Database=TELCO
Protocol=TCPIP
DriverUnicodeType=1
Port=50000
Hostname=IP_Address_of_DB_server
UID=<username>
PWD=<password>
```

11. Restart the `modelersrv` when the steps are completed.

Go to `/usr/IBM/SPSS/ModelerServer/17.1`

Start the modeller by running the following command:

```
./modelersrv.sh start
```

Validate the connection

To check the connection, you must complete the next step.

3.9.4 Configuring the SPSS Modeler Client connection to DB2

1. Open the SPSS Modeler Client.
2. Select **File > New Stream**
If prompted, connect to the SPSS Modeler Server.
If you are not prompted after creating the new stream, click the **Server** button on the bottom left corner of the Modeler Client user interface and create the connection to Modeler Server.
3. Select the **Sources** tab, and drag a Database node onto the white blank stream.
4. Right-click the stream and select **Edit**.
5. Click **Select** beside the Datasource field.
6. Select the <Add new database connection> option.
7. Click **Refresh**.
8. Ensure the Telco database is listed.

Validate the connection

To further validate the connection, complete earlier validation steps and the following steps.

1. Select the Telco Database.
2. Enter the db2 authentication details.

3. Select **Connect**. Ensure the connection succeeds.
4. Click OK.
5. In the Database node window which should remain open, click **Select** beside the Table Name.
6. Select a TNF table. For example, select table: CP_APPLICATION_MAPPING.
7. Click OK.
8. Right-click the database node and select Preview.

If there is data in the table you selected and the data is returned then the connection between Modeler Client and DB2 is successful.

3.10 Deploying the SPSS Modeler

3.10.1 Before you begin

Ensure that you have:

- Deployed the solution datasets as described in Section 3.4.
- Configured and run the Sqoop job to extract data from Hadoop to DB2 as described in Section 3.8.

3.10.2 Deploying the Analytic Server datasources

Analytic Server datasources are required by the customer profiles models in order to enable all the models to read and export data on the Hadoop node (aafnode).

Installing the Analytic Server datasources must be completed on the main Hadoop node (aafnode) where Analytic Server is installed.

1. Launch the IBM Analytic Server at `http://host:port/analyticserver/admin/ibm`.
Note the default port on a single node system is 9080.
2. Log in using your Analytic Server credentials.
3. Click **Datasources**.
4. Import each of the zip files located at the:
`IS_CSP_Customer_Insight_1.0.4/ analytics-platform/analyticserver_datasources/ directory`.
5. Select **Actions > Import > Browse** and select a data source.
6. Repeat steps 4-5 to import each of the data sources in the `analyticserver_datasources` directory.

Validate the installation

1. Launch the IBM Analytic Server at `http://host:port/analyticserver/admin/ibm`.
Note the default port on a single node system is 9080.
2. Log in using your Analytic Server credentials.

3. Click **Datasources**.
4. Select a Datasource and Preview the Content. Previewing the content of the datasource ensures that the connection to the Hadoop has succeeded.

3.10.3 Deploying models in SPSS Collaboration and Deployment Services

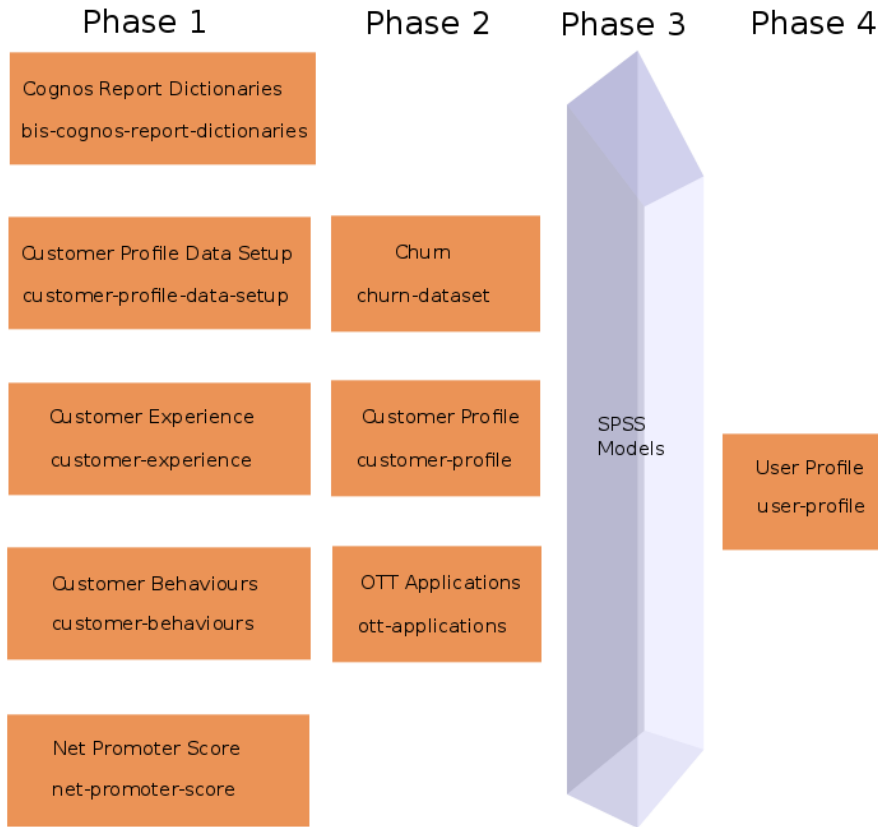
The SPSS models are packaged in a .pes file within `spss.zip` that must be imported into the Collaboration and Deployment Services client on the Windows node on the deployment machines.

1. Start Deployment Manager.
2. Click **File > New > Content Server Connection**.
3. In the Connection Name field, enter a name that identifies the Predictive Analytics node.
4. In the Server URL field, enter `http://analytics_node_IP:9081` and click **Finish**.
5. Right-click Content Repository and click Import.
6. Browse to and select the `CSP_CustomerInsights_CDS.pes` file.
7. Click Open.
8. Accept the default options in the Import window, and click OK. The `CSP_CustomerInsights_NPS` folder, a job in the Jobs folder, and streams in the Modeler Streams folder are created.
9. Repeat steps 6-8 for the `CSP_CustomerInsights_Profile.pes` file.
10. Update the credential information for the admin, db2inst1, and root users so that the job and streams run successfully.
 - In the Content Explorer tab, under Resource Definitions, open the Credentials folder.
 - Update the admin user with the credentials for the SPSS Collaboration and Deployment Services user that has access to the content repository and runs the job.
 - Update the db2inst1 user with the credentials for the user that has access to the IBM DB2 database. Update the root user with the credentials for the user that has access to the Modeler Server.
11. Open the job in the Jobs folder and ensure that the user credentials match the credentials of the IBM user: db2inst1.
12. Verify the server connections are correct for your environment. In the Content Explorer, open the Servers folder and verify the collaboration and deployment services and modeler connections.

3.10.4 Scheduling SPSS Job Triggers

A series of data processing activities is required to perform data analysis in the Customer Insight for CSP solution. Figure 7 SPSS Dataset and Model Activity Flow Sample shows a sample activity flow.

Figure 7 SPSS Dataset and Model Activity Flow Sample



Note: Figure 7 displays the order of execution but does not display exact dependencies. ETL for example is not included. The Churn and NPS models are triggered by the Database Loader. Customer Profile SPSS jobs are triggered from the `customer-profile-data-setup` dataset.

Revising Customer Profile SPSS job schedules

The `customer-profile-data-setup` dataset is scheduled to run at 00.30 on a Monday, two hours before the next dataset is run.

If a dataset's models are processing a large amount of data, and data processing does not complete before subsequent and dependent datasets are triggered, the dependent datasets will run with old data. Check the hive log files on the Hadoop node (aafnode) at `/tmp/boss/hive.log`.

It may be necessary to adjust the schedules of those datasets to ensure they are not being triggered before all necessary processing completes. To do so, edit the dataset run-schedules:

1. Log on to the Hadoop Master node as `boss` user.
2. Run the command:

```
crontab -e
```

3. Press `a` while in `crontab` to make the entries editable.
4. Modify either of the first two fields (minute / hour), save updates and exit.

The `customer-profile-data-setup` pre-processes the input data to the SPSS models before triggering the models.

The Customer Profile SPSS models generate output from the tables populated by the `customer-profile-data-setup` dataset.

When the dataset is run, the models that depend on this dataset are triggered automatically through the SPSS Collaboration and Deployment Services Client (C&DS).

In order for the datasets to be automatically triggered configuration files are required to be updated in order to allow the dataset to trigger the model. The configuration file is located in the `/opt/tnf/apps/customer-profile-data-setup/config` folder on the Hadoop master node (`aafnode`). The server connection properties are found in the `spss_server.properties` file.

After setting the username and password properties, the `encodePasswordProperties.sh` script (also located in the `config` directory) should be run on this file in order to base64 encode credentials (see Base encoding the passwords in the configuration files).

Model trigger information is located in the `.job_properties` files. Each `.job_properties` file corresponds to a single job to be run through C&DS.

The properties to configure are as follows:

Property	Description
SPSSID	The C&DS job ID, located in the C&DS dashboard. To locate the SPSS ID, select the job in the Content Repository and then select Properties from the menu.
SPSSTriggerModel	A boolean value. Can be set to false in order to quickly disable the automatic execution of this job (if required).
SPSSRunSchedule	A comma-separated list of the days on which this model should be run, using the three-letter code for a day (such as <code>mon, tue, thu</code>). Note: A model is only triggered after <code>customer-profile-data-setup</code> is run. A model will trigger only if set to run on the same day as <code>customer-profile-data-setup</code> .
PrerequisiteJobs	(Optional) Enables you to specify job dependencies. Some models rely on the output of other models, so those models must not be triggered until all the models on which they depend have been triggered and completed. The value is a comma-separated string of job names. For example, the <code>customer_profile_best_time_and_medium</code> models depend on output from the

	<p>customer_profile_location_affinity model, which is also triggered by customer-profile-data-setup.</p> <p>In this case, customer_profile_best_time_and_medium.job_properties specify a PrerequisiteJobs value of customer_profile_location_affinity to ensure that it waits for customer_profile_location_affinity to complete before running.</p> <p>Note: Jobs that do not specify any prerequisites run as normal, without dependencies.</p>
--	--

3.10.5 Validating the installation

In order to verify the installation of the models the following verification checks should be completed.

Verify a SPSS Model that reads from DB2

1. Open Collaboration and deployment Services on the Windows node.
2. Navigate to **Content Repository > CSP_CustomerInsight_NPS > Jobs**.
3. Right click the Training Churn Prediction Model.
4. Select Run Job.
5. In a few minutes right click the Job and select Show Job History.
6. Monitor the logs to ensure there are no error.

Note: At this point in the installation the models have not been trained. If the job fails ensure it is not failing for connectivity reasons. If there are no connectivity issues then the SPSS connections have been configured correctly and the SPSS models have been installed correctly.

Verify a SPSS Model that reads from Analytic Server

1. Open Collaboration and deployment Services on the Windows node.
2. Navigate to **Content Repository > CSP_CustomerInsight_Profile > Jobs**.
3. Right click the Customer Profile Lifestyle Mobility Job.
4. Select Run Job.
5. In a few minutes right click the Job and select Show Job History.
6. Monitor the logs to ensure there are no error.

Note: At this point in the installation the models have not been trained. If the job fails ensure it is not failing for connectivity reasons. If there are no connectivity issues then the SPSS connections have been configured correctly and the SPSS models have been installed correctly.

3.11 Deploying the visualization component

Complete the following steps to install the Cognos reports and dashboards.

3.11.1 Before you begin

Verifying the tables or views required by the dashboards are visible using a BigSQL connection

On the Hadoop node (aafnode), connect to bigsql using the db2 command line, as shown in the following commands:

1. Switch to the bigsql user:

```
su bigsql
```

2. Connect to bigsql:

```
db2 connect to bigsql
```

3. List the tables synchronized:

```
db2 list tables for schema tnf
```

The output should match the following:

Table/View	Schema	Type	Creation time
CB_CATEGORY	TNF	T	2016-08-26-15.31.11.531522
CGR_APPLICATION	TNF	T	2016-08-26-15.31.47.956727
CGR_APPLICATIONS_CONFIG	TNF	T	2016-08-26-15.31.48.126860
CGR_CRM	TNF	T	2016-08-26-15.31.48.299948
CGR_DEVICE	TNF	T	2016-08-26-15.31.48.508268
CGR_PERF_ASSIGN_CONFIG	TNF	T	2016-08-26-15.31.48.704083
CGR_UNACCEPTABLE_TREND_CONFIG	TNF	T	2016-08-26-15.31.48.867846
CONNECTION_EXPERIENCE_PROFILE_>	TNF	T	2016-08-26-15.31.18.645216
CP_APPLICATION_USAGE_W	TNF	T	2016-08-26-15.31.12.101273
CP_APPLICATION_USAGE_W_LATEST	TNF	T	2016-08-26-15.31.17.858681
CUSTOMER_BEHAVIOUR_W	TNF	T	2016-08-26-15.31.11.338262
GN_CP_2_H_1	TNF	T	2016-08-26-15.31.49.027877
GN_CP_ST_2_H_1	TNF	T	2016-08-26-15.31.49.221128
MOBILE_UP_3_H_1	TNF	T	2016-08-26-15.31.49.430280
OTT_APPLICATION_AGGREGATION	TNF	T	2016-08-26-15.30.54.207120
OTT_APPLICATION_MUSIC	TNF	T	2016-08-26-15.30.54.285733
OTT_APPLICATION_UNION	TNF	V	2016-08-26-15.31.08.043951
OTT_APPLICATION_VIDEO	TNF	T	2016-08-26-15.30.54.353019
OTT_APPLICATION_VOICE_AND_MESS>	TNF	T	2016-08-26-15.30.54.428011
OTT_APPLICATION_W	TNF	T	2016-08-26-15.30.54.494038
OTT_APP_DAILY_TOP_VOLUMES	TNF	T	2016-08-26-15.30.53.968564
OTT_APP_TOP_CUSTOMERS	TNF	T	2016-08-26-15.30.54.049108
OTT_APP_TOP_VOLUMES	TNF	T	2016-08-26-15.30.54.130465
OTT_DEVICE_DAILY_TOP_VOLUMES	TNF	T	2016-08-26-15.30.54.572865
OTT_DEVICE_TOP_CUSTOMERS	TNF	T	2016-08-26-15.30.54.646994
OTT_DEVICE_TOP_VOLUMES	TNF	T	2016-08-26-15.30.54.713263
OTT_PERCENTAGE_CHANGE_W	TNF	T	2016-08-26-15.30.54.779990
OTT_TOP_N_APP_PERCENTAGE_CUSTO>	TNF	T	2016-08-26-15.30.54.853759
OTT_TOP_N_APP_PERCENTAGE_VOLUM>	TNF	T	2016-08-26-15.30.54.923037
OTT_TOP_N_DEVICES_APP_AGG	TNF	T	2016-08-26-15.30.55.176676
OTT_TOP_N_DEVICE_PERCENTAGE_CU>	TNF	T	2016-08-26-15.30.55.039077
OTT_TOP_N_DEVICE_PERCENTAGE_VO>	TNF	T	2016-08-26-15.30.55.109154
OTT_TOP_N_PRICE_PLANS_APP_AGG	TNF	T	2016-08-26-15.30.55.276894
OVERALL_EXPERIENCE_PROFILE_J	TNF	T	2016-08-26-15.31.18.426044
PERFORMANCE_ASSIGN_CONFIG	TNF	T	2016-08-26-15.31.11.152388
SMS_EXPERIENCE_PROFILE_J	TNF	T	2016-08-26-15.31.19.005402
SUBSCRIBER_DOMAIN	TNF	T	2016-08-26-15.31.18.049222
SUBSCRIBER_INTEREST	TNF	T	2016-08-26-15.31.18.237207
SUBSCRIBER_LEVEL_CX_SCORE_DAIL>	TNF	T	2016-08-26-15.31.49.925678
SUBSCRIBER_LEVEL_CX_SCORE_WEEK>	TNF	T	2016-08-26-15.31.50.129472
SUBSCRIBER_PROFILE_W	TNF	T	2016-08-26-15.31.11.911526
SUBSCRIBER_PROFILE_W_LATEST	TNF	T	2016-08-26-15.31.17.675560
USER_PROFILE_W	TNF	T	2016-08-26-15.31.11.719573
USER_PROFILE_W_LATEST	TNF	T	2016-08-26-15.31.17.453013
VIDEO_EXPERIENCE_PROFILE_J	TNF	T	2016-08-26-15.31.19.402207
VOICE_EXPERIENCE_PROFILE_J	TNF	T	2016-08-26-15.31.18.832182
VOICE_SMS_2_H_1	TNF	T	2016-08-26-15.31.50.324686
WEB_BROWSING_EXPERIENCE_PROFIL>	TNF	T	2016-08-26-15.31.19.191099

48 record(s) selected.

If tables or views are missing, then complete the following steps:

1. Switch to the bigsql user:

```
su bigsql
```

2. Connect to bigsql:

```
db2 connect to bigsql
```

3. Ensure the table or view exists in hive prior to running the command. Run the command for each table or view missing:

```
db2 "CALL SYSHADOOP.HCAT_SYNC_OBJECTS ('TNF',
'<TABLE_OR_VIEW_NAME>', '<t or v>', 'REPLACE', 'CONTINUE')"
```

4. Replace <TABLE_OR_VIEW_NAME> with the required table or view name to be synchronized to bigsql and update <t or v> with either t or v depending on if the synchronization is being run for a table or view.

Setting the report server execution mode

To set the report server execution mode:

1. Open a command line terminal and log in as root user on to the PCI/Cognos node (pcibinode).
2. Go to the following directory:

```
cd /opt/ibm/cognos/analytics/bin64
```
3. Start the Cognos configuration wizard:

```
sh cogconfig.sh
```

Note: This launches a graphical configuration tool and requires a direct connection to the server or an X-windows capability.
4. In the wizard, navigate to:
Local Configuration -> Environment -> Report Server execution mode
5. Change the mode from 32bit to 64bit and save the configuration.
6. Restart Cognos using the wizard or terminal:

```
/opt/ibm/cognos/analytics/bin64/cogconfig.sh -stop
```

```
/opt/ibm/cognos/analytics/bin64/cogconfig.sh -s
```
7. Close the wizard.

Setting up the folder structure in Cognos

To set up the folder structure in Cognos:

1. As root user, in your Firefox or Chrome browser open the IBM Cognos Analytics page:

```
<hostname where Cognos is running>:9300/bi/
```
2. Select **Team content** from the left hand menu.
3. Click the **arrow** icon to change the view.
4. Click the **new folder** icon to create a new folder.
5. Create the following folder structure:
Team content -> CI -> reports
6. Create a similar folder structure for dashboards. A dashboard represents an assembled view that contains visualizations such as a graph, chart, plot, table, map, or any other visual representation of data.
Team content -> CI-> dashboards
Team content -> CI-> dashboards-mobile

3.11.2 Deploying the reports

Complete the following steps to deploy reports.

Importing CI reports and images

1. Open a command line terminal and log in as root user on to the PCI/Cognos node (pcibinode).

2. Go to the following directory:

```
cd /opt/IBM/IS_CSP_Customer_Insight_1.0.4/
```

3. Copy the report zips to the Cognos deployment directory:

```
cp CI-churn-report-1.0.4.zip  
/opt/ibm/cognos/analytics/deployment/
```

```
cp CI-nps-report-1.0.4.zip  
/opt/ibm/cognos/analytics/deployment/
```

4. Copy the images zip to the correct Cognos directory:

```
cp CI-nps-images-1.0.4.zip  
/opt/ibm/cognos/analytics/webcontent/bi/samples/images/
```

5. Unzip all the images into the images folder, so that the path to an image is as follows:

```
/opt/ibm/cognos/analytics/webcontent/bi/samples/images  
/<image>.png
```

Installing the reports

1. As root user, in your Firefox or Chrome browser open the IBM Cognos Analytics page:

```
<hostname where Cognos is running>:9300/bi/
```

2. Select **Manage->Administration Console** to open the IBM Cognos Administration Console in a new tab.

3. In the IBM Cognos Administration Console, select **Configuration -> Content Administration** to open a view on the content imported and exported to and from Cognos.

4. Select the **New Import** symbol from the symbol menu on the right side of the page to launch the **New Import** wizard.

5. Select "<report-name>-1.0.4" and click **Next**.

6. On the "Specify a name and description" page, accept the defaults and click **Next**.

7. On the "Select the public folders, directory and library content" select all the folders/items which appear and click **Next**.

8. On the "Specify the general options" page accept the defaults and click **Next**.

9. On the "Review the summary" page accept the defaults and click **Next**.

10. On the "Select an action" page accept the defaults and click **Finish**.

11. On the "Run with options" page accept the defaults and click **Run**.

12. On the final wizard page click "OK".

13. The import should now be visible on the "Content Administration" page.

14. Switch back to the IBM Cognos Analytics page, select "Manage" and then select "Data servers".
Note: If the required datasources (correct type and name) exist on the system, it may only be required to edit the username and password in steps 16, 17 and 18.
15. In the "Data servers" menu click the "plus" symbol to create a new data server.
16. For CI reports do the following:
 - a) In the "Select a type" menu select "DB2"
 - b) In the "Connection" menu please specify the connection name <TELCO>, server, port<50000>, database name <TELCO>
 - c) Select "Use the following saved credentials" and enter the username <db2inst1>, password <*****>
 - d) Test the connection and if its successful click "OK" to save the data source connection
 - e) Select "Team content" -> "CI" -> "reports" and navigate to the desired report to open it
17. Once the correct data has been loaded in the database, each page of the reports will display correctly.
18. Cognos logs activity to the following logs:

```
/opt/ibm/cognos/analytics/logs/cogserver.log
```

```
/opt/ibm/cognos/analytics/logs/p2pd_messages.log
```

3.11.3 Deploying the dashboards

Importing the CI dashboards

1. Open a command line terminal and log in as root user on to the PCI/Cognos node (pcibinode).
2. Run this command to go to the correct directory:

```
cd /opt/IBM/IS_CSP_Customer_Insight_1.0.4/
```
3. Run these commands to copy the report zips to the Cognos deployment directory:

```
cp CI-dashboards-1.0.4.zip  
/opt/ibm/cognos/analytics/deployment/
```

Installing the CI dashboards

1. As root user in your Firefox or Chrome browser navigate to the IBM Cognos Analytics page:

```
<hostname where Cognos is running>:9300/bi/
```
2. From the menu select **Manage -> Administration Console** to open the IBM Cognos Administration Console in a new tab.

3. In the IBM Cognos Administration Console, select **Configuration** -> **Content Administration** to open a view on the content imported and exported to and from Cognos.
4. Select the "New Import" symbol from the symbol menu on the RHS of the page - to launch the "New Import wizard"
5. Select "CI-dashboards-1.0.4" and click **Next**.
6. On the "Specify a name and description" page accept the defaults and click **Next**.
7. On the "Select the public folders, directory and library content" select all the folders/items which appear and click **Next**.
8. On the "Select the directory content" page accept the defaults and click **Next**.
9. On the "Specify the general options" page accept the defaults and click **Next**.
10. On the "Review the summary" page accept the defaults and **Finish**.
11. On the "Select an action" page accept the defaults and click **Next**.
12. On the "Run with options" page accept the defaults and click **Run**.
13. On the final wizard page click **OK**.
14. The import should now be visible on the "Content Administration" page.

Note: If the "bbci-ott-datasource" or "bbci-user-profile-datasource" do not already exist on your system (pre-dashboard installation), do steps:16, 17 and 18 - otherwise the existing data source login will be okay.
15. Switch back to the IBM Cognos Analytics page, select "Manage" & then select "Data servers"
16. For the ott-dashboard do the following:
 - a) Select the "bbci-ott-datasource" and select it again in the next tab that opens on the page
 - b) Select "Connection details" from the 3-dot menu
 - c) Edit the connection: select "Use the following saved credentials" and enter the username, password, host & port
 - d) Test the connection and if successful click "OK" to save the data source connection
 - e) Select "Team content" -> "CI" -> "dashboards" -> "ott-dashboard" to open the dashboard
 - f) If using an iPad select "Team content" -> "CI" -> "dashboards-mobile" -> "ott-dashboard-mobile"
17. For the user-profile-dashboard do the following:
 - a) Select the "bbci-user-profile-datasource" and select it again in the next tab that opens on the page
 - b) Select "Connection details" from the 3-dot menu
 - c) Edit the connection: select "Use the following saved credentials" and enter the username, password, host & port

- d) Test the connection and if successful click "OK" to save the data source connection
 - e) Select "Team content" -> "CI" -> "dashboards" -> "user-profile-dashboard" to open the dashboard
 - f) If using an iPad select "Team content" -> "CI" -> "dashboards-mobile" -> "user-profile-dashboard-mobile"
18. Once the correct data has been loaded in the database, each tab of the dashboards will display correctly
19. Cognos logs activity to the following files:
- ```
/opt/ibm/cognos/analytics/logs/cogserver.log
/opt/ibm/cognos/analytics/logs/p2pd_messages.log
```

### 3.11.4 Verifying deployment of reports and dashboards

Verify deployment of reports and dashboards.

1. Start Cognos:

```
cd /opt/ibm/cognos/analytics/bin64/
export JAVA_HOME=/opt/ibm/cognos/analytics/jre
./cogconfig.sh -s
```
2. Open Cognos Analytics in your browser:

```
http://<hostname>:9300/bi
```
3. Select Team content -> CI -> dashboards to view CI dashboards.
4. Select Team content -> CI -> dashboards-mobile to view CI mobile dashboards.
5. Select Team content -> CI -> reports to view CI reports.

To stop Cognos:

```
./cogconfig.sh -stop
```

To check the Cognos log files go to `/opt/ibm/cognos/analytics/logs` to view:

```
cogconfig_response.csv.*.log
cogserver.log
```



## 4 Configuring the Customer Insight for CSP solution

### 4.1 Managing resources assigned to dataset processing

To manage the resources assigned to the dataset execution, configure the Yarn queues to assign specific resources. Complete the steps in this section to set up Yarn queues, specify queries to run on specific Yarn queues, and test queries are run on the correct Yarn queue.

#### 4.1.1 Configuring the Yarn queues

Use the Ambari dashboard to configure Yarn queues.

1. Open a browser and access the Ambari server dashboard.  
`http://<server-name>:8080`
2. Select the button in the Ambari menu bar with the 6 boxes.
3. Select Yarn Queue manager from the drop down.
4. Select Add Queue.
5. Enter the name of the new queue usecases.
6. Populate the resource percentages. Updating the usecases queue will require updates to the other configured queues.
7. Ensure hive query is set to map reduce/tez depending on the systems configuration.

#### 4.1.2 Updating solution datasets to run on a specific queue

To update the dataset configuration to run on a specific queue, modify the properties file in the `/opt/tnf/apps/dataset-common/master-config` folder.

For Tez:

```
set hive.execution.engine=tez;
set yarn.queue.name=usecases;
```

For Map Reduce:

```
Set hive.execution.engine=mr;
set yarn.queue.name=usecases;
```

#### 4.1.3 Queue management summary commands

|                                          |                                                        |
|------------------------------------------|--------------------------------------------------------|
| List the queues configured in a terminal | <code>&gt;hadoop queues -list</code>                   |
| Show jobs in the queue                   | <code>&gt;hadoop queue -info usecases -showJobs</code> |

### 4.1.4 Queue file locations

|                                                      |                                                                                                                                                                                                                                                               |
|------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Hadoop scheduler configuration file (stored locally) | /etc/hadoop/4.1.0.0/capacity-scheduler.xml                                                                                                                                                                                                                    |
| Map reduce query logs                                | http://<hostname>:8088/logs/                                                                                                                                                                                                                                  |
| Job queries (for all queries)                        | <p>http://&lt;hostname&gt;: 8088/cluster/apps</p> <p>The URL shows all jobs (both Map Reduce and Tez) or alternatively 19888/jobhistory for just map reduce jobs.</p> <p>The cluster apps shows the queue and application type (either Map Reduce or Tez)</p> |

### 4.1.5 Verifying a query has run on the correct queue

After running a simple query on the map reduce configured environment, the query goes to the usecases queue. Verify that the query is sent to the correct queue by checking the All applications URL

http://<hostname>:8088/cluster/apps.

Sending datasets queries to a specific queue enables greater control over system use.

Viewing all jobs history



### All Applications

Logged in as: dr.who

- Cluster
- About Nodes
- Node Labels
- Applications
- NEW
- NEW SAVING
- SUBMITTED
- ACCEPTED
- RUNNING
- FINISHED
- FAILED
- KILLED
- Scheduler
- Tools

| Cluster Metrics |              |              |                |                    |             |              |                 |             |              |                 |              |                      |            |                 |                |
|-----------------|--------------|--------------|----------------|--------------------|-------------|--------------|-----------------|-------------|--------------|-----------------|--------------|----------------------|------------|-----------------|----------------|
| Apps Submitted  | Apps Pending | Apps Running | Apps Completed | Containers Running | Memory Used | Memory Total | Memory Reserved | VCores Used | VCores Total | VCores Reserved | Active Nodes | Decommissioned Nodes | Lost Nodes | Unhealthy Nodes | Rebooted Nodes |
| 14              | 0            | 1            | 13             | 1                  | 2 GB        | 27.34 GB     | 0 B             | 1           | 8            | 0               | 1            | 0                    | 0          | 0               | 0              |

| Scheduler Metrics  |  |                          |  |  |  |                         |  |                         |  |  |  |  |  |
|--------------------|--|--------------------------|--|--|--|-------------------------|--|-------------------------|--|--|--|--|--|
| Scheduler Type     |  | Scheduling Resource Type |  |  |  | Minimum Allocation      |  | Maximum Allocation      |  |  |  |  |  |
| Capacity Scheduler |  | [MEMORY]                 |  |  |  | <memory:1024, vCores:1> |  | <memory:4096, vCores:8> |  |  |  |  |  |

| ID                             | User | Name                                                  | Application Type | Queue    | StartTime                     | FinishTime                    | State    | FinalStatus | Progress                         | Tracking UI |
|--------------------------------|------|-------------------------------------------------------|------------------|----------|-------------------------------|-------------------------------|----------|-------------|----------------------------------|-------------|
| application_1472828311239_0004 | boss | select * from tnf.dim_cell un...tnf.dim_cell(Stage-1) | MAPREDUCE        | usecases | Fri Sep 2 16:01:14 +0100 2016 | Fri Sep 2 16:01:36 +0100 2016 | FINISHED | SUCCEEDED   | <div style="width: 100%;"></div> | History     |
| application_1472828311239_0003 | boss | select * from tnf.dim_cell un...tnf.dim_cell(Stage-1) | MAPREDUCE        | usecases | Fri Sep 2 16:00:32 +0100 2016 | Fri Sep 2 16:00:54 +0100 2016 | FINISHED | SUCCEEDED   | <div style="width: 100%;"></div> | History     |
| application_1472828311239_0002 | boss | select * from tnf.dim_cell un...tnf.dim_cell(Stage-1) | MAPREDUCE        | usecases | Fri Sep 2 15:59:57 +0100 2016 | Fri Sep 2 16:00:24 +0100 2016 | FINISHED | SUCCEEDED   | <div style="width: 100%;"></div> | History     |
| application_1472828311239_0011 | boss | select * from tnf.dim_cell un...tnf.dim_cell(Stage-1) | MAPREDUCE        | usecases | Fri Sep 2 16:03:15 +0100 2016 | Fri Sep 2 16:03:38 +0100 2016 | FINISHED | SUCCEEDED   | <div style="width: 100%;"></div> | History     |

Viewing Map reduce job history



### JobHistory

Logged in as: dr.who

- Application
- About Jobs
- Tools

| Retired Jobs            |                         |                         |                        |                                                 |      |          |           |            |                |               |                   |  |  |
|-------------------------|-------------------------|-------------------------|------------------------|-------------------------------------------------|------|----------|-----------|------------|----------------|---------------|-------------------|--|--|
| Submit Time             | Start Time              | Finish Time             | Job ID                 | Name                                            | User | Queue    | State     | Maps Total | Maps Completed | Reduces Total | Reduces Completed |  |  |
| 2016.09.02 15:20:02 IST | 2016.09.02 15:20:07 IST | 2016.09.02 15:20:21 IST | job_1472821868776_0015 | select distinct(d) from tnf.fdr6021(Stage-1)    | boss | usecases | SUCCEEDED | 1          | 1              | 1             | 1                 |  |  |
| 2016.09.02 15:19:15 IST | 2016.09.02 15:19:21 IST | 2016.09.02 15:19:36 IST | job_1472821868776_0014 | select distinct(d) from tnf.fdr6021(Stage-1)    | boss | usecases | SUCCEEDED | 1          | 1              | 1             | 1                 |  |  |
| 2016.09.02 15:13:54 IST | 2016.09.02 15:14:06 IST | 2016.09.02 15:14:21 IST | job_1472821868776_0013 | select count(*) from tnf.dim_cell(Stage-1)      | boss | usecases | SUCCEEDED | 1          | 1              | 1             | 1                 |  |  |
| 2016.08.30 10:24:14 IST | 2016.08.30 10:24:19 IST | 2016.08.30 10:24:41 IST | job_1472499702105_0012 | select count(*) from mobile_up_2_min_5(Stage-1) | boss | default  | FAILED    | 1          | 1              | 0             | 0                 |  |  |
| 2016.08.30 09:57:59 IST | 2016.08.30 09:58:04 IST | 2016.08.30 09:58:28 IST | job_1472499702105_0011 | select count(*) from mobile_up_2_min_5(Stage-1) | boss | default  | FAILED    | 1          | 1              | 0             | 0                 |  |  |

## 5 Troubleshooting installation

### 5.1 Problems and solutions during installation

#### 5.1.1 Port already in use when running setup.sh

If the `setup.sh` script fails due to a port already in use error, then the clients and deployment nodes must be cleaned using the `clean.sh` and `cleanClient.sh` scripts provided. The `clean.sh` script must be run on the deployment node and `cleanClient.sh` script on the client nodes. See Section 3.3.3.

#### 5.1.2 Deployment of the solution fails

If the solution fails to deploy to the nodes provided check the solution installer logs. The main reasons for content failing to deploy to nodes are that the system clocks are not in sync across the machines or invalid machine details were provided.

#### 5.1.3 Telco database fails to create

##### Ensure DB2 is started

1. Logon to the PCI DB2 node (`pcidbnode`).
2. Switch to the user `db2inst1`: `su db2inst1`
3. Start the database: `db2start`
4. List the active databases: `db2 list active databases`

##### Verify that the `db2inst1` password has not expired

1. Open a terminal as root.
2. List the status of the `db2` user: `chage -l db2inst1`.
3. Check the expiration information listed. If the password has expired, then reset it.

**Important:** If the `db2` user password changes, then the Cognos Analytics connection, the SPSS Collaboration & Deployment Services connection and the Database Loader configuration settings for DB2 must be reconfigured.

#### 5.1.4 Cannot locate the Database Loader solution content

The Database Loader solution content is packaged in the Customer Insight download.

**Note:** The location of the Database Loader will be in `/opt/tnf/apps/telco-dbexport` on the `aafnode`.

When deploying content using the Solution Installer the database content contains the Database Loader. In the sample Solution Installer deployment, the content is located in the DB2 node (`pcidbnode`) in the directory `/opt/ibm/telco-dbexport`.

The original rpm is located at `/opt/IBM/IS_CSP_Customer_Insight_1.0.4/`.

#### 5.1.5 Churn/NPS Database Loader fails to run due to no records

The Database Loader will not run if there are no records in the Churn and NPS source tables. Sqoop does not support compressed tables in the format of ORC or Parquet. Due

to this reason when the Churn dataset runs staging tables are created. The following tables must be populated prior to the Churn or NPS Database Loader running:

```
stage.subscriber_billing_etl
stage.subscriber_care_etl
stage.subscriber_crm_etl
stage.cgr_device_etl
tnf.churn_data
tnf.subscriber_level_cx_score_weekly
tnf.nps_score_table
tnf.nps_score_table_latest
```

### 5.1.6 Churn/NPS Models do not execute

The Database Loader script automatically triggers the Churn/NPS Models. If the models fail to run, there can be a number of reasons. Firstly, verify the credentials are correct and base64 encoded. Manually try and login to the Collaboration and Deployment Services Windows client with the credentials. The first time the Database Loader triggers the Churn model the Model will fail to execute successfully because the model has not been trained. Ensure after the first execution of the loading of the data to DB2 that the model is trained.

To verify that the model did execute, complete the following steps:

1. Open the Collaboration and Deployment services Windows client.
2. Right click the job that failed and select Show Job History.
3. Expand the correct job in the Job History view and locate the Log file.
4. Verify there are no errors in the log file.

### 5.1.7 Database Loader Cron Jobs are not visible as boss

Complete section 3.8.8 and 3.8.9 to ensure the Database Loader Cron jobs are created and visible as boss.

### 5.1.8 Cron Jobs not running

If no cron jobs run ensure the cron job service is started. To verify if it is started run `pgrep cron`. If a number is returned then the cron service is running if not then start cron by running `sudo /etc/init.d/cron start`.

### 5.1.9 Cron Job removed or does not exist

Recreate the cron job by completing the steps in section 3.8.8 and 3.8.9.

### 5.1.10 Collaboration and Deployment Services Windows client cannot connect to the Collaboration and Deployment Services repository

**Verify that the db2 instance containing the Collaboration and Deployment services database is running:**

1. Log on to the PCI DB2 node.
2. Switch to the db2inst1 user: `su db2inst1`.

3. List all active databases: `db2 list active databases.`

Ensure the output contains SPSSDB which is the Collaboration and Deployment services database.

If there are issues connecting or listing active databases ensure that the `db2inst1` user has not expired. To determine the expiration details of the db2 user complete the following steps:

1. Open a terminal as root.
2. List the status of the db2 user: `chage -l db2inst1.`
3. Check the expiration information listed. If the password has expired, then reset it.

**Note:** Be careful if the db2 user password changes then Cognos Analytics and the SPSS Collaboration and Deployment Services connections to DB2 will be required to be reconfigured.

### **Verify that Collaboration and Deployment Services application server is running.**

1. Navigate to the Collaboration and Deployment services WebSphere administration user interface.
2. To determine the URL of the administration console, do the following:
  - o Open a virtual network connection to the SPSS PCI node
  - o Select Applications > IBM WebSphere > IBM WebSphere Application Server V 8.5 > profiles > CNDS profile > Administration Console. On click a browser window is displayed with the URL loaded.
3. If the URL does not load then WebSphere is not started.

Refer to the PCI 1.1.1 documentation for details on starting the SPSS WebSphere server.

### **Verify that the Collaboration and Deployment Services application server can connect to the db2 instance.**

1. Logon to the SPSS WebSphere Administration console as per above.
2. Navigate in the left hand menu to Resources > JDBC > Data sources.
3. Select the check box beside CDS\_Datasource and select the Test connection button.
4. If the connection succeeds, then the Collaboration and Deployment services application can connect to DB2. If the connection fails then ensure db2 is started. If db2 is started verify the credentials are correct. If they have changed since installing PCI then move to step 5.
5. Select the CDS\_Datasource.
6. In the right panel select the JAAS – J2C authentication data link.
7. Select the CDS\_Auth\_Alias.
8. Enter the correct password for the db2inst1 user in the password textbox.

Ensure the Collaboration and deployment services credentials are valid.

Ensure that the Collaboration and Deployment Services client can ping the Collaboration and Deployment Services server.

#### 5.1.11 No DB2 datasource is available to select in Modeler Client

Repeat steps in section 3.9.3 ensuring the symbolic link created correctly in step 10. Generally, the main issue with connecting Modeler Client to a configured DB2 instance is that the symbolic link has been created incorrectly.

#### 5.1.12 Analytic Server connection fails in Modeler Client

Ensure Analytic Server is installed on the Hadoop node and that the service is running. To start Analytic Server:

1. Login to the Ambari user interface: `https://host:port/#/login`.
2. Select the SPSS Analytic Server service in the left hand panel. If it does not exist as a service then SPSS Analytic Server is not installed.

**Note:** Analytic Server is an optional installation component in PCI 1.1.1 and it must be installed in order to use the CI solution. If the service is installed proceed to step 3.

3. Select Service Actions > Start.

#### 5.1.13 The SPSS model fails to run

There are a number of reasons why a model might fail to run. In order to determine the reason, check the job run logs. Open the Collaboration and Deployment services Windows client. Right click the job that failed and select Show Job History. Expand the correct job in the Job History view and locate the log file. The reason for the failure is displayed in the log file.

If the issue relates to credentials, ensure the Resource Definitions > Credentials are set correctly in the Collaboration and Deployment services Windows client.

#### 5.1.14 Jobs are queued in Collaboration and Deployment Services

If Jobs on trigger are moving into a Queued state in the Job History windows then refer to the following tech note: <http://www-01.ibm.com/support/docview.wss?uid=swg21673950>

#### 5.1.15 Provisioning a CSV fails

If a CSV file fails to provision, the `install_telsol` log file displays an error. The provisioning file that failed to load can be loaded manually. Determine the error from the log, fix the error and then reload the file. To reload the file, run the following commands:

```
su boss
cd /opt/tnf/apps/bis-main-var/bis-provisioning-tool/
./load.sh -f <file name> -t <table name> where file name is the name of the
csv provisioning file and table name is the name of the table that the csv file will populate.
```

#### 5.1.16 How to verify that all RPMs are installed

Log in to a terminal as boss and run the following command:

```
sudo rpm -qg 'Application/TNF'
```

The result of the command should return the following:

```
customer-profile-data-setup-<version>-r<timestamp>..x86_64
user-profile-<version>-r<timestamp>..x86_64
net-promoter-score-<version>-r<timestamp>..x86_64
customer-profile-<version>-r<timestamp>..x86_64
customer-experience-<version>-r<timestamp>..x86_64
churn-dataset-<version>-r<timestamp>..x86_64
bis-cognosreport-dictionaries-<version>-r<timestamp>..x86_64
customer-behaviours-<version>-r<timestamp>..x86_64
ott-applications-<version>-r<timestamp>..x86_64
```

### 5.1.17 How to verify cron jobs are set up correctly

Each dataset installed configures a cron job to run the dataset at a preconfigured time interval. To list the cron jobs that are configured, run the following commands at the UNIX prompt on the Hadoop Master node (aafnode):

```
su boss
crontab -l
```

The response should contain:

```
30 0 * * * /opt/tnf/apps/customer-profile-data-setup/run_customer-
profile-data-setup.sh > /tmp/cpds_daily_cron_all.sh.log 2>&1
30 4 * * Mon /opt/tnf/apps/user-profile/run_user-profile.sh >
/tmp/userprofile_daily_cron_all.sh.log 2>&1
30 0 * * Mon /opt/tnf/apps/net-promoter-score/run_nps.sh >
/tmp/nps_weekly_cron_all.sh.log 2>&1
30 2 * * Mon /opt/tnf/apps/churn-dataset/run_churn_dataset.sh >
/tmp/cp_weekly_cron_all.sh.log 2>&1
45 * * * * /opt/tnf/apps/customer-experience/run_cea_hourly.sh >
/tmp/cea_hourly_cron_all.sh.log 2>&1
30 03 * * * /opt/tnf/apps/customer-experience/run_cea_daily.sh >
/tmp/cea_daily_cron_all.sh.log 2>&1
30 00 * * MON /opt/tnf/apps/customer-experience/run_cea_weekly.sh >
/tmp/cea_weekly_cron_all.sh.log 2>&1
30 05 * * MON /opt/tnf/apps/churn-dataset/run_churn_dataset.sh >
/tmp/churn_cron_all.sh.log 2>&1
45 * * * * /opt/tnf/apps/customer-
behaviours/customer_behaviour/run_all_hourly.sh >
/tmp/cb_cron_all.sh.log 2>&1
```

```
30 03 * * * /opt/tnf/apps/customer-
behaviours/customer_behaviour/run_daily_rollup.sh >
/tmp/cb_daily_cron_all.sh.log 2>&1

15 03 * * * /opt/tnf/apps/customer-
behaviours/weighted_interest/run_weighted_interest_daily.sh >
/tmp/wi_daily_cron_all.sh.log 2>&1

30 00 * * MON /opt/tnf/apps/customer-
behaviours/customer_behaviour/run_weekly_rollup.sh >
/tmp/cb_weekly_cron_all.sh.log 2>&1

30 2 * * Mon /opt/tnf/apps/ott-applications/run_ott-applications.sh
> /tmp/ott_weekly_cron_all.sh.log 2>&1
```

### 5.1.18 What to do if a dataset fails to run

If a dataset fails to run, data will not be entered into the output table of the dataset and an error/failure message will be logged to the datasets log file. For example if the Churn dataset failed, then the log file located at `/opt/tnf/apps/churn-dataset/log` on the Hadoop Master node (aafnode) would contain a failure notice and the `churn_data` table would not be populated in the `tnf` schema.

The reason for the failure is located in the Hadoop hive log located on the aafnode at `/tmp/boss/hive.log`.



## 6 Appendices

### 6.1 Sample Deployment Sequence Worksheet

| Preparation Tasks    | Description                                                                                                                    | Details / Notes |
|----------------------|--------------------------------------------------------------------------------------------------------------------------------|-----------------|
| Review tech notes    | Ensure you have read any technical notes associated with the software release published on the download page                   |                 |
| Review release notes | Ensure you have read the release notes in the software package                                                                 |                 |
| Hardware check       | VM                                                                                                                             |                 |
|                      | Blade                                                                                                                          |                 |
| Software check       | IBM Predictive Customer Intelligence: (Cognos Analytics, SPSS, DB2)                                                            |                 |
|                      | IBM Big Insights 4.1 (BigSQL and BigSheets)                                                                                    |                 |
|                      | IBM Open Platform (Ambari, Hadoop, Hive, Knox, Parquet, Sqoop)                                                                 |                 |
|                      | IBM Streams                                                                                                                    |                 |
|                      | Analytics Accelerator Framework                                                                                                |                 |
| Server preparation   | Red Hat<br>You have a Red Hat Enterprise Linux operating system that you can install to.                                       |                 |
|                      | Networking<br>You have configured your computer firewall settings for the installation of Linux                                |                 |
|                      | Security settings                                                                                                              |                 |
|                      | Service setup<br>You understand your deployment environment. There are a combination of nodes that you must install to, so you |                 |

|                           |                                                                                                                                                                                                                                                 |  |
|---------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|
|                           | need to determine where you want the various components of the solution to reside.                                                                                                                                                              |  |
|                           | <p>Configure users</p> <p>Ensure you are either root user or have sudo permission on each node computer</p> <p>Disable requiretty during the installation.</p> <p>You have modified the sudoers file for the user who runs the installation</p> |  |
|                           | <p>Solution Installer</p> <p>Check the Solution Installer works only on that version of Linux.</p>                                                                                                                                              |  |
| Other                     |                                                                                                                                                                                                                                                 |  |
|                           |                                                                                                                                                                                                                                                 |  |
| <b>Prerequisites</b>      | <b>Before deploying the CI solution</b>                                                                                                                                                                                                         |  |
| Install PCI               | Follow this <a href="#">document</a>                                                                                                                                                                                                            |  |
| IBM Big Insights          | Follow this <a href="#">document</a> .                                                                                                                                                                                                          |  |
| IBM Open Platform         | Follow this <a href="#">document</a> .                                                                                                                                                                                                          |  |
| IBM Streams               | Follow this <a href="#">document</a> .                                                                                                                                                                                                          |  |
| Other                     |                                                                                                                                                                                                                                                 |  |
|                           |                                                                                                                                                                                                                                                 |  |
| <b>Deployment</b>         | <b>Deploying the solution</b>                                                                                                                                                                                                                   |  |
| Download CI               | Downloading the Customer Insight for CSP solution                                                                                                                                                                                               |  |
| Deploy Solution Installer | Deploying the Customer Insight for CSP s                                                                                                                                                                                                        |  |
| Deploy CI datasets        | Deploying the datasets                                                                                                                                                                                                                          |  |
| Deploy Telco database     | Creating the Telco Database                                                                                                                                                                                                                     |  |

|                           |                               |  |
|---------------------------|-------------------------------|--|
| Deploy Database Loader    | Deploying the Database Loader |  |
| Configure SPSS components | Configuring SPSS components   |  |
| Deploy SPSS Modeler       | Deploying the SPSS Modeler    |  |
| Deploy Cognos             | Deploying the visualization   |  |
| Other                     |                               |  |

## 6.2 Sample Data Required Worksheet

| Installation Node           | Description                            | Details/Notes |
|-----------------------------|----------------------------------------|---------------|
| Node                        |                                        |               |
| Hostname                    |                                        |               |
| Username                    |                                        |               |
| Password                    |                                        |               |
| Other                       | Root directory packages download       |               |
|                             |                                        |               |
| <b>Solution Installer</b>   | <b>Description</b>                     |               |
| URL                         | Solution Installer web server URL      |               |
| Username                    | Solution Installer web server username |               |
| Password                    | Solution Installer web server password |               |
| Other                       |                                        |               |
|                             |                                        |               |
| <b>PCI / Windows Client</b> | <b>Description</b>                     |               |
| Node                        |                                        |               |

|                        |                    |  |
|------------------------|--------------------|--|
| Hostname               |                    |  |
| Username               |                    |  |
| Password               |                    |  |
| Other                  |                    |  |
|                        |                    |  |
| <b>PCI/SPSS Node</b>   | <b>Description</b> |  |
| Node                   |                    |  |
| Hostname               |                    |  |
| Username               |                    |  |
| Password               |                    |  |
| Other                  |                    |  |
|                        |                    |  |
| <b>PCI/DB2 Node</b>    | <b>Description</b> |  |
| Node                   |                    |  |
| Hostname               |                    |  |
| Username               |                    |  |
| Password               |                    |  |
| Other                  |                    |  |
|                        |                    |  |
|                        |                    |  |
| <b>PCI/Cognos Node</b> | <b>Description</b> |  |
| Node                   |                    |  |
| Hostname               |                    |  |
| Username               |                    |  |
| Password               |                    |  |

|                            |                    |  |
|----------------------------|--------------------|--|
| Other                      |                    |  |
| <b>Hadoop Master Node</b>  | <b>Description</b> |  |
| Node                       |                    |  |
| Hostname                   |                    |  |
| Username                   |                    |  |
| Password                   |                    |  |
| Other                      |                    |  |
|                            |                    |  |
| <b>Hadoop Slave Node 0</b> | <b>Description</b> |  |
| Node                       |                    |  |
| Hostname                   |                    |  |
| Username                   |                    |  |
| Password                   |                    |  |
| Other                      |                    |  |
|                            |                    |  |
|                            |                    |  |
| <b>Hadoop Slave Node 1</b> | <b>Description</b> |  |
| Node                       |                    |  |
| Hostname                   |                    |  |
| Username                   |                    |  |
| Password                   |                    |  |
| Other                      |                    |  |
|                            |                    |  |
|                            |                    |  |
|                            |                    |  |

| Hadoop Slave Node 2 | Description |  |
|---------------------|-------------|--|
| Node                |             |  |
| Hostname            |             |  |
| Username            |             |  |
| Password            |             |  |
| Other               |             |  |

Notes:

- Use fully qualified hostname (xxxxxxx.xxx.xx.xxx.com:1) defined in the /etc/hosts file.
- Log in to installation node as root user or as a user with sudo permissions. The working directory for solution installation files is opt/IBM
- Log on to the data node with a user ID that has access to the IBM DB2® database. For example, db2inst1.
- Log on to the BI node as the root user or as a user with sudo permissions.

## 7 Glossary

### **AAF Platform**

Analytics Acceleration Framework (formally Analytics Platform).

### **Ambari**

Apache Ambari is an Apache Hadoop open source component and part of the IBM Open Platform. Ambari is a system for provisioning, managing, and monitoring Apache Hadoop clusters.

### **CI**

Customer Insight (formally Behavioural-Based Customer Insight)

### **Dashboards**

IBM® Cognos® Analytics provides dashboards and stories to communicate your insights and analysis. A dashboard represents an assembled view that contains visualizations such as a graph, chart, plot, table, map, or any other visual representation of data.

### **DB Loader**

DB Loader is a streams application that is part of the CNA 9.1 Mediation Layer system, and loads aggregated records and raw TDRs to the Hadoop DB. The DB Loader can be either centralized (recommended in order to reduce the number of parallel Hadoop loaders) or local (running in every Streams host where ITE or RawTDR is deployed).

### **Database Loader**

Database Loader is used by Customer Insight for CSPs in the ETL process (to extract, transform and load data).

### **IBM Analytic Server**

IBM Analytic Server enables the IBM predictive analytics platform to use data from Hadoop distributions.

### **IBM Collaboration and Deployment Services**

IBM Collaboration and Deployment Services also known as C&DS manages analytical assets such as models, automates processes for example the running of models and efficiently shares results widely and securely.

### **IBM Modeler Client**

IBM SPSS Modeler Client is a powerful, versatile data mining workbench that helps you build accurate predictive models quickly and intuitively, without programming.

### **Mediation**

The process of collecting TDRs from data sources (e.g., TNF Data Collectors (probes), 3rd party Data Source Adaptors), aggregating, enriching and then uploading data into the data

layer (that is, some database like Hadoop, Netezza, Oracle, etc.) for processing by the applications in the application layer (CNA applications, Telco Solutions use cases, etc.).

## **OTT**

Over-the-Top describes a scenario in which a telecommunications service provider delivers one or more of its services across all IP network

## **PCI**

Predictive Customer Intelligence

## **PES**

PES is a file extension for a repository export file. A repository export file packages SPSS models, SPSS jobs and their associated configurations.

## **RPM**

RPM is a file extension for a Red Hat Package Manager file. The file extension is associated with Linux packages and is an archive management package that you can install or deinstall in a Linux environment.

## **Solution**

Refers to the IBM Customer Insight for Communication Service Providers product.

## **Solution Installer**

The Solution Installer is used to deploy solution content to the PCI and Hadoop nodes in your deployment.

## **Sqoop**

Sqoop is a tool used for efficiently transferring bulk data from Hadoop to other datasources. The Customer Insight solution uses Sqoop to transfer data from Hadoop to DB2.

## **TA**

Telecoms Architecture

## **TDR**

Transaction Data Records. Files produced by SourceWorks Data Collectors for input into Mediation for ultimate storage in Hadoop.

## **VM**

Virtual Machine