

Trafodion Install Guide

Version 1.3.0, January 2016

Table of Contents

- 1. About This Document 2
 - 1.1. Intended Audience 2
 - 1.2. New and Changed Information..... 2
 - 1.3. Notation Conventions 2
 - 1.4. Publishing History 5
 - 1.5. Comments Encouraged 5
- 2. Requirements 6
 - 2.1. General Hardware and OS Requirements and Recommendations 6
 - 2.1.1. IP Ports 6
 - 2.2. Prerequisite Software 8
 - 2.2.1. Hadoop Services 8
 - 2.2.2. Software Packages 8

License Statement

Licensed to the Apache Software Foundation (ASF) under one or more contributor license agreements. See the NOTICE file distributed with this work for additional information regarding copyright ownership. The ASF licenses this file to you under the Apache License, Version 2.0 (the "License"); you may not use this file except in compliance with the License. You may obtain a copy of the License at <http://www.apache.org/licenses/LICENSE-2.0>

Unless required by applicable law or agreed to in writing, software distributed under the License is distributed on an "AS IS" BASIS, WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied. See the License for the specific language governing permissions and limitations under the License.

| Version | Date |
|---------|---------------|
| 1.3.0 | January, 2016 |

Chapter 1. About This Document

This guide describes how to install Trafodion for end-user environments. This install allows you to store and query data using Trafodion, either via Trafodion clients (see [Trafodion Client Installation Guide](#) or via application code you write.

If you want to install a Trafodion developer-build environment, then please refer to the [Trafodion Contributors' Guide](#).

1.1. Intended Audience

This guide is intended for:

- **Single-Node Install:** Typically used when you want to evaluate Trafodion.
- **Cluster (Multi-Node) Install:** Typically used when you deploy Trafodion for application usage.



Trafodion can be installed on a single-node or multi-node environment. Unless specifically noted, the term **cluster** is used to mean both single- and multi-node environments.

The installation instructions apply to a diverse set of platforms:

- **Virtual Machines:** Often used for evaluations and Trafodion development.
- **Cloud:** Used for Product Environments as well as for Developer Environments.
- **Bare Metal:** Used for Product Environments as well as for Developer Environments.



The term **node** is used to represent a computing platform on which operating system, Hadoop, and Trafodion software is running. Unless specifically qualified (bare-metal node, virtual-machine node, or cloud-node), **node** represents a computing platform in your cluster regardless of platform type.

1.2. New and Changed Information

This is a new guide.

1.3. Notation Conventions

This list summarizes the notation conventions for syntax presentation in this manual.

| Notation | Usage | Example | Notes |
|-----------------------|--|--|---|
| UPPERCASE LETTERS | Uppercase letters indicate keywords and reserved words. Type these items exactly as shown. Items not enclosed in brackets are required. | SELECT | |
| <i>Italic Letters</i> | Italic letters, regardless of font, indicate variable items that you supply. Items not enclosed in brackets are required. | <i>file-name</i> | |
| Computer Type | Computer type letters within text indicate case-sensitive keywords and reserved words. Type these items exactly as shown. Items not enclosed in brackets are required. | myfile.sh | |
| Bold Text | Bold text in an example indicates user input typed at the terminal. | ENTER RUN CODE: ?123 CODE RECEIVED: 123.00 | The user must press the Return key after typing the input. |
| [] Brackets | Brackets enclose optional syntax items. | DATETIME [<i>start-field</i> TO] <i>end-field</i> | A group of items enclosed in brackets is a list from which you can choose one item or none. The items in the list can be arranged either vertically, with aligned brackets on each side of the list, or horizontally, enclosed in a pair of brackets and separated by vertical lines. For example: DROP SCHEMA schema [CASCADE] DROP SCHEMA schema [CASCADE RESTRICT] |

| Notation | Usage | Example | Notes |
|---------------|---|---|---|
| { } Braces | Braces enclose required syntax items. | <code>FROM { grantee[, grantee]...}</code> | A group of items enclosed in braces is a list from which you are required to choose one item. The items in the list can be arranged either vertically, with aligned braces on each side of the list, or horizontally, enclosed in a pair of braces and separated by vertical lines. For example: <code>INTERVAL { start-field TO end-field } { single-field }</code> <code>INTERVAL { start-field TO end-field single-field } —</code> |
| Vertical Line | A vertical line separates alternatives in a horizontal list that is enclosed in brackets or braces. | <code>{expression NULL}</code> | |
| ... Ellipsis | An ellipsis immediately following a pair of brackets or braces indicates that you can repeat the enclosed sequence of syntax items any number of times. | <code>ATTRIBUTE[S] attribute [, attribute]... {, sql-expression}...</code> | An ellipsis immediately following a single syntax item indicates that you can repeat that syntax item any number of times. For example: <code>expression-n...</code> |
| Punctuation | Parentheses, commas, semicolons, and other symbols not previously described must be typed as shown. | <code>DAY (datetime-expression) `@script-file</code> | Quotation marks around a symbol such as a bracket or brace indicate the symbol is a required character that you must type as shown. For example: <code>"{" module-name [, module-name]... "}"</code> |
| Item Spacing | Spaces shown between items are required unless one of the items is a punctuation symbol such as a parenthesis or a comma. | <code>DAY (datetime-expression)</code> <code>DAY(datetime-expression)</code> | If there is no space between two items, spaces are not permitted. In this example, no spaces are permitted between the period and any other items: <code>myfile.sh</code> |

| Notation | Usage | Example | Notes |
|--------------|--|---|-------|
| Line Spacing | If the syntax of a command is too long to fit on a single line, each continuation line is indented three spaces and is separated from the preceding line by a blank line. This spacing distinguishes items in a continuation line from items in a vertical list of selections. | <code>match-value [NOT] LIKE pattern [ESCAPE esc- char-expression]</code> | |

1.4. Publishing History

| Product Version | Publication Date |
|-------------------------|------------------|
| Trafodion Release 1.3.0 | To be announced. |

1.5. Comments Encouraged

The Trafodion community encourages your comments concerning this document. We are committed to providing documentation that meets your needs. Send any errors found, suggestions for improvement, or compliments to:

issues@trafodion.incubator.apache.org

Include the document title and any comment, error found, or suggestion for improvement you have concerning this document.

Chapter 2. Requirements

The current release of Trafodion has been tested with:

- 64-bit Red Hat Enterprise Linux (RHEL) 6.5, 6.6, and 6.7
- SUSE SLES 11.3
- Cloudera CDH 5.2
- Cloudera CDH 5.3
- Hortonworks HDP 2.2

Other OS releases may work, too. The Trafodion project is currently working on better support for non-distribution version of Hadoop.

2.1. General Hardware and OS Requirements and Recommendations

Trafodion assumes an environment based on the requirements of the supported Hadoop distributions. Please verify these assumptions on each server you will install Trafodion on:

- `bash` is available for shell-script execution.
- `sshd` (the `ssh` daemon) is running on each node in the cluster.
- `ntpd` (the `ntp` daemon) is running and synchronizign time on each node in the cluster.
- `/etc/hosts` is set up for fully-qualified node names (FQDN).
- `/etc/resolv.conf` is configured to use a name server tat can resolve commands such as `host -T <FQDN>`. Please attempt to use such a command to validate host-name resolution.
- The Linux Kernel Firewall (`iptables`) has either been disabled (to check: `service iptables status` (requires sudo access) or [ports required by Trafodion](#) have been opened

2.1.1. IP Ports

The following table lists the default ports used by the different Trafodion components plus the configuration file and configuration attribute associated with each port setting.

| Default Port | Configuration File | Configuration Entry | Require d | Range | Protocol | Comment |
|--------------|--------------------|----------------------------|-----------|-------|----------|--------------------------------|
| 4200 | rest-site.xml | trafodion.rest.port | Yes | 1 | REST | Trafodion REST Server. |
| 4201 | rest-site.xml | trafodion.rest.https .port | Yes | 1 | HTTPS | Trafodion REST Server (HTTPS). |

| Default Port | Configuration File | Configuration Entry | Required | Range | Protocol | Comment |
|--------------|--------------------|---------------------------------|----------|----------|----------|---|
| 23400 | dcs-site.xml | dcs.master.port | Yes | <i>n</i> | binary | Start of Trafodion DCS port range. (37800 for Trafodion 1.1) |
| 24400 | dcs-site.xml | dcs.master.info.port | Yes | 1 | HTTP | DCS master web GUI. (40010 for Trafodion 1.1) |
| 24410 | dcs-site.xml | dcs.server.info.port | Yes | <i>n</i> | HTTP | Start of range for DCS server web GUIs. (40020 for Trafodion 1.1) |
| 50030 | mapred-site.xml | mapred.job.tracker.http.address | No | 1 | HTTP | MapReduce Job Tracker web GUI. |
| 50070 | hdfs-site.xml | dfs.http.address | No | 1 | HTTP | HDFS Name Node web GUI. |
| 50075 | hdfs-site.xml | dfs.datanode.http.address | No | 1 | HTTP | HDFS Data Node web GUI. |
| 50090 | hdfs-site.xml | dfs.secondary.http.address | No | 1 | HTTP | HDFS Secondary Name Node web GUI. |
| 60010 | hbase-site.xml | hbase.master.info.port | No | 1 | HTTP | HBase Master web GUI. |
| 60030 | hbase-site.xml | hbase.regionserver.info.port | No | 1 | HTTP | HBase Region Server web GUI. |

There are two port ranges used by Trafodion.

- 23400 is a range, to allow multiple mxosrvr processes on each node. Allow a range of a few ports, enough to cover all the servers per node that are listed in the "servers" file in the DCS configuration directory.
- 24410 is a range as well, enough to cover the DCS servers per node, usually 1 or 2.

On top of the ports identified above, you also need the ports required by your Hadoop distribution. For example:

- [Cloudera Ports](#)
- [Hortonworks Ports](#)

Although not all the ports will be used on every node of the cluster, you need to open most of them for all the nodes in the cluster that have Trafodion, HBase, or HDFS servers on them.

2.2. Prerequisite Software

2.2.1. Hadoop Services

The following Hadoop services must be running on the cluster where you install Trafodion:

2.2.2. Software Packages

You need to install the following software and dependencies before installing Trafodion.