ThoughtSpot

ThoughtSpot Deployment Guide for VMware

Release 6.0

December, 2019

© COPYRIGHT 2015, 2019 THOUGHTSPOT, INC. ALL RIGHTS RESERVED.

910 Hermosa Court, Sunnyvale, California 94085

This document may not, in whole or in part, be copied, photocopied, reproduced, translated, or reduced to any electronic medium or machine-readable form without prior consent in writing from ThoughtSpot, Inc.

All rights reserved. The ThoughtSpot products and related documentation are protected by U.S. and international copyright and intellectual property laws. ThoughtSpot and the ThoughtSpot logo are trademarks of ThoughtSpot, Inc. in the United States and certain other jurisdictions. ThoughtSpot, Inc. also uses numerous other registered and unregistered trademarks to identify its goods and services worldwide. All other marks used herein are the trademarks of their respective owners, and ThoughtSpot, Inc. claims no ownership in such marks.

Every effort was made to ensure the accuracy of this document. However, ThoughtSpot, Inc., makes no warranties with respect to this document and disclaims any implied warranties of merchantability and fitness for a particular purpose. ThoughtSpot, Inc. shall not be liable for any error or for incidental or consequential damages in connection with the furnishing, performance, or use of this document or examples herein. The information in this document is subject to change without notice.

Table of Contents

Configuration overview 2

Set up ThoughtSpot in VMware 5

Configure ThoughtSpot nodes in VMware 11

Install ThoughtSpot clusters in VMware 18

VMware configuration overview

Congratulations on purchasing the ThoughtSpot instance. This section is an overview of the ThoughtSpot AI-Driven analytics platform hosted on the VMware vSphere Hypervisor (ESXi) 6.5 environment.

About ThoughtSpot in VMware

The VMware virtualization platform provides highly scalable and efficient memory and CPU resources management that can be used by ThoughtSpot instances. Additionally, the VMware virtualization environment is an easy transition between development and production environments. The following diagram shows the components of a VMware and ThoughtSpot architecture:

Note: This is a generic representation; Only CentOS-based virtual machines are supported with ThoughtSpot.

Your database capacity will determine the number of ThoughtSpot instances and the instance network/storage requirements. In addition, you can scale your ThoughtSpot VMs as your dataset size grows.

Supported configurations

ThoughtSpot Engineering has performed extensive testing of the ThoughtSpot platform in VMware for the best performance, load balancing, scalability, and reliability. Based on this testing, ThoughtSpot recommends the following *minimum specifications* for an individual VMware ESXi host machine:

Per VM user data capacity	CPU/RAM	Data disk
20 GB	16/128 GB	800 GB
100 GB	32/256 GB	800 GB
256 GB	72/512 GB	6 TB
Note: All cores must be hyperthreaded. 200GB SSD boot disk required for all configurations.		

Locally attached storage provides the best performance.

SAN can be used, but must comply with the following requirements:

- 136 MBps minimum random read bandwidth
- 240 random IOPS (~4ms seek latency)

NAS/NFS is not supported since its latency is so high that it tends to be unreliable.

All virtualization hosts should have VMware vSphere Hypervisor (ESXi) 6.5 installed.

ThoughtSpot provides a VMware template (OVF) together with a VMDK (Virtual Machine Disk) file for configuring a VM. VMDK is a file format that describes containers for virtual hard disk drives to be used in virtual machines like VMware Workstation or VirtualBox. OVF is a platform-independent, efficient, extensible, and open packaging distribution format for virtual machines.

The ThoughtSpot VM configuration uses thin provisioning and sets the recommended reserved memory, among other important specifications. You can obtain these files from your ThoughtSpot Customer Success Engineer.

Questions or comments?

We hope your experience with ThoughtSpot is excellent. Please let us know how it goes, and what we can do to make it better. You can contact ThoughtSpot [See page 0] by filing a support ticket, email or phone.

Set up ThoughtSpot in VMware

Summary: Learn how to install a ThoughtSpot cluster in a VMware environment.

This page explains how to install a ThoughtSpot cluster in a VMware VSphere Hypervisor (ESXi) 6.5 environment. For each hardware node, you must:

- Complete the prerequisites
- Use the ThoughtSpot Open Virtualization Format (OVF) file to create a virtual machine (VM)
- Add hard disks to the VM

Prerequisites

This installation process assumes you have already acquired your host machines. You can install on a cluster with any number of nodes. A one node cluster is suitable for a sandbox environment but is insufficient for a production environment. You need at least three nodes for high availability (HA), but there is no limit on the number of nodes.

1. Make sure you have installed the Hypervisor on each of your nodes.

   The VM template, by default, captures a 72-core configuration. If your physical host has more than 72 cores, you may want to edit VM to have (`n–2`) cores (for a physical host with n cores) to fully take advantage of computing power of the physical host. Extra cores help performance.

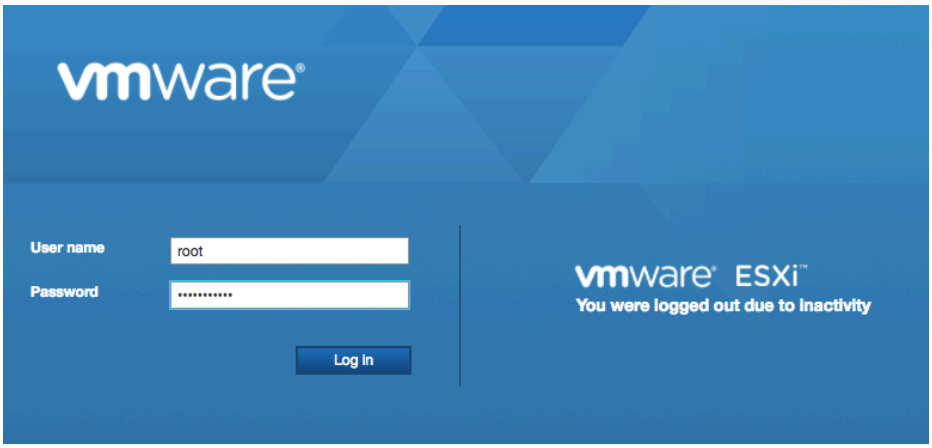
   You should aim to allocate 490 GB or more RAM.

2. Create datastores for all solid-state drive (SSD) and hard drive devices.

Use the OVF to Create a VM

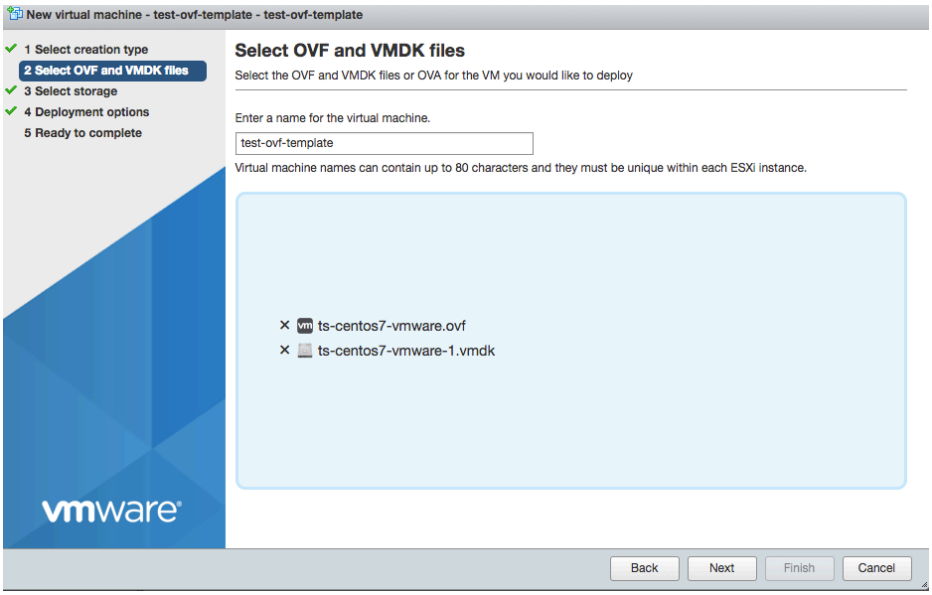
1. **Download [See page 0]** the `ThoughtSpot OVF` to a location on an accessible disk.

2. Log into the ESXi web portal.



3. Select **Virtual Machines > Create/Register VM**.

   The system displays the dialog for selecting an OVF template.



4. Choose the OVF template and click **Next**.

   The system prompts you to select a storage.

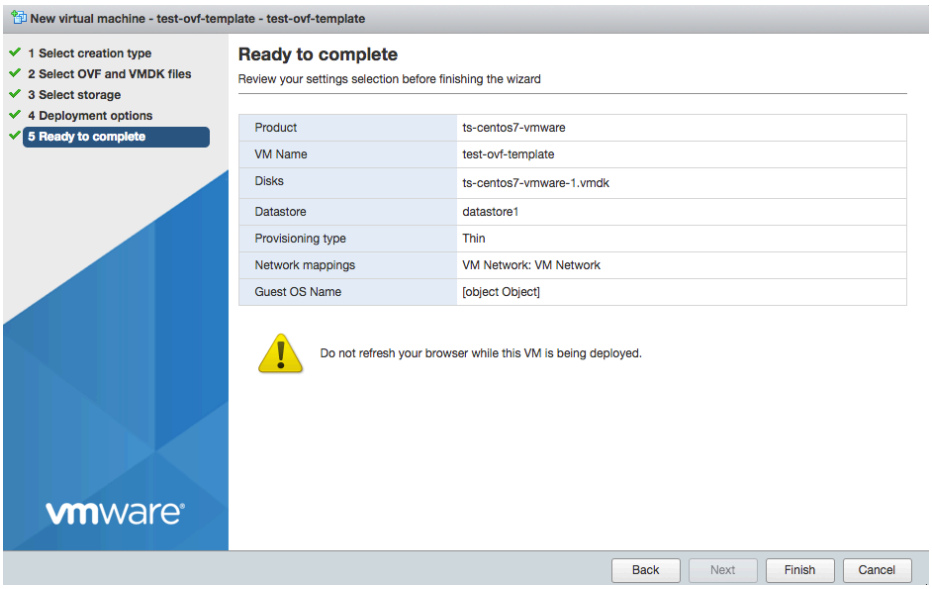
5. Choose the SSD as the destination and click **Next**.

   The system displays the **Deployment Options** dialog.

6. Enter the options and click **Next**.

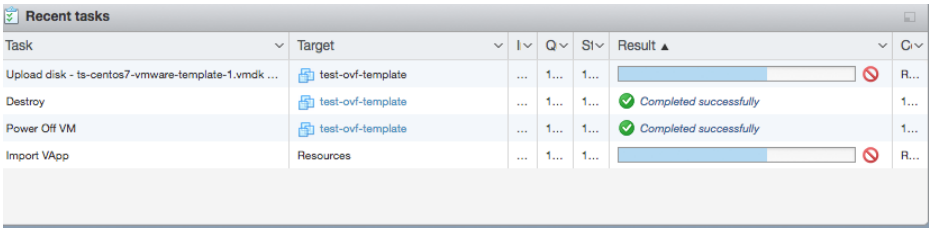
Setting	Value
Network mappings	Select the correct network for your installation.
Disk provisioning	Choose Thin.
Power on automatically	Check this box.

7. Review your selection and click **Finish**.



8. Wait for the template to be loaded.

   Depending on your network speed, loading can take several minutes.



9. Make sure that VM is powered off.

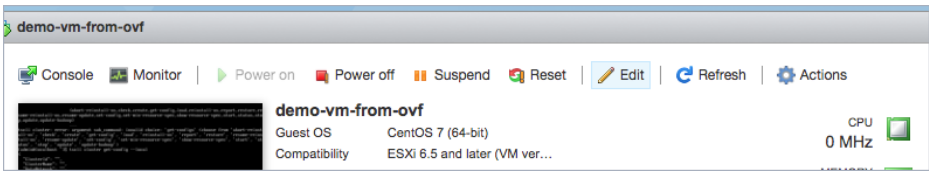
Add hard disks to the VM

Use Case	HDFS Disk Requirements
POC	2 x 1 TB on HDD
Production	3 x 2 TB on HDD

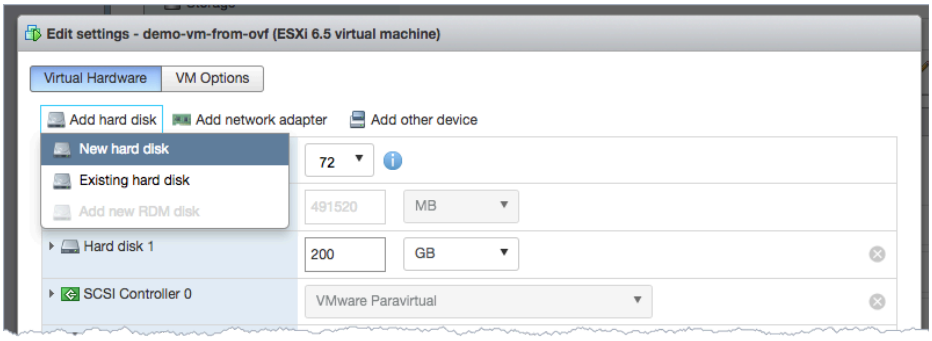
For a proof of concept (POC), follow these steps to create two 1TB HDFS disks on HDD storage, as shown here (2 x 1TB).

For production deployments, ThoughtSpot requires you to have three 2TB HDFS disks on HDD (3 x 2TB). For this use case, follow these same steps to create the additional, larger capacity disks.

1. Edit the VM you just created.



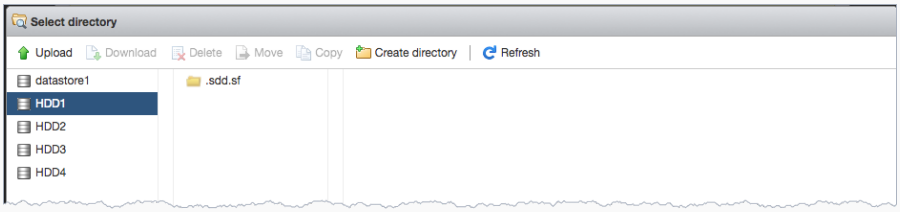
2. Select **Add hard disk > New hard disk**.



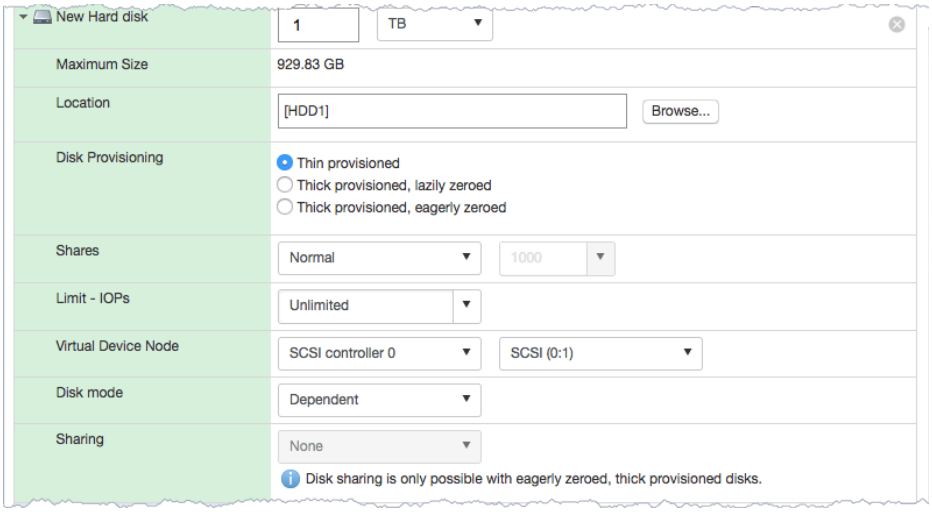
   The new disk appears as a new row under the only existing SSD row.

3. Click the **New Hard disk** to expand the detailed configuration options.

4. For a proof of concept, set the options as follows. (For production deployments, set the size to 2TB.)

Setting	Value
size	1 TB
Location	Use the **Browse** button to select the hard disk store. 
Thin provisioned	Check this box.

You should see something similar to the following:



5. Save your changes.
6. Repeat steps 1-5 to create more hard disks.
7. Power on the VM
8. After the VM is online, run the following command to prepare the HDFS disks:

```
sudo /usr/local/scaligent/bin/prepare_disks.sh
```

Next steps

There is no network at this point on your VMs. As a prerequisite:

1. Verify that Network Adapter type is set to VMware vmxnet3 (Recommended).
2. Verify that all ESXi hosts in your VMware farm for ThoughtSpot have been trunked to the VLAN assigned to your ThoughtSpot VMs.
3. Verify that the console of all ThoughtSpot VMs is accessible in VMware vCenter Server.

After you finish, go to the ThoughtSpot Support website [See page 0] and use the support ticket for installation tasks. If necessary, create a new ticket.

Configure ThoughtSpot Nodes in VMware

Summary: Prepare to install your ThoughtSpot cluster by configuring nodes.

Before you can install a ThoughtSpot cluster in VMware, you must configure your nodes.

Installation Prerequisites

Ensure the successful creation of the virtual machines (VMs) before you install the ThoughtSpot cluster in VMware.

1. **Download the OVF** Download the Open Virtualization Format [See page 0] (OVF) file.
2. **Review configuration overview** Refer to VMware configuration overview [See page 2] for detailed instance specs.
3. **Create the instance** Refer to Set up VMware for ThoughtSpot [See page 5] to create and launch your instance.
4. **Review required ports** Refer to Network Policies [See page 0] to view the required ports for successful operation of ThoughtSpot.

Configure Nodes

After creating the instance, you must configure the nodes. Follow the steps in this checklist.

- ❑ Step 1: Log into your cluster [See page 0]
- ❑ Step 2: Get a template for network configuration [See page 0]
- ❑ Step 3: Prepare node configuration [See page 0]
- ❑ Step 4: Configure the nodes [See page 0]
- ❑ Step 5: Confirm node configuration [See page 0]

Step 1: Log into your cluster

Log into your cluster with admin credentials from Terminal on a Mac or a terminal emulator on Windows. Ask your network administrator if you do not know the admin credentials.

1. Run `ssh admin@<clusterIP>` or `ssh admin@<hostname>`.

   Replace `clusterIP` or `hostname` with your specific network information.

   ```
   $ ssh admin@<clusterIP>
   ```

2. Enter your admin password at the prompt.

   Ask your network administrator if you don't know the password.

> **Note:** The password does not appear on the screen as you type it.

Step 2: Get a template for network configuration

Run the `tscli cluster get-config` command to get a template for network configuration for the new cluster. Redirect it to the file `nodes.config`.

You can find more information on this process in the `nodes.config` file reference [See page 0].

```
$ tscli cluster get-config |& tee nodes.config
```

Step 3: Prepare node configuration

1. Add your specific network information for the nodes in the `nodes.config` file, as demonstrated in the autodiscovery of one node example [See page 0].
2. Fill in the areas specified in Parameters of the nodes.config file [See page 0] with your specific network information.

   If you have additional nodes, complete each node within the nodes.config file in the same way.

Do not edit any part of the `nodes.config` file except the sections described in Parameters of the nodes.config file [See page 0]. If you delete quotation marks, commas, or other parts of the code, it may cause setup to fail.

Step 4: Configure the nodes

Configure the nodes in the `nodes.config` file using the `set-config` command.

1. Disable the `firewalld` service by running `sudo systemctl stop firewalld` in your terminal. The `firewalld` service is a Linux firewall that must be off for ThoughtSpot installation. After the cluster installer reboots the nodes, `firewalld` automatically turns back on.

   ```
   $ sudo systemctl stop firewalld
   ```

2. To make sure you temporarily disabled `firewalld`, run `sudo systemctl status firewalld`. Your output should specify that `firewalld` is inactive. It may look something like the following:

   ```
   $ sudo systemctl status firewalld
     ● firewalld.service - firewalld - dynamic firewall da
   emon
        Loaded: loaded (/usr/lib/systemd/system/firewalld.s
   ervice; disabled; vendor preset: enabled)
        Active: inactive (dead)
   ```

3. Run the configuration command: `$ cat nodes.config | tscli cluster set-config`.

   If the command returns an error, refer to set-config error recovery [See page 15].

   After you run the node configuration command, your output appears similar to the following:

```
$ cat nodes.config | tscli cluster set-config

Connecting to local node-scout
Setting up hostnames for all nodes
Setting up networking interfaces on all nodes
Setting up hosts file on all nodes
Setting up IPMI configuration
Setting up NTP Servers
Setting up Timezone
Done setting up ThoughtSpot
```

Step 5: Confirm node configuration

Use the `get-config` command to confirm node configuration.

Your output may look similar to the following:

```
$ tscli cluster get-config

{
  "ClusterId": "",
  "ClusterName": "",
  "DataNetmask": "255.255.252.0",
  "DataGateway": "192.168.4.1",
  "IPMINetmask": "255.255.252.0",
  "IPMIGateway": "192.168.4.1",
  "Timezone": "America/Los_Angeles",
  "NTPServers": "0.centos.pool.ntp.org,1.centos.pool.ntp.or
g,2.centos.pool.ntp.org,3.centos.pool.ntp.org",
  "DNS": "192.168.2.200,8.8.8.8",
  "SearchDomains": "example.company.com",
  "Nodes": {
        "ac:1f:6b:8a:77:f6": {
          "NodeId": "ac:1f:6b:8a:77:f6",
          "Hostname": "Thoughtspot-server1",
          "DataIface": {
            "Name": "eth2",
            "IPv4": "192.168.7.70"
          },
          "IPMI": {
            "IPv4": "192.168.5.70"
          }
        }
  }
}
```

Install ThoughtSpot software

Next, install your ThoughtSpot clusters [See page 18].

Error recovery

`Set-config` error recovery

If you get a warning about node detection when you run the `set-config` command, restart the node-scout service.

Your error may look something like the following:

```
Connecting to local node-scout WARNING: Detected 0 nodes, but f
ound configuration for
only 1 nodes.
Continuing anyway. Error in cluster config validation: [] is no
t a valid link-local
IPv6 address for node: 0e:86:e2:23:8f:76 Configuration failed.
Please retry or contact support.
```

Restart the node-scout service with the following command.

```
$ sudo systemctl restart node-scout
```

Ensure that you restarted the node-scout by running `sudo systemctl status node-scout`. Your output should specify that the node-scout service is active. It may look something like the following:

```
$ sudo systemctl status node-scout
  ● node-scout.service - Setup Node Scout service
    Loaded: loaded (/etc/systemd/system/node-scout.service; ena
bled; vendor preset: disabled)
    Active: active (running) since Fri 2019-12-06 13:56:29 PS
T; 4s ago
```

Next, retry the `set-config` command.

```
$ cat nodes.config | tscli cluster set-config
```

The command output should no longer have a warning.

Related information

Use these references for successful installation and administration of ThoughtSpot.

- the nodes.config file [See page 0]
- Parameters of the nodes.config file [See page 0]
- Using the cluster create command [See page 0]

- Parameters of the `cluster create` command [See page 0]
- ThoughtSpot Documentation [See page 0]
- Contact Support [See page 0]

Install ThoughtSpot Clusters in VMware

Summary: Learn how to install ThoughtSpot clusters in VMware.

Prerequisites

Before you can install your ThoughtSpot clusters in VMware, complete these prerequisites.

1. **Download the OVF** Download the Open Virtualization Format [See page 0] (OVF) file.
2. **Review configuration overview** Refer to VMware configuration overview [See page 2] for detailed instance specs.
3. **Create the instance** Refer to Set up VMware for ThoughtSpot [See page 5] to create and launch your instance.
4. **Review required ports** Refer to Network Policies [See page 0] to view the required ports for successful operation of ThoughtSpot.
5. **Configure nodes** Refer to Configure ThoughtSpot Nodes in VMware [See page 11] to configure your nodes.

Install ThoughtSpot Software

Install the cluster using the release tarball. The estimated installation time is one hour. Follow the steps in this checklist.

- ❏ Step 1: Run the installer [See page 0]
- ❏ Step 2: Check cluster health [See page 0]
- ❏ Step 3: Finalize installation [See page 0]

Refer to your welcome letter from ThoughtSpot to find the link to download the release tarball. If you do not have a link, open a support ticket at ThoughtSpot Support [See page 0] to request access to the release tarball.

Step 1: Run the installer

1. Copy the downloaded release tarball to `/home/admin` using the following command:

```
$ scp <release-number>.tar.gz admin@<hostname>:/home/admin/<file-name>
```

Note the following parameters:

- `release-number` is the release number of your ThoughtSpot instance, such as 5.3, 6.0, and so on.
- `hostname` is your specific hostname.
- `file-name` is the name of the tarball file on your local computer.

2. Create the cluster.

   Run `tscli cluster create` to create the cluster.

```
$ tscli cluster create <release-number>.tar.gz
```

3. Edit the output with your specific cluster information.

   For more information on this process, refer to Using the cluster create command [See page 0] and Parameters of the `cluster create` command [See page 0].

The cluster installer automatically reboots all the nodes after a successful install. The `firewalld` service automatically turns on. At this time, the system is rebooting, which may take approximately 15 minutes.

Log into any node to check the current cluster status:

```
$ tscli cluster status
```

Step 2: Check cluster health

After the cluster installs, check its status using the `tscli cluster status` command.

Your output may look similar to the following:

```
$ tscli cluster status
Cluster: RUNNING
Cluster name    : thoughtspot
Cluster id      : 1234X11111
Number of nodes : 3
Release         : 6.0
Last update     = Wed Oct 16 02:24:18 2019
Heterogeneous Cluster : False
Storage Type    : HDFS

Database: READY
Number of tables in READY state: 2185
Number of tables in OFFLINE state: 0
Number of tables in INPROGRESS state: 0
Number of tables in STALE state: 0
Number of tables in ERROR state: 0

Search Engine: READY
Has pending tables. Pending time = 1601679ms
Number of tables in KNOWN_TABLES state: 1934
Number of tables in READY state: 1928
Number of tables in WILL_REMOVE state: 0
Number of tables in BUILDING_AND_NOT_SERVING state: 0
Number of tables in BUILDING_AND_SERVING state: 128
Number of tables in WILL_NOT_INDEX state: 0
```

Step 3: Finalize installation

After the cluster status changes to `READY`, sign into ThoughtSpot on your browser. Follow these steps:

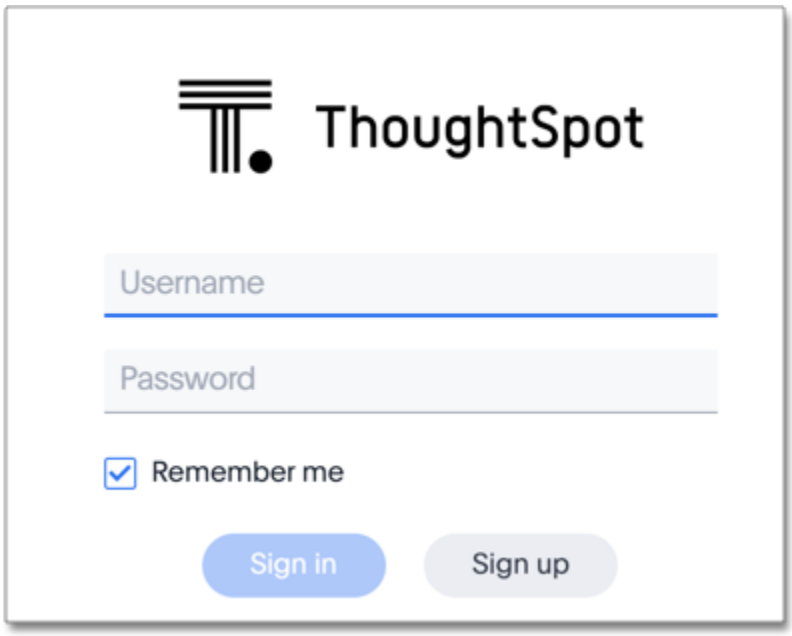
1. Start a browser from your computer.
2. Enter your secure IP information on the address line.

   ```
   https://<IP-address>
   ```

3. If you don't have a security certificate for ThoughtSpot, you must bypass the security warning:
   - Click **Advanced**
   - Click **Proceed**
4. The ThoughtSpot sign-in page appears.

5. In the ThoughtSpot sign-in window [See page 21], enter admin credentials, and click **Sign in**.

   ThoughtSpot recommends changing the default admin password.



ThoughtSpot's sign-in window

Related information

Use these references for successful installation and administration of ThoughtSpot:

- the nodes.config file [See page 0]
- Parameters of the nodes.config file [See page 0]
- Using the cluster create command [See page 0]
- Parameters of the `cluster create` command [See page 0]
- Deployment Overview [See page 0]
- Contact Support [See page 0]