ThoughtSpot

ThoughtSpot Deployment Guide for Microsoft Azure

Release 6.0
December, 2019

Table of Contents

Azure configuration options

ThoughtSpot can be deployed in your Azure environment by deploying compute (VM) instances in your VNET as well as an underlying persistent storage infrastructure. Currently we support Premium SSD Managed Disks for persistent storage. For more information, see Managed Disks pricing [See page 0] in Microsoft's Azure documentation.

All Azure VMs (nodes) in a ThoughtSpot cluster must be in the same availability zone (and, therefore, also in the same region). ThoughtSpot does not support deploying VMs(nodes) of the same cluster across availability zones. For more information, see What are Availability Zones in Azure? [See page 0] in Microsoft's Azure documentation.

ThoughtSpot Azure instance types

Per VM user data capacity	Instance type	CPU/ RAM	Recommended per-VM Premium SSD Managed Disk volume
200 GB	E64sv3	64/432	2x1 TB
100 GB	E32sv3	32/256	2X 400 GB
20 GB	E16sv3	16/128	2X 400 GB
120 GB	D64v3	64/256	2X 1 TB

Set up ThoughtSpot in Azure

Summary: After you determine your configuration options, you must set up your virtual machines using a ThoughtSpot image for Azure.

About the ThoughtSpot image

To provision ThoughtSpot in the Azure portal, access the ThoughtSpot Virtual Machine in the Azure Marketplace.

The ThoughtSpot Virtual Machine comes provisioned with the custom ThoughtSpot image to make hosting simple. A virtual machine is a preconfigured template that provides the information required to launch an instance of ThoughtSpot. It includes the following:

- A template for the root volume for the instance (for example, an operating system, an appliance server, and applications).

The ThoughtSpot Virtual Machine has the ThoughtSpot software installed and configured, on a CentOS base image. Check with your ThoughtSpot contact to learn about the latest version of the ThoughtSpot Virtual Machine.

Due to security restrictions, the ThoughtSpot Virtual Machine does not have default passwords for the administrator users. When you are ready to obtain the password, contact ThoughtSpot Support [See page 0].

Set up ThoughtSpot in Azure

Follow these steps to provision and set up the VMs and launch ThoughtSpot.

Prerequisites

Complete these steps before launching your ThoughtSpot Virtual Machine:

1. Obtain an Azure login account.
2. Set up usage payment details with Microsoft Azure.
3. Set up a Resource Group [See page 4].

Create an instance

Create a resource group and a resource based on the ThoughtSpot Virtual Machine [See page 0] on the Azure Marketplace [See page 0].

1. Log into the Azure portal.

   In a browser, go to http://azure.microsoft.com [See page 0], and log into your Azure account.

2. Create a Resource Group.

   Specify a name, subscription type, and the region in which you are creating your VMs.
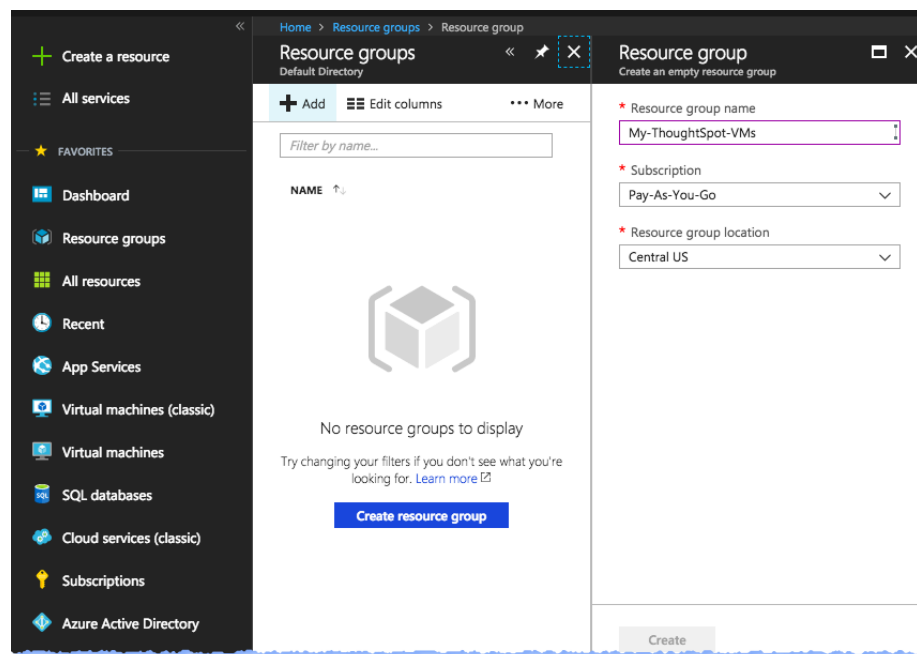
3. Next, create a resource based on the ThoughtSpot Virtual Machine.

   a. Click **Create a resource**. If you already have a resource within your company, use that one.

   b. Search the Marketplace [See page 0] for the ThoughtSpot Virtual Machine, and select it.
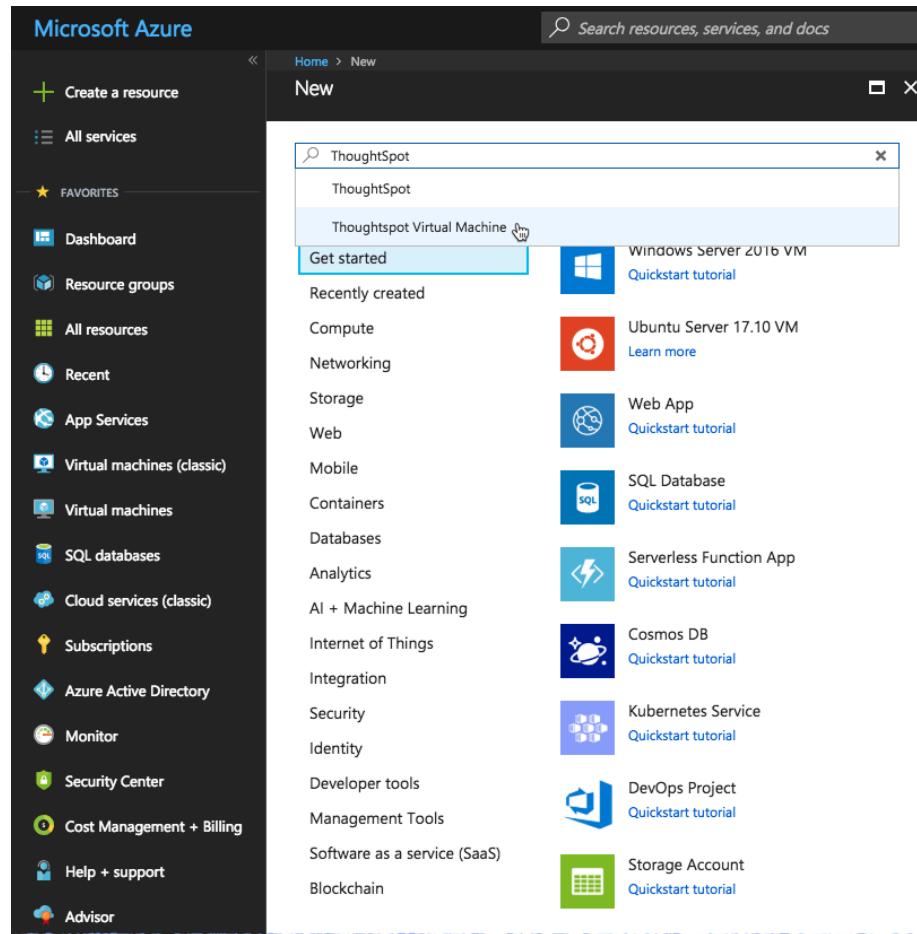
b. On the ThoughtSpot Virtual Machine page, click **Create**.

Configure basic settings

1. Provide a username for your new virtual machine.

2. Select either **SSH public key** or **Password**.

   - If you select **Password**, supply a password and confirm it by typing it again.

   - If you select **SSH public key**, contact ThoughtSpot Support [See page 0] for a key.

3. Choose a disk type.

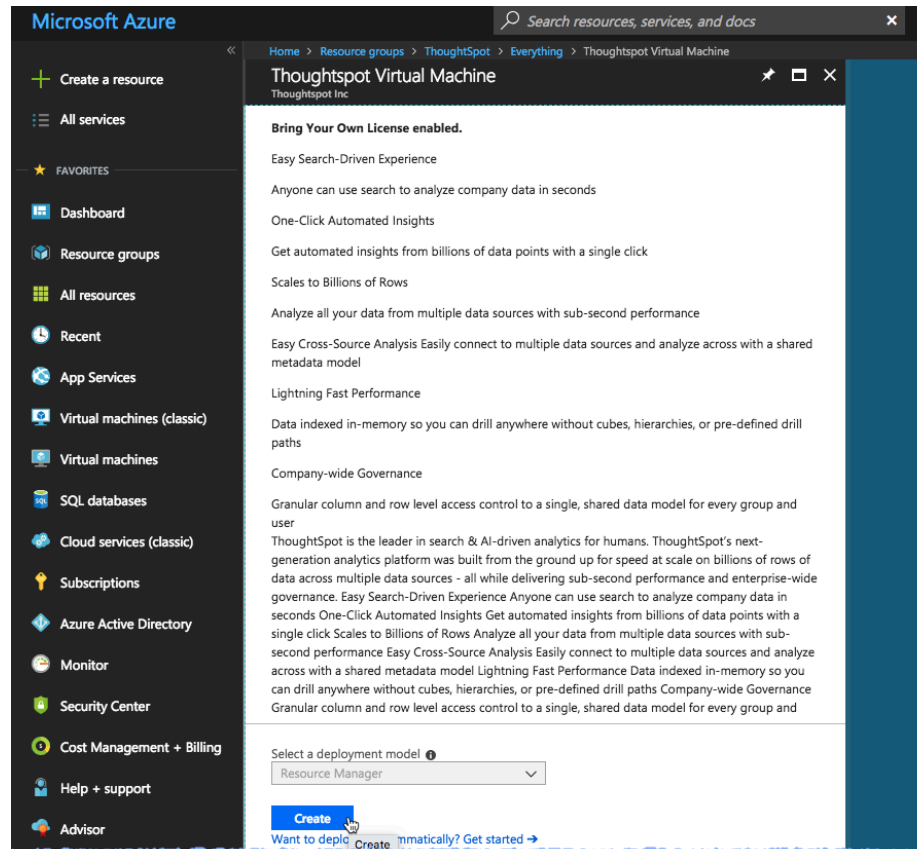   > ☑ **Tip:** The new Standard SSD disk types are only available for particular regions. Make
   >
   > sure this disk type is supported in the region you chose for your VM before selecting it.

   See Standard SSD Disks for Virtual Machine workloads [See page 0] for more on SSD disks.

   ThoughtSpot recommends the Premium SSD disks.

4. Provide a Resource Group, by clicking `existing` and selecting the one you just created.

5. Select a location.

6. Click **OK** to save the Basics, which should look similar to the following example.

Choose a machine size

Under **Choose a size**, select `E64S_V3 standard` . For more information, refer to Azure configuration options [See page 2].

Configure network settings, storage, and other options

Prerequisite: Get the details needed for setting up the Virtual Network, Subnet, and Network Security Group from your Azure support team.

1. For storage, select **Yes** to **use managed disks**.

2. Under **Network**, select **Virtual network**, then **Subnet**, then **Public IP addresses**, and set those names, addresses, and ranges approriately for your network.

3. Open the necessary Inbound and Outbound ports to ensure that the ThoughtSpot processes do not get blocked.

   The minimum ports needed are:

Port	Protocol	Service
22	SSH	Secure Shell access
80	HTTP	Web access
443	HTTPS	Secure Web access
12345	TCP	ODBC and JDBC drivers access
2201	HTTP	Cluster Debugging
2101	HTTP	Node daemon Debugging
4001	HTTP	Data Cache Debugging

   ❶ **Note:** Nodes purchased from Azure must be reachable to each other so that they can communicate and form a distributed environment. ThoughtSpot requires that these ports be accessible between nodes within a cluster. Use your discretion about whether to restrict public access or not for all nodes and all ports.

   Refer to network policies [See page 0] for more information.

4. Configure `auto shutdown`, `monitoring`, `guest OS diagnostics`, and any other settings to your preference. If you have no preference, you can leave them on their default settings.
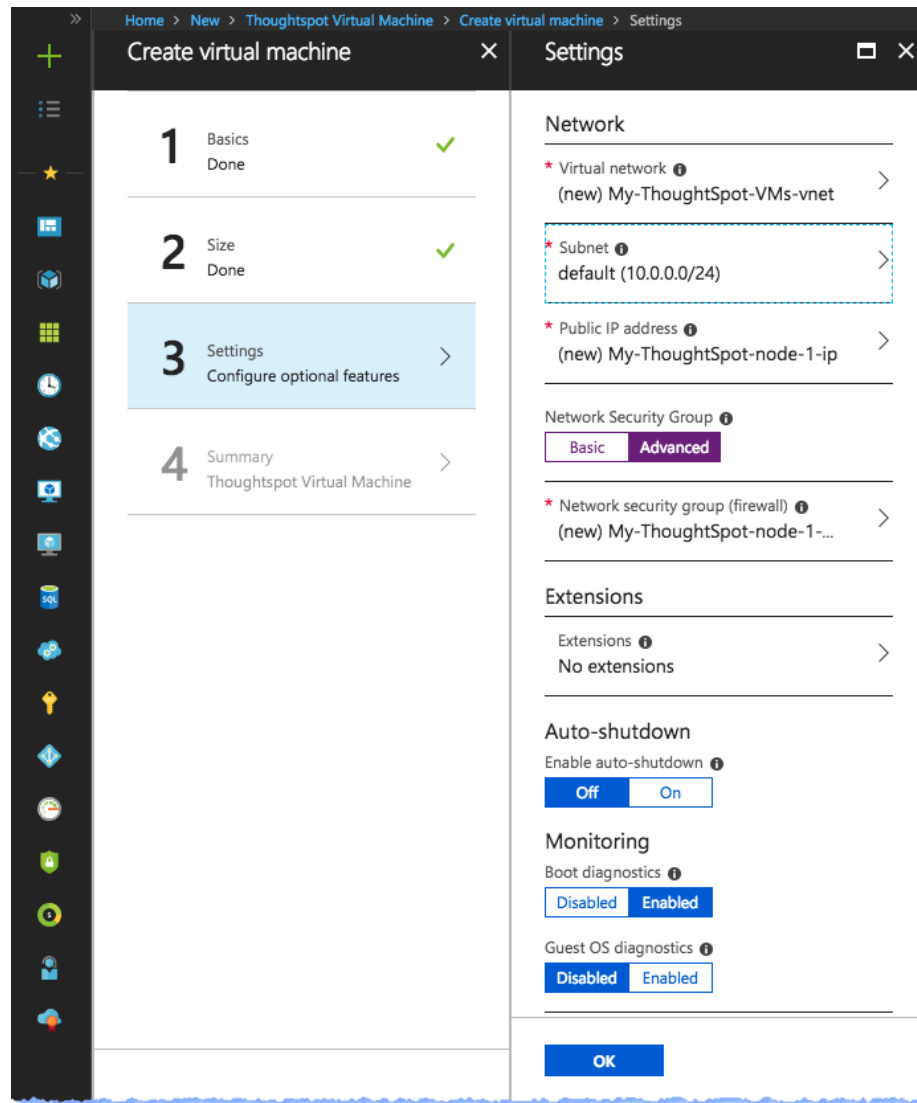
5. Click **OK**.

   Azure will do the final validation check.

Review the Summary

Verify that the validation check succeeded and that summary of information shown is correct. If you find errors, reconfigure as needed.

When you are satisfied with the virtual machine setup, click **Create**.

Prepare for starting up ThoughtSpot

Prerequisite: To log into the VM, you need the private key that is available in the image. You can obtain this from your ThoughtSpot contact.

1. Obtain the VM's public and private IP addresses.

    - To see the public IP, click the VM name link. This will show the public IP of the VM.
    - To see the private IP click Networking (under SETTINGS on the left side of the screen).

2. In a terminal application, connect to the VM through SSH. Use the private key provided for the admin user.

    - You must file a support ticket to obtain this private key; it is necessary for the first login.
    - This key is different from the credentials, or the private keys supplied in earlier steps, which do not work in this context.

    ```
    $ ssh admin@<VM-IP>
    ```

3. Update the password for both the `admin` and the `thoughtspot` users.

    The command prompts you to type in a new password, and then to confirm the password.

    ```
    $ sudo passwd admin
    Changing password for user admin
    $ sudo passwd thoughtspot
    Changing password for user thoughtspot
    ```

    ❶ **Warning:** If you do not change the password, you cannot log back into your Azure VMs. Your private key does not work after initial installation.

4. Update the file `/etc/hosts` with all the node IP addresses for the other VMs that will be part of the ThoughtSpot cluster.

Add Storage Disks

1. Go back to the VM and click it.

2. Add 2 SSD disks of 1TB each.

3. Click **Add data disk** and choose **Create disk from the menu**.

4. Create one mode data disk (demo-disk2) and save them both.

5. Click **Save** to add the disks to the VM.

6. Verify that the disks were added by issuing `lsblk` in your terminal application:

```
$ lsblk
```

Your result may look something like the following:

```
NAME      MAJ:MIN RM  SIZE RO TYPE MOUNTPOINT
fd0         2:0    1    4K  0 disk
sda         8:0    0  200G  0 disk
├─sda1      8:1    0    1G  0 part /mntboot
├─sda2      8:2    0   20G  0 part /
├─sda3      8:3    0   20G  0 part /update
└─sda4      8:4    0  159G  0 part /export
sdb         8:16   0    1T  0 disk
└─sb1       8:17   0    1T  0 part /mnt/resource
sdc         8:32   0    1T  0 disk
sdd         8:48   0    1T  0 disk
sr0        11:0    1  628K  0 rom
```

7. Unmount the temporary disk by issuing the following command:

```
$ sudo umount /mnt/resource
```

8. Prepare the disks /dev/sdc and /dev/sdd for ThoughtSpot by issuing the following command:

> ❶ **Warning:** Do not use the disk `/dev/sdb`. This disk is reserved for ThoughtSpot use.

```
$ sudo /usr/local/scaligent/bin/prepare_disks.sh /dev/sdc /d
ev/sdd
```

1. Check the disks' status by issuing the following command:

```
$ df -h
```

2. Repeat the steps in this section for each node in your cluster.

Create network support settings

> ☑ **Tip:** All changes in this section must be re-applied each time after a cluster is created or
>
> updated. If these changes are not present, a reboot of the VMs will not have network access. So
>
> when updating these files, keep a backup to copy after any subsequent cluster creation or update.

1. SSH into one of your VMs.

```
ssh admin@<VM-IP>
```

2. Update the VM's hostname:

```
$ sudo hostnamectl set-hostname <HOSTNAME>
```

If you are using a static name, you can issue:

```
sudo hostnamectl set-hostname <HOSTNAME> --static
```

3. Update `/etc/hosts` with the IP and hostname:

```
$ sudo vi /etc/sysconfig/network-scripts/ifcfg-eth0

DEVICE=eth0 ONBOOT=yes BOOTPROTO=dhcp HWADDR=<Add eth0 M
AC> TYPE=Ethernet USERCTL=no PEERDNS=yes IPV6INIT=no
```

4. Repeat this process for each node.

5. Do not reboot any of the nodes, until these changes are made to each node:

   a. Open the grub file /update/etc/default/grub in an editor:

   ```
   $ sudo vi /update/etc/default/grub
   ```

   b. Change the line:

   ```
   GRUB_CMDLINE_LINUX="console=tty0 console=ttyS1,115200
   n8"
   ```

   to:

   ```
   GRUB_CMDLINE_LINUX="console=tty0 console=ttyS1,115200
   n8 net.ifnames=0"
   ```

   c. Save your changes.

6. Issue these commands:

   ```
   $ sudo cp /update/etc/default/grub /etc/default/
   $ rm /usr/local/scaligent/bin/setup-net-devices.sh
   ```

7. Reboot the nodes.

Configure ThoughtSpot nodes in Azure

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

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

Installation Prerequisites

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

1. **Review configuration options** Refer to Azure configuration options [See page 2] for detailed instance specs.
2. **Create the instance** Refer to Set up Azure for ThoughtSpot [See page 3] to create and launch your instance.
3. **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

Use Terminal on a Mac or a terminal emulator on Windows to log into your cluster. Log in using the ssh

private key provided by ThoughtSpot.

If you do not have a private key, contact ThoughtSpot Support [See page 0] by email or through the

support portal.

To log into your cluster, run `ssh -i <private-key> admin@<public-vm-ip>` .

```
$ ssh -i <private_key> admin@<public-vm-ip>
```

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 18].

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,

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 Azure

> **Summary:** Learn how to install ThoughtSpot clusters in Azure.

Prerequisites

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

1. **Review configuration options** Refer to Azure configuration options [See page 2] for detailed instance specs.

2. **Create the instance** Refer to Set up Azure for ThoughtSpot [See page 3] to create and launch your instance.

3. **Review required ports** Refer to Network Policies [See page 0] to view the required ports for successful operation of ThoughtSpot.

4. **Configure nodes** Refer to Configure ThoughtSpot nodes in Azure [See page 14] 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/ad
min/<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
```
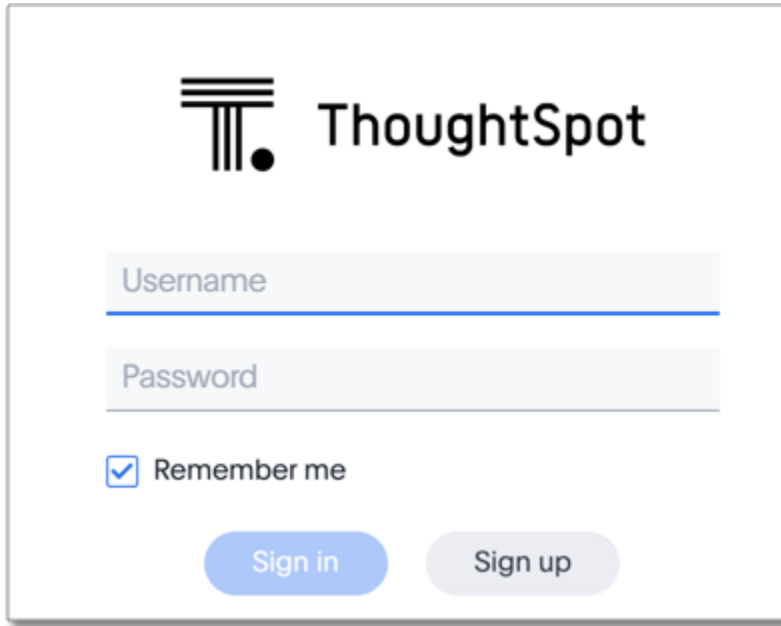
Step 3: Finalize installation

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 24], 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]