

Installing Drill on MapR

You can use the instructions in this document to install the latest version of Drill on MapR.

MapR categorizes certain releases of Drill as "production-ready" and certain releases as "developer previews." For more information about developer preview releases, see [MapR Developer Preview](#).

Prerequisites

Before you install Drill, make sure that the node meets the following prerequisites:

- Installation of [Oracle Java SE Development Kit \(JDK\) 1.7](#) or [OpenJDK 7](#)
- One of the following operating systems for which MapR provides packages:
 - Red Hat 6.1-7.0
 - CentOS 6.1-7.0
 - Ubuntu 12.04, 14.04
 - SUSE Server 11 SP3
- Member of a MapR 4.x or 5.0.0 cluster

Refer to the [Drill Release Notes](#) for information about known issues. Do not add HADOOP_HOME to Drill's classpath. In fact, you do not have to set HADOOP_HOME in current releases of Drill.

Hive and HBase Support

Installation of Hive is optional. Support differs, depending on the release:

- Drill 1.0 supports Hive 0.13.
- Drill 1.1.0 supports Hive 1.0.
- The latest version of Drill on the MapR Distribution for Hadoop 1.2.0 supports all Hive versions available on the corresponding MapR Platform.

Installation of a [supported HBase](#) version is optional.

Configuring the maprtech.repo File

To install Drill on Red Hat, CentOS, Ubuntu or SUSE, a repository file must be configured for the Drill package. The repository for the Drill package is different from the repository for the MapR package. The following sections briefly describe how to set up the repositories on each platform. For more information about setting up repositories, see "[MapR Repositories and Package Archives](#)."

RHEL or CentOS

On RedHat Enterprise Linux (RHEL) or CentOS, a `maprtech.repo` file should be in the following directory:

```
/etc/yum/repos.d
```

The `maprtech.repo` file should contain repository specifications necessary for installing MapR and it might also contain the ecosystem repository that you need to install Drill. Open the file and check the specifications. The following example shows a MapR 5.0.0 `maprtech.repo` file. The first set of properties, `maprtech_v5.0.0`, install MapR 5.0.0. The second set, `maprecosystem`, installs an ecosystem component, such as the production-ready version of Drill.

```
[maprtech_v5.0.0]
name=MapR Technologies
baseurl=http://package.mapr.com/releases/v5.0.0/redhat/
enabled=1
gpgcheck=0
protect=1

[maprecosystem]
name=MapR Ecosystem Components
baseurl=http://package.mapr.com/releases/ecosystem-5.x/redhat/
enabled=1
gpgcheck=0
protect=1
```

To install the developer preview release of Drill 1.3.0 on MapR instead of Drill 1.4.0 on MapR, use the following baseurl in the `maprecosystem`

properties: <http://package.mapr.com/labs/drill/redhat/>

Ubuntu

On Ubuntu, a `sources.list` file should be in the following directory:

`/etc/apt/`

The `sources.list` file should contain a repository specification for installing MapR and it might also contain the specification for the ecosystem repository that you need for installing the latest version of Drill. Open the file and check for the specifications. The following example shows the specifications for a MapR 5.0.0 and version 5.x ecosystem. The first deb specification installs MapR 5.0.0. The second deb specification installs an ecosystem component, such as the production-ready version of Drill:

```
deb http://package.mapr.com/releases/v5.0.0/ubuntu/ mapr optional
deb http://package.mapr.com/releases/ecosystem-5.x/ubuntu binary/
```

To install the developer preview release of Drill 1.3.0 on MapR instead of Drill 1.4.0 on MapR, use the following URL in the second deb specification: <http://package.mapr.com/labs/drill/ubuntu/>

SUSE

As root user, or using `sudo`, issue the following command to configure the repository for MapR ecosystem packages, which includes the latest Drill package, on a MapR 5.0.0 cluster:

```
zypper ar http://package.mapr.com/releases/ecosystem-5.x/suse/ maprecosystem
```

To install the developer preview release of Drill 1.3.0 on MapR instead of Drill 1.4.0 on MapR, use the following URL and change `maprecosystem` to `maprdrill`: <http://package.mapr.com/labs/drill/suse/>

Installing Drill

Using the repositories previously described, complete the following steps to install the latest version of Drill:

1. Issue the following command to install the `mapr-drill` package on a node:

RedHat or CentOS

```
$ sudo yum install mapr-drill
```

Ubuntu

```
$ sudo apt-get install mapr-drill
```

SUSE

```
$ sudo zypper install mapr-drill
```

2. Run `configure.sh` to refresh the node configuration.

Example

```
$ /opt/mapr/server/configure.sh -R
```

3. Start or restart the warden to make configuration changes effective:

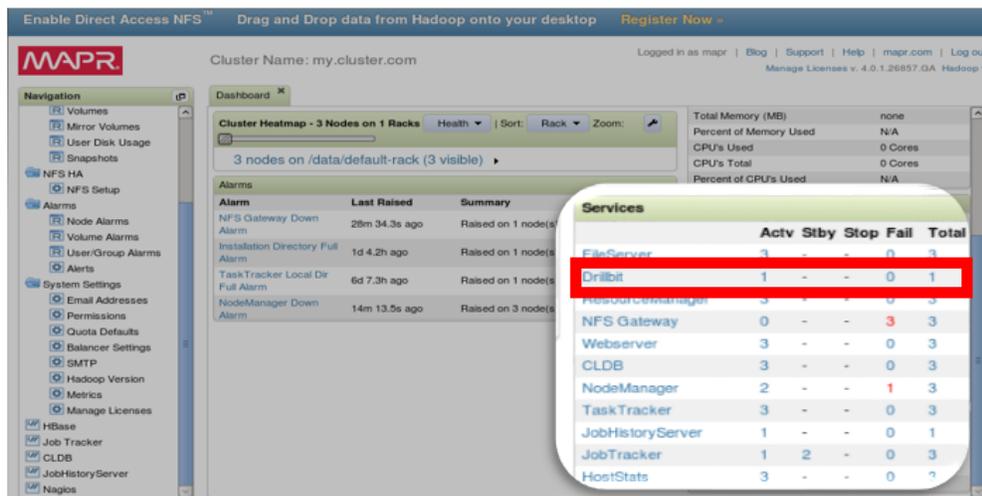
```
service mapr-warden restart
```

- Verify that the Drillbit service is running on the node. It might take a minute or so for the Drillbit to start after starting warden. You can issue the following command to verify the status of the Drillbit service from the command line:

```
jps
```

You should see Drillbit as one of the services listed.

Alternatively, login to the MCS at `https://<host name>:8443` to verify the status of the Drillbit service:



The screenshot shows the MapR Management Console (MCS) interface. The main content area displays a 'Services' table with the following data:

Service	Actv	Stby	Stop	Fail	Total
FileServer	3	-	-	0	3
Drillbit	1	-	-	0	1
ResourceManager	3	-	-	0	3
NFS Gateway	0	-	-	3	3
Webserver	3	-	-	0	3
CLDB	3	-	-	0	3
NodeManager	2	-	-	1	3
TaskTracker	3	-	-	0	3
JobHistoryServer	1	-	-	0	1
JobTracker	1	2	-	0	3
HostStats	3	-	-	0	?

- Repeat the installation process on any other nodes that you want to run Drill.

Starting the Drillbit Service Manually

You can start the Drillbit service on one or more nodes using MCS or the following command:

```
$ maprcli node services -name drill-bits -action start -nodes <node host names separated by a space>
```

Use the host name if possible. Using host names instead of IP addresses is a best practice. Replace `start` with `restart` to restart the service.

Accessing Drill Logs

You can access log files for Drill in `/opt/mapr/drill/drill-<version>/logs/drillbit.log`.

When you have Drill installed and running on selected nodes, you can perform any of the following tasks:

- Review the [Drill Architectural Overview](#).
- [Connect Drill to data sources](#).
- Start the Drill shell and query sample data.
- Install the MapR Drill ODBC Driver to connect to Drill from BI tools.
- Install the JDBC driver and connect to Drill from SquirrelL on Windows.
- Visit the [Drill documentation web site](#) for more information.