ORACLE®

Quick Deployment
Step-by-step instructions to deploy Oracle Big Data Lite Virtual Machine

Version 4.11

Last Updated: 1/10/2018

Please note*:* This appliance is for testing and educational purposes only; it is unsupported and not to be used in production.

CONTENTS

INTRODUCTION

Oracle Big Data Lite Virtual Machine provides an integrated environment to help you get started with the Oracle Big Data platform. Many Oracle Big Data platform components have been installed and configured - allowing you to begin using the system right away. See the Big Data Lite landing page (http://www.oracle.com/technetwork/database/bigdata-appliance/oracle-bigdatalite-2104726.html) on OTN for tutorials, videos, white papers and more.

The following components are included on Oracle Big Data Lite:

- Oracle Enterprise Linux 6.9

Hardware and Software
ORACLE®
Engineered to Work Together

- Oracle Database 12c Release 1 Enterprise Edition (12.1.0.2) - including Oracle Big Data SQL-enabled external tables, Oracle Multitenant, Oracle Advanced Analytics, Oracle OLAP, Oracle Partitioning, Oracle Spatial and Graph, and more.
- Cloudera Distribution including Apache Hadoop (CDH5.13.1)
- Cloudera Manager (5.13.1)
- Oracle Big Data Spatial and Graph 2.4
- Oracle Big Data Connectors 4.11
    - Oracle SQL Connector for HDFS 3.8.1
    - Oracle Loader for Hadoop 3.9.1
    - Oracle Data Integrator 12c (12.2.1.3.0)
    - Oracle R Advanced Analytics for Hadoop 2.7.1
    - Oracle XQuery for Hadoop 4.9.1
    - Oracle Data Source for Apache Hadoop 1.2.1
    - Oracle Shell for Hadoop Loaders 1.3.1
- Oracle NoSQL Database Enterprise Edition 12cR1 (4.5.12)
- Oracle JDeveloper 12c (12.2.1.2.0)
- Oracle SQL Developer and Data Modeler 17.3.1 with Oracle REST Data Services 3.0.7
- Oracle Data Integrator 12cR1 (12.2.1.3.0)
- Oracle GoldenGate 12c (12.3.0.1.2)
- Oracle R Distribution 3.3.0
- Oracle Perfect Balance 2.10.0

Big Data Lite includes software products that are optional on the Oracle Big Data Appliance (BDA), including Oracle NoSQL Database Enterprise Edition, Oracle Big Data SQL, Oracle Big Data Spatial and Graph and Oracle Big Data Connectors.

To get started, you should perform the following high level tasks:

- Ensure your host computer meets the requirements for running the virtual machine
- Download and install VirtualBox and VirtualBox Extension Pack – which is a cross-platform virtualization application
- Download the zip files from OTN and use 7-zip to extract the appliance file
- Create the VM by importing the BigDataLite-xxx.ova file into VirtualBox
- Start the machine!
- Log into Big Data Lite using:

Hardware and Software
ORACLE
Engineered to Work Together

       user id: **oracle**

       password: **welcome1**

- Open the Start Here document on the desktop for details about the environment.

INSTALLATION STEPS

1.   **Prepare your host system.**

- Minimum 8GB of real memory; more is better. 5G memory will be dedicated to the VM. Some details:
  - Big Data Lite w/o Cloudera Manager (CM) requires 5GB memory dedicated to the VM
  - Big Data Lite w/CM running requires 10GB memory
- Turn on Virtual Assist features in the BIOS (usually done by default). Refer to Troubleshooting tips in Step 12 for more details.
- ~73GB disk space needed to download and install:
  - 11 zip files: ~24GB (Can be removed after extraction)
  - .ova file: ~24GB (Result of unzip. It can be removed after importing – but nice to keep around)
  - imported image: ~49GB.
- Download and install 7Zip.
- Download and install md5sum (optional)

2.   **Download and install** Oracle VirtualBox and VirtualBox Extension Pack (version 4.3.26 and above is rec-ommend).

    It is highly recommended that you also install **the Virtual Box Guest Additions** – which provides en-hanced mouse support, clipboard support and sharing of files between the virtual machine and its host.

3.   **Download Big Data Lite files from Oracle Technology Network**

- Go to the Big Data Lite landing page at:
  http://www.oracle.com/technetwork/database/bigdata-appliance/oracle-bigdatalite-2104726.html
- Click to accept the OTN license.
- Download all of the files to a single directory.

4.   **After all zip files are downloaded, extract them with 7zip.**

    In Windows:
- Right-click on file **BigDataLite-xxx.7z.001** file and select **7-zip->Extract Here**
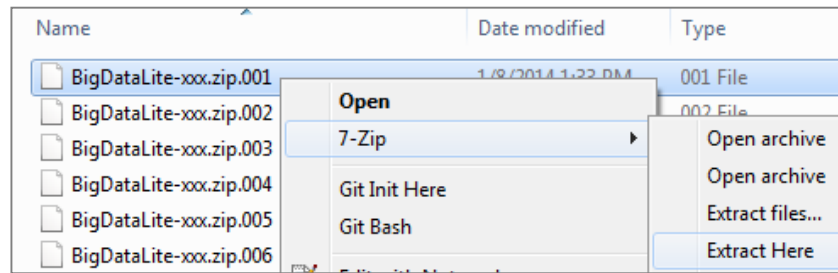
In Linux:

- Run the following command to extract the file:

```
7za e BigDataLite-xxx.7z.001
```

Result:

The extraction will create the **BigDataLite-xxx.ova** appliance file. This file will be used to create a new machine in Virtual Box. This single file contains the entire machine definition, including the physical disks as well as defaults for the machine configuration (e.g. network, CPUs, memory, etc.).

5. **Start Oracle Virtual Box Manager and Import the Appliance**

After starting Virtual Box, create the new virtual machine using the import wizard:
- Click **File -> Import Appliance** to launch the import wizard
- In the Import Virtual Appliance page, click **Open appliance…**
- Locate the **BigDataLite-xxx.ova** file and click **Open**. Click **Next**.
- **Appliance settings** give you an overview of the configuration. In this step, you may need to update the location of the **Virtual Disk Images**. These files will initially use approximately 50GB of disk – and can expand as you use the machine. Change the default location of the Virtual Disk Images if the current location does not have enough capacity. Note, the VM uses 4 separate disks. These disks should not be moved after importing the VM:
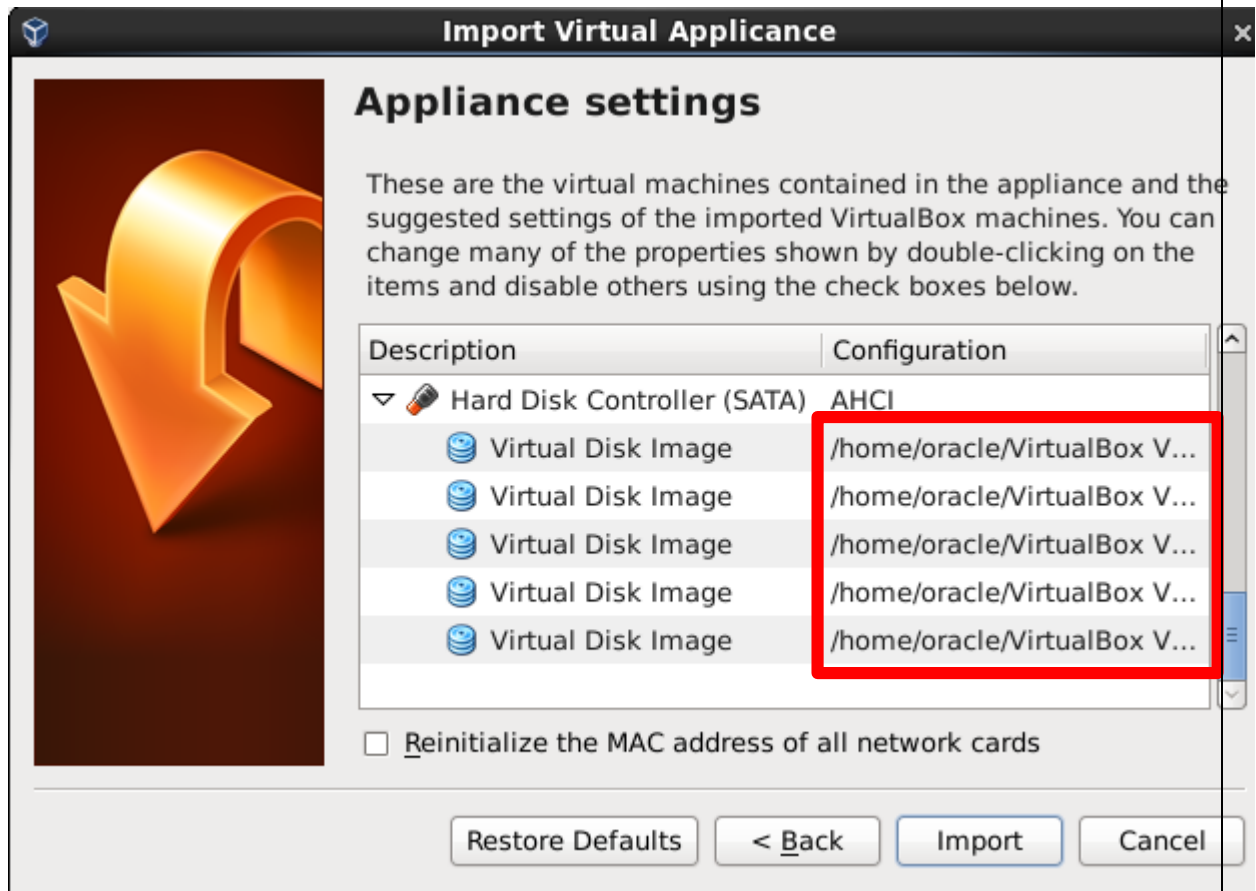
Click **Import**

Result:

A new BigDataLite-xxx Virtual Machine is available in the Virtual Box Manager:
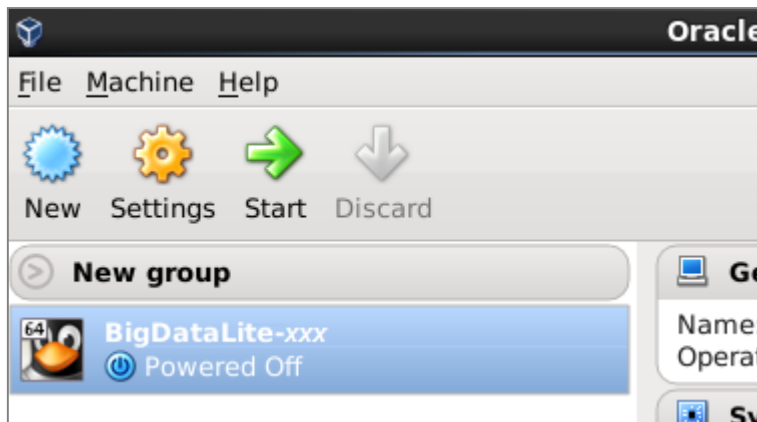
6. **Virtual Machine Settings**

Big Data Lite is configured with the following parameters, which is ideal for an 8GB host system:
- 5GB RAM
- 2 processors
- NAT Network Adapter

As mentioned earlier, if you would like to run Cloudera Manager, you will need to increase the memory allocated to the VM. You can update these default settings by selecting the Big Data Lite VM in Virtual Box Manager and clicking **Settings**:
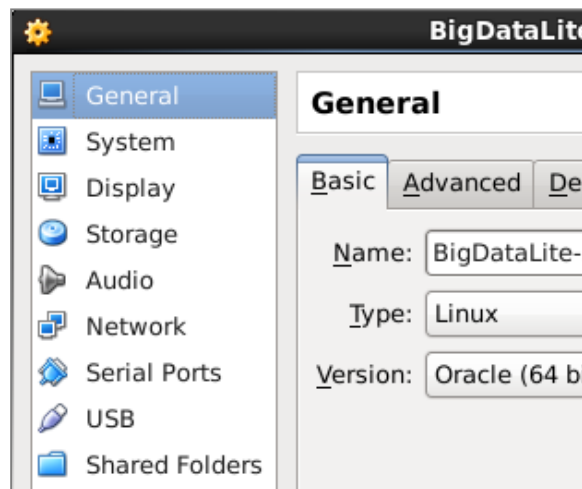
You can update the amount of memory and CPUs allocated to the machine from the **System** tab. Please review the Virtual Box documentation to learn details about updating the configuration.

7. **Start your Big Data Lite!**

Double-click on BigDataLite-xxx to start the VM. Log on as the oracle user to get started:

id:	`oracle`
password:	`welcome1`

This will bring you to a Linux desktop. The **Start Here** document on the desktop will provide you details about the installed software, ids and passwords, and more.
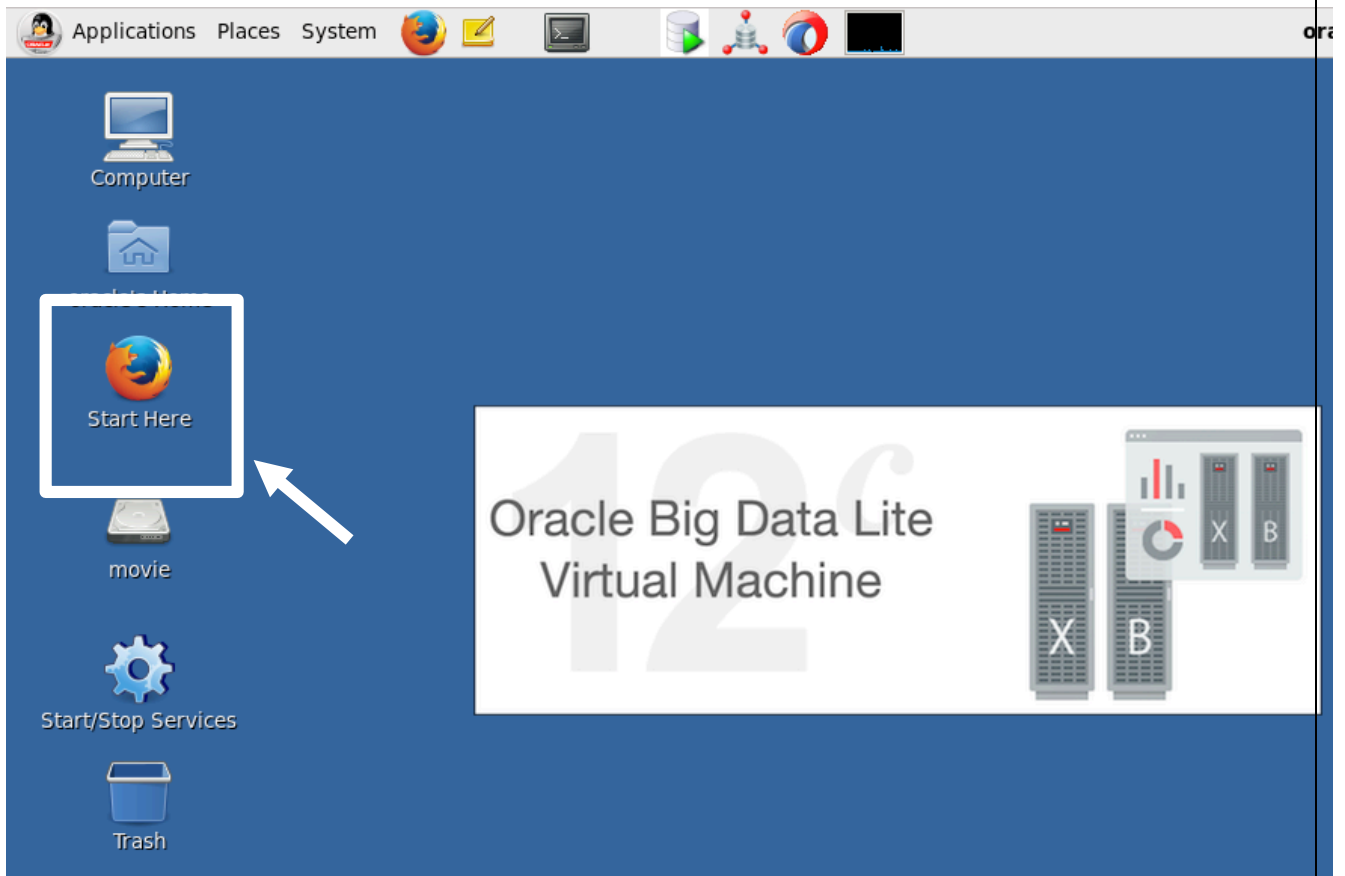
Big Data Lite Desktop

8. **Start/Stop Services**

   Use the Start/Stop Services script to start/stop your services (you can also type `services` at the linux prompt). **Services only appear if the minimum memory requirement is met.** Ensure that you start all services required for an application (e.g. Hive requires HDFS).

```
                     Start/Stop Oracle Big Data Lite Services

 Use the arrow keys to navigate between services. Use the spacebar to select services to start.
 Deselect services to stop them. Hit <enter> to accept the changes, <CTRL>-c to cancel.

 Note: Using Cloudera Manager and manually starting services are mutually exclusive; you must stop
 one to access the other. Also, please perfom a system reboot after shutting down Cloudera Manager.
 _____
 |  [*] ORCL                Oracle Database 12c(on)                                                  |
 |  [ ] Zookeeper           zookeeper(on)                                                             |
 |  [*] HDFS                namenode(on) secondarynamenode(on) hadoop-httpfs(on) datanode(on)         |
 |  [ ] HBase               master(off) regionserver(off) thrift(off)                                |
 |  [*] Hive                metastore(on) hive-server2(on)                                            |
 |  [*] Hue                 Hue(on)                                                                   |
 |  [ ] Impala              impala-server(off) catalog(off) state-store(off)                          |
 |  [ ] Kafka               kafka-server(off)                                                         |
 |  [*] NoSQL               Oracle NoSQL Database(on)                                                 |
 |  [ ] Oozie               oozie(off)                                                                |
 |  [ ] ORDS-Apex           Oracle Rest Data Service-Apex(off)                                        |
 |  [ ] PGX-Server          PGX Server(off)                                                           |
 |  [*] Solr                solr-server(on)                                                           |
 |  [ ] Spatial             bdsg(off)                                                                 |
 |  [ ] Sqoop2              sqoop2-server(off)                                                        |
 |  [ ] WebLogic-MovieDemo  WebLogic-MovieDemo(off)                                                   |
 |  [*] YARN                resourcemanager(on) nodemanager(on) history-server(on) spark-history(on) |
 |  [ ] RStudio             RStudio server(off)                                                       |
 |  [ ] Zeppelin            Zeppelin Notebooks(off)                                                   |
 |  [*] Jupyter             Jupyter Notebooks(on)                                                     |
 |_____|


                        <  OK  >                      <Cancel>
```

Starting/Stopping services

9. **Big Data Lite Samples**

   Big Data Lite samples are now available on github (https://github.com/oracle/BigDataLite). This makes
   it easy to update your VM with the latest collateral available. There is a branch for each Big Data Lite
   release – starting with Big Data Lite 4.3.0. Double-click the **Refresh Samples** icon and follow the in-
   structions to refresh your local copy of the samples with the hosted repository:

Hardware and Software
ORACLE
Engineered to Work Together
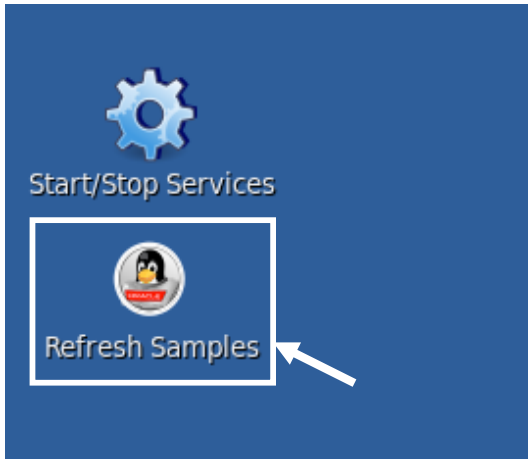
Refresh Samples

The samples are downloaded to the /home/oracle/src directory. **Note**: the contents in this directory will be deleted as part of the refresh process – so remember to save any customizations you made in this directory. Also, directory /home/oracle/movie is a bind mount that refers to directory /home/oracle/src/movie. This means that these two folders are equivalent; any changes to the contents in one folder is reflected in the other.

TROUBLESHOOTING TIPS

1. **Failed to import appliance <path>/BigDataLite-xxx.ova. Could not open the current file in the OVA package (VERR_TAR_CHKSUM_MISMATCH).**

    This error occurs when attempting to import an image. To solve the problem, simply update your version of VirtualBox.

2. **ERROR: "*Failed to open a session for the virtual machine*" when attempting to start the machine**

    The VM requires that hardware virtualization is enabled for the host computer. If it is not enabled, you will receive the following error when attempting to start the machine:

    Failed to open a session for the virtual machine BigDataLite-xxx.
    VT-x is disabled in the BIOS
    (VERR_VMX_MSR_VMXON_DISABLED)

    You can update the BIOS at boot time for the host. Ensure that both "Intel(R) Virtualization Technology" and "Intel(R) VT-d Feature" are enabled. Please review the instructions provided by your hardware manufacturer that describe updating BIOS settings.

3. **ERROR: "This kernel requires an x86-64 CPU, but only detected an i686 CPU. Unable to boot - please use a kernel appropriate for your CPU"**

    Ensure that your machine is running Oracle (64-bit) Linux. In the VM settings select "General" and set the Version to Oracle (64-bit).

General		
Basic	Advanced	Description
Name:	BigDataLite-2.4.1	
Type:	Linux	
Version:	Oracle (64 bit)	
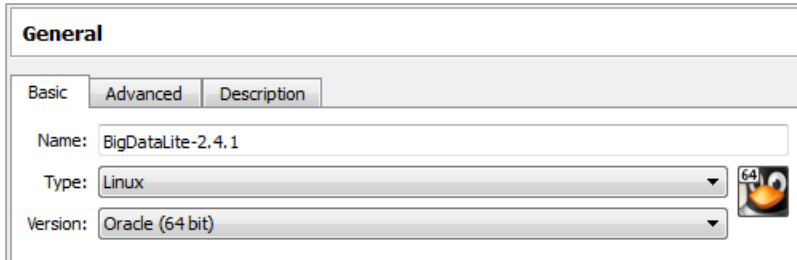
4. **ERROR: "Implementation of the USB 2.0 controller not found!"**

    Install the appropriate VirtualBox Extension Pack for your version of VirtualBox.

5. **While the image is starting up, you initially see a few messages and then a blank screen**

    If this happens, then stop the image. Go back to Virtual Box Manager and click on the image's **Set-tings**. Go to **Display** and increase the **Video memory** to at least 6MB. Start the image and you should no longer see the blank screen.

6. **Unable to copy/paste between the host and Big Data Lite**

    First, ensure that you have installed the Virtual Box Guest Additions. If you have installed Guest Additions and copy/paste still doesn't work – go to the **Devices** menu and ensure that **Shared Clipboard** is enabled.

7. **Incorrect JDK used by Hadoop. Note: this will not occur unless you add JDK 1.7 to the VM.**

    There are numerous errors that can occur because CDH is using JDK 1.7 instead of JDK 1.8 (see below). To use the correct JDK:
    - Open a terminal window by single-clicking the Terminal icon in the task bar.
    - Enter the following command:
        - sudo vi /etc/default/bigtop-utils
        - Add the following line at the end of the file:
   export JAVA_HOME=/usr/java/latest
    - Restart the VM to make sure that the updated settings are used.

    Many of these errors occur when ODI is attempting to run jobs on the VM. For example:
    - com.sunopsis.core.SnpsForeignKeyReferencedObjectMissingException: ODI-10244: Cannot import DataType Conversion : 949-62-184 (with global id: 7663ea6c-f7c3-4d0b-bd80-

3117c1e06868 and internal id: 949 | 62 | 184) because referenced Technology with global id 7ffd3917-2b27-35a3-96a7-3bde26098531 does not exist.

- Error: QueryResult : Unsupported major.minor version 52.0
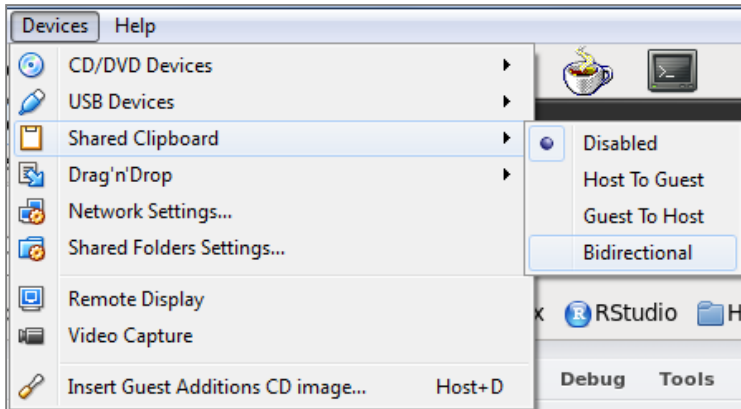
USEFUL TIPS

**1. Sharing folders between Big Data Lite and its host**

Virtual Box Guest Additions provides a useful feature that enables you to share a host's folder within the virtual machine. To configure a shared folder:

- Select BigDataLite-xxx from the list of machines and click **Settings**

- Go to **Shared Folders** and click **Add Folder**
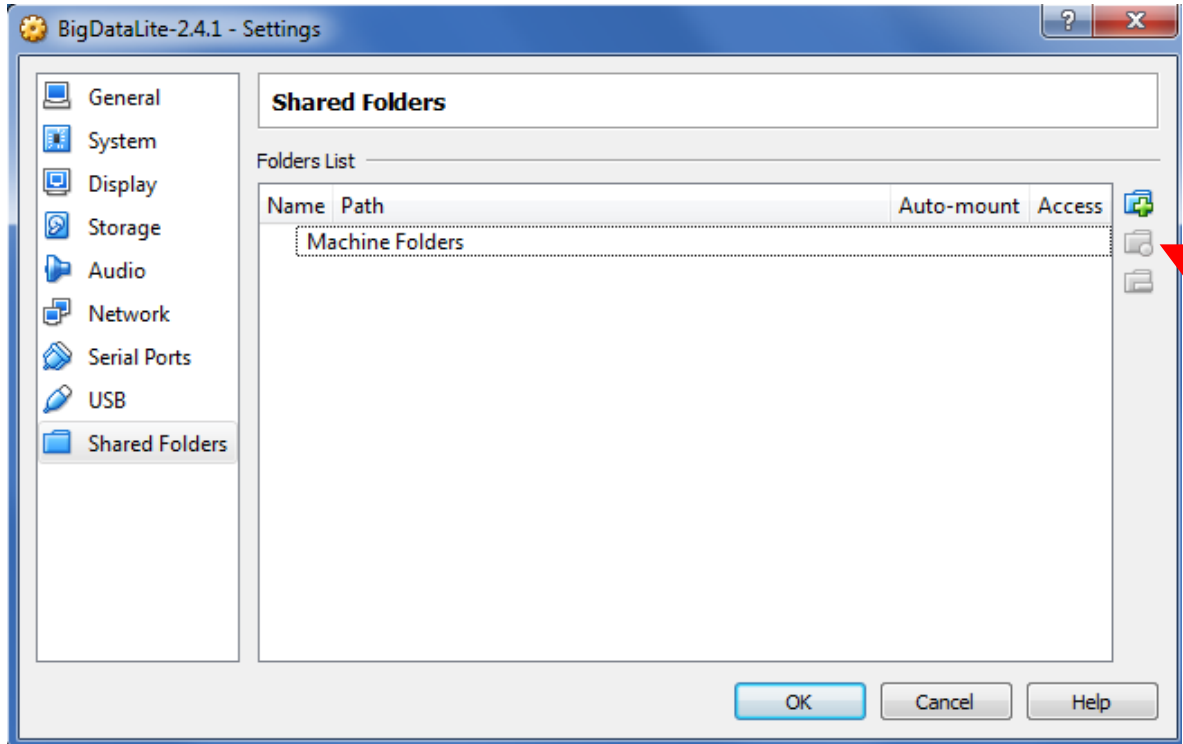
- Specify the folder you would like to share and select **Auto-mount**. Then click **OK**. Below, the H:\ drive is being shared.
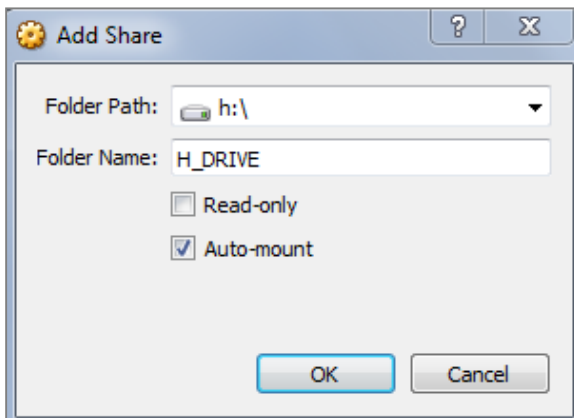
You can repeat this process for multiple shared folders.

- To access the shared folder from Big Data Lite, navigate to /media/sf_<folder_name>

2. **Using the VM as a Hadoop Client**

You can use Big Data Lite as a Hadoop client – allowing you to use it as a development environment for the Oracle Big Data Appliance. To connect the Big Data Lite to BDA, log into the Cloudera Manager instance that is managing your Hadoop cluster and select **Actions -> View Client Configuration URLs**:
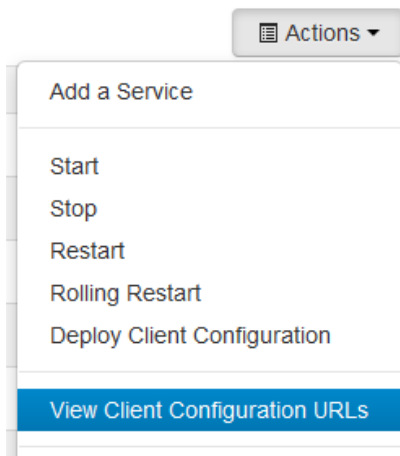
This will take you to the Client Configuration URLs page. Select the appropriate Client Configuration (e.g. MapReduce):

Client Configuration URLs   ✕

Name	Type	URL
📄 hdfs ▾	HDFS	⏱ /cmf/services/1/client-config
▦ mapreduce ▾	MapReduce	⏱ /cmf/services/2/client-config
🐝 hive ▾	Hive	⏱ /cmf/services/5/client-config
H hbase1 ▾	HBase	⏱ /cmf/services/9/client-config
🌅 solr1 ▾	Solr	⏱ /cmf/services/14/client-config

Close

This will download a zip file containing all the files required to connect to the cluster. After download-ing the file, execute the following two steps:

- Unzip the file into /home/oracle.
- Set the appropriate environment variable that will point Big Data Lite to the remote cluster. For example, if you downloaded both hive and mapreduce configurations:

```
export HADOOP_CONF_DIR=/home/oracle/hadoop-conf
export HIVE_CONF_DIR=/home/oracle/hive-conf
```

3. **Connecting external clients to Big Data Lite**

You may want to connect to Big Data Lite from your host computer. VirtualBox port forwarding has been implemented allowing you to connect the guest using a browser or SSH. You will connect to server "localhost:<port#>" and VirtualBox will forward the request to the guest. Listed below are the port mappings for services running in Big Data Lite:

Service	Port
NameNode	50070
Secondary NameNode	50090
ResourceManager	8088
NodeManager	8042
JobHistory Server	19888
Hue	8888
Oozie Server	11000
History Server	18088

Hardware and Software
ORACLE
Engineered to Work Together

Cloudea Manager	7180
Movie demo	7001
Solr Admin	8983
Oracle	1521
Apex	8080
SSH	2222
NoSQL DB Admin	5001
Big Data Spatial & Graph	8045

Hardware and Software
ORACLE
Engineered to Work Together