New Data Lake/Tutorial 1 ...

Tutorial 1 - Notebook Basics

Welcome to Oracle Big Data Cloud Service - Compute Edition (BDCS-CE).

This tutorial was built for BDCS-CE version 17.4.1 as part of the New Data Lake User Journey: here (https://github.com/oracle/learning-library /tree/master/workshops/journey2-new-data-lake). Questions and feedback about the tutorial: david.bayard@oracle.com (mailto:david.bayard@oracle.com)

Contents

- About Zeppelin and Interpreters
- About the Markdown Interpreter
- About the Shell Interpreter
- A few Zepellin tricks (printing, restarting, viewing logs)
- Next Steps

As a reminder, the documentation for BDCS-CE can be found here (https://docs.oracle.com/cloud/latest/big-data-compute-cloud/index.html)

Took 0 sec. Last updated by anonymous at November 15 2017, 1:47:00 PM.

About Zeppelin Notes and Paragraphs

BDCS-CE includes a built-in Zeppelin notebook. Currently, BDCS-CE is using Zeppelin version 0.7.0. You can learn more about Zeppelin here (https://zeppelin.apache.org/docs/0.7.0/).

To get started, please review the quick overview about Notes and how to interact with them here (https://zeppelin.apache.org/docs/0.7.0

/quickstart/explorezeppelinui.html#note-layout).

```
Make sure you have reviewed the above overview and know about Zeppelin Notes and Paragraphs and the UI controls to work with them befor
```

About Interpreters

The Zeppelin notebook is powered by Interpreters, which allow various technologies to be combined into a single, easily documented environment. The Interpreters currently offered with this version of BDCS-CE are

- Markdown (%md)
- Shell (%sh)
- Spark with Scala (%spark)
- Spark with Python (%pyspark)
- Spark SQL (%sql)
- Spark with R (%r)
- JDBC (%jdbc)

In this tutorial, we will introduce you to the Markdown and Shell Interpreters.

About the Markdown Interpreter (%md)

The Markdown interpreter is used to add documentation to your notebooks. To create a paragraph using the Markdown interpreter, use %md for the first line in the code editor.

Details about the markdown interpreter can be found here (https://zeppelin.apache.org/docs/0.7.0/interpreter/markdown.html)

Some simple examples...

Add Emphasis using one or more *

Create a list using +

- item

- item

- item

For more, check out the Markdown cheatsheet here (https://github.com/adam-p/markdown-here/wiki/Markdown-Cheatsheet)

Hint: Show the editor for this paragraph to see the actual Markdown code used

FINISHED ▷ ⛶ ▤ ⚙

About the Markdown Interpreter (%md)

The Markdown interpreter is used to add documentation to your notebooks. To create a paragraph using the Markdown interpreter, use %md for the first line in the code editor.

Details about the markdown interpreter can be found here: https://zeppelin.apache.org/docs/0.6.0/interpreter/markdown.html

Some simple examples…

Add Emphasis using one or more *

Create a list using +

- item
- item
- item

For more, check out the Markdown cheatsheet at https://github.com/adam-p/markdown-here/wiki/Markdown-Cheatsheet

Hint: Show the editor for this paragraph to see the actual Markdown code used

Took 0 seconds. Last updated by anonymous at time Jun 12, 2017 3:17:45 PM. (outdated)

READY ▷ ⛶ ▤ ⚙

Before you continue…

Before you continue, be sure you know how to work with Notes. Here is quick tutorial on working with the Note layout: https://zeppelin.apache.org/docs/0.6.0/quickstart/explorezeppelinui.html#note-layout

You will want to make sure you know how to:

Before you continue...   READY

Before you continue, be sure you know how to work with Notes. Here is quick tutorial on working with the Note layout: here (https://zeppelin.apache.org/docs/0.7.0/quickstart/explorezeppelinui.html#note-layout)

You will want to make sure you know how to:

- Show the code editor for a paragraph
- Run a paragraph

Now is a good time to **practice**. Go ahead and show the code editor for this paragraph and your name to the line below. Then run this paragraph to see the change.

Hello your_name_here

About the Shell Interpreter (%sh)

The shell interpreter runs linux shell commands as the zeppelin user on the BDCS-CE server running zepellin. It can be handy way to check zeppelin log files, pull files into BDCS-CE, interact with HDFS, or change/install software. SSH can also be used, but running shell commands via the notebook can be more natural and self-documenting.

The following paragraph has some sample shell command examples. Run the following paragraph to see the shell interpeter in action.

Example: Shell Interpreter commands

```
%sh
echo "running whoami"
whoami
echo ".."
echo "running pwd"
pwd
echo ".."
echo "running a hadoop command"
hadoop fs -ls /user
echo ".."
echo ".."
echo ".."
echo "done"
```

Before you continue...

Run the above paragraph. Did you notice that the above Shell example paragraph originally had no output? (At least it didn't when you first looked at it). Go ahead and run the above paragraph to execute the sample shell commands and see the output.

Next Steps

- Review the above links to learn more about Zeppelin and Markdown
- Refer to the below tips for some additional tricks on working with Zeppelin in BDCS-CE
- Do the BDCS-CE Working with Notebook tutorial from the documentation if you have not yet done it: here (http://www.oracle.com /webfolder/technetwork/tutorials/obe/cloud/bdcsce/notebook/notebook.html)
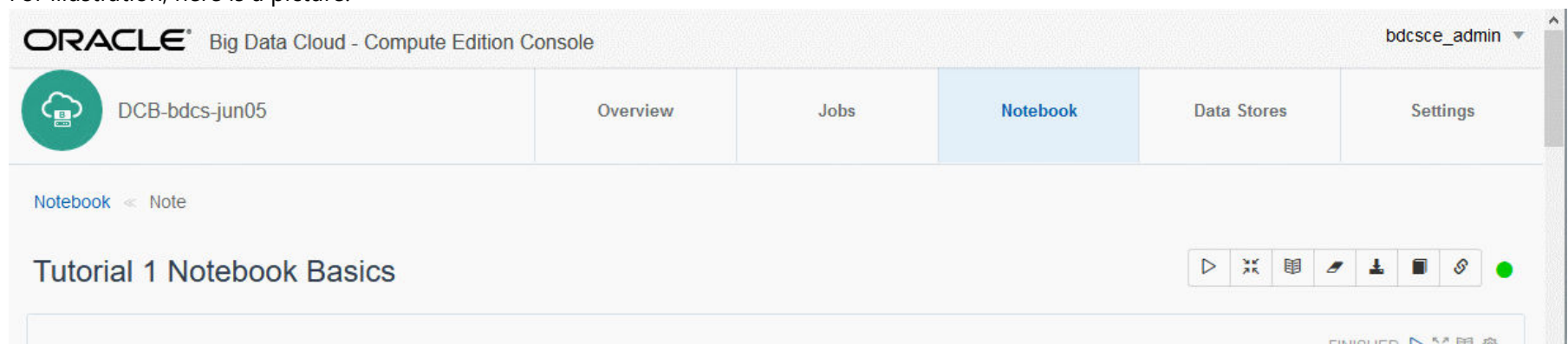- Move on to the next tutorial

Tip: How to Print Notebooks (Save as PDF)

Here is a trick to pretty print the Notebooks so that you can easily share them via email. Note: you can also export notebooks, but the recipient would need to import the notebook into another Zepellin instance to view it.

The trick is that the main body of the notebook is enclosed in an iframe. To best print the notebook, you want to just print the main iframe. For instance, using the Firefox browser, right-click on the open area between the Notebook title and the set of action icons. In the pop-up menu, then choose This Frame... Print Frame... Then choose a printer driver that can print to PDF.

For illustration, here is a picture:

Tutorial 1 - Notebook Basics

Welcome to Oracle Big Data Cloud Service - Compute Edition (BDCS-CE).

This tutorial was built for BDCS-CE version 17.2.5. If you are using a later version of BDCS-CE, please look for a newer version of this tutorial notebook in the New Data Lake section at https://github.com/oracle/learning-library/tree/master/workshops/journey2-new-data-lake. Questions and feedback about the tutorial: david.bayard@oracle.com

Contents

- About Zeppelin and Interpreters
- About the Markdown Interpreter
- About the Shell Interpreter
- A few Zepellin tricks (printing, restarting, viewing logs)
- Next Steps

As a reminder, the documentation for BDCS-CE can be found here: http://docs.oracle.com/cloud/latest/big-data-compute-cloud/index.html

Took 0 seconds. Last updated by anonymous at time Jun 9, 2017 4:36:38 PM. (outdated)

READY ▷ ⤢ ▦ ⚙

About Zeppelin and Interpreters

BDCS-CE includes a built-in Zeppelin notebook. Currently, BDCS-CE is using Zeppelin version 0.6.0. You can more about Zeppelin here https://zeppelin.apache.org/docs/0.6.0/

Tip: How (and when) to Restart Zeppelin Notebooks   READY

From time to time, you may want to restart the Zepellin Notebook process. To do so,

- Click on the Settings tab at the top of the browser
- Then, click on Notebook on the left-hand column
- Then click Restart

Reasons why you may want to restart the Zeppelin process.

- You changed a configuration and want to pick up the new value. For instance, once you set certain swiftfs parameters (as would be needed to reference the Oracle Storage Cloud Object Store), these values will be used for the rest of the session. You restart Zeppelin to pick up the new value.

Tip: How to look at the Zeppelin log

The zeppelin log is at /u01/bdcsce/data/var/log/zeppelin

The following paragraph has some examples showing how to review the logs.

Example: Viewing the Zeppelin Log File

FINISHED

```
%sh
cd /u01/bdcsce/data/var/log/zeppelin
ls -ltr
echo ".."
echo ".."
tail -20 zeppelin-zeppelin-*.log
```

Took 0 sec. Last updated by anonymous at November 15 2017, 1:47:22 PM. (outdated)

READY