BIRT Connectivity Tutorial

V1.2
December 12, 2013

Introduction

Summary

This tutorial will demonstrate how to configure BIRT (**B**usiness **I**ntelligence and **R**eporting **T**ools) to access data from the Hortonworks Sandbox and its integrated Hive Server. BIRT is an open source Eclipse-based reporting system that integrates with your Java/Java EE application to produce compelling reports. BIRT provides core reporting features such as report layout, data access and scripting.

Introductory

This tutorial is a prerequisite tutorial for the "BIRT Integration and Report" tutorial.

Prerequisites:

To complete this tutorial, and the "BIRT Integration and Report" tutorial you will need to install Eclipse by downloading and extracting the open source Eclipsed-based BIRT version 4.3.2 project at the following link.

- http://www.birt-exchange.org/get-birt/432

Overview

Since this tutorial is designed to simply configure connectivity between BIRT and the Hortonworks Sandbox, it is relatively short with very few steps.

1. Download and install the jar's required by the ODA
2. Test the ODA connectivity by:
    a. Creating a Data Source
    b. Creating a Data Set
    c. Displaying the data

Step 1 – Download Dependent Jars from Apache

In this section we will download the dependent jars from Apache and copy them to the appropriate directories under your eclipse installation.

1. Download an extract Apache Hive version 0.12.0 from

   http://hive.apache.org/releases.html

2. Download and extract hadoop-core-1.0.3.jar from
   http://hadoop.apache.org/releases.html

3. To install we simply copy the following jar's

```
commons-logging-1.1.1.jar
hadoop-core-1.0.3.jar
hive-exec-0.12.0.jar
hive-jdbc-0.12.0.jar
hive-metastore-0.12.0.jar
hive-service-0.12.0.jar
httpcore-4.2.4.jar
httpclient-4.2.5.jar
libfb303-0.9.0.jar
slf4j-api-1.6.1.jar
```

slf4j-log4j12-1.6.1.jar
log4j-1.2.16.jar

To "`plugins\org.eclipse.birt.report.data.oda.jdbc_`*`version-number`*" directory.

e.g.

- "`c:\eclipse\plugins\org.eclipse.birt.report.data.oda.jdbc_4.3.2.v20131024-0257\drivers`" (Windows)
- "`/Applications/eclipse/plugins/org.eclipse.birt.report.data.oda.jdbc_ 4.3.2.v20131024-0257/drivers`" (Mac)

If you are connecting to Hortonworks HDP version 1.3 use 0.11.0 instead of 0.12.0

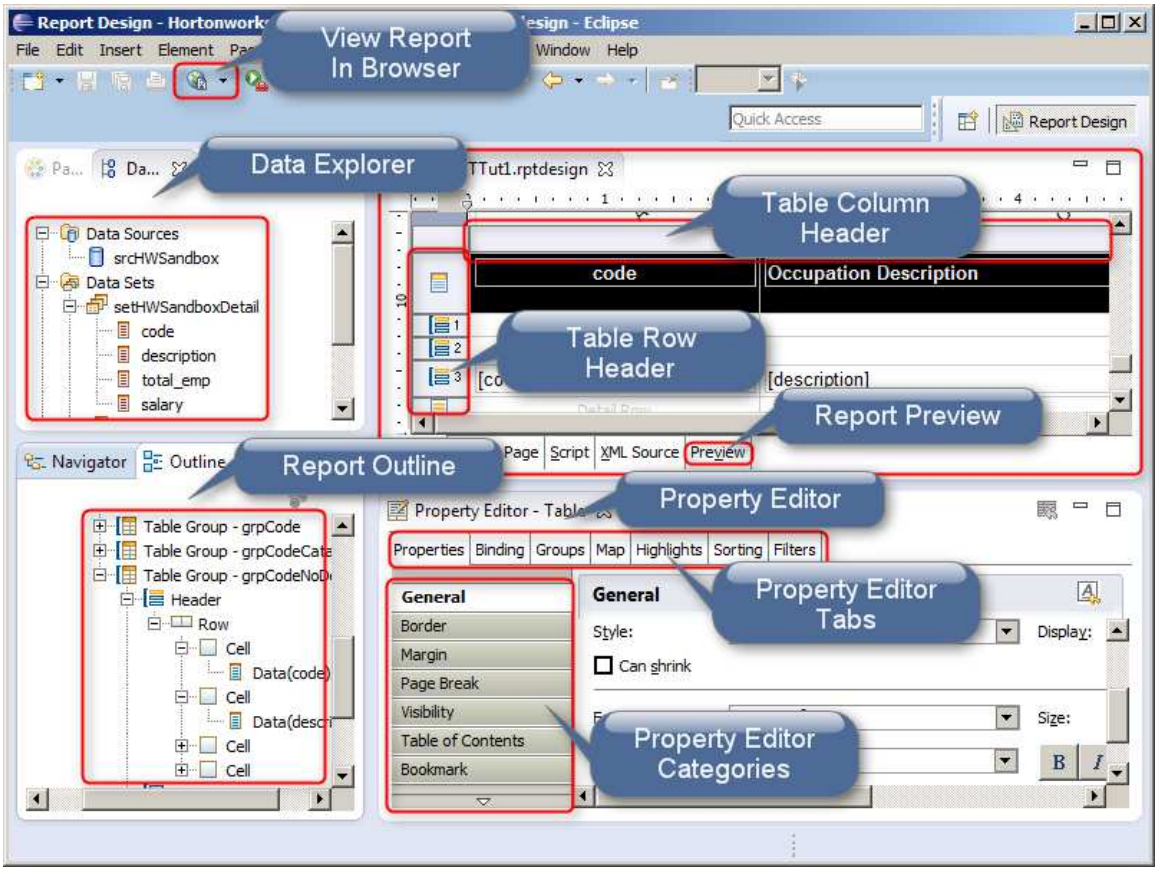
Now that this section is complete, we have the required configuration to access the Hortonworks Sandbox from BIRT.

Step 2 – Test BIRT / Hortonworks Sandbox Connectivity

In this section, we will test that we can use BIRT to report from data located in the Hive Server on the Hortonworks Sandbox.

1. Create a BIRT project
2. Create a new BIRT report
3. Create a Hive Data Source
4. Create a Data Set using the Hive Data Source
5. Display the data

The following image is an overview of the Eclipse BIRT IDE with callouts identifying some of the major IDE components.



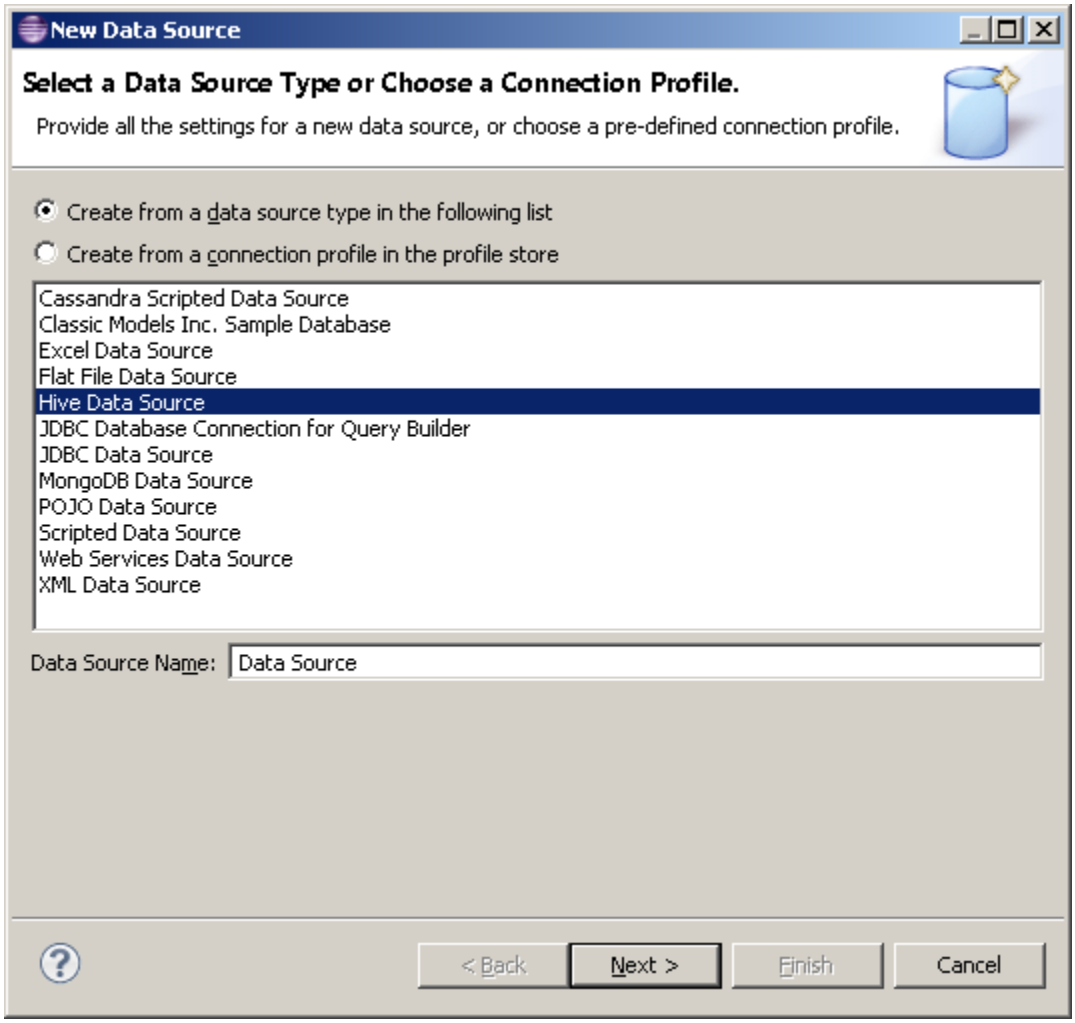
To test the connectivity to the Hortonworks Sandbox, we'll first need to create a BIRT reporting project and a BIRT report in eclipse.

1. To create a new BIRT project, select **File > New > Project...**, in the displayed dialog, expand the "Business Intelligence and Reporting Tools" section, click "Report Project", then click **Next**. Give the project a name, e.g. "Test", click **Finish**.

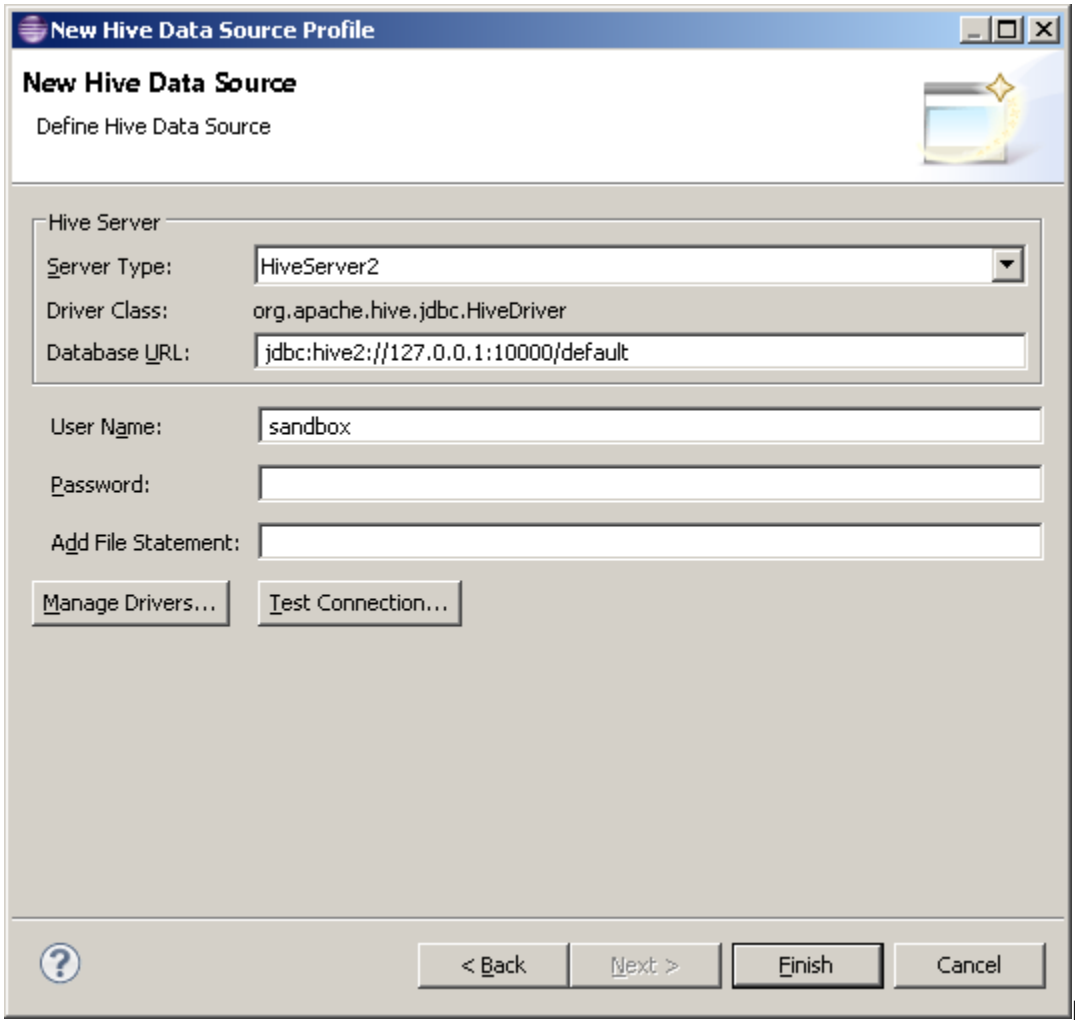
2. To create a new BIRT report, select **File > New > Report**. Give the report a name, e.g. "BIRT+Hortonworks", click **Finish**.

Next we need to create the Data Source that will connect to the Hortonworks Sandbox.

To create our Hive Data Source, navigate to the **Data Explorer** tab, right click **Data Sources**, and select **New Data Source**. In the displayed dialog select "Hive Data Source" and click **Next >** as shown in the following image.



In the following dialog specify the **Database URL** using the IP address displayed in the Hortonworks Sandbox virtual machine. In my case it is "127.0.0.1". Leave the rest of the URL alone as it should already be correct. Enter "Sandbox" as the **User Name**, and click **Test Connection...**. You should see a "Connection Successful" confirmation. Click **Finish**.

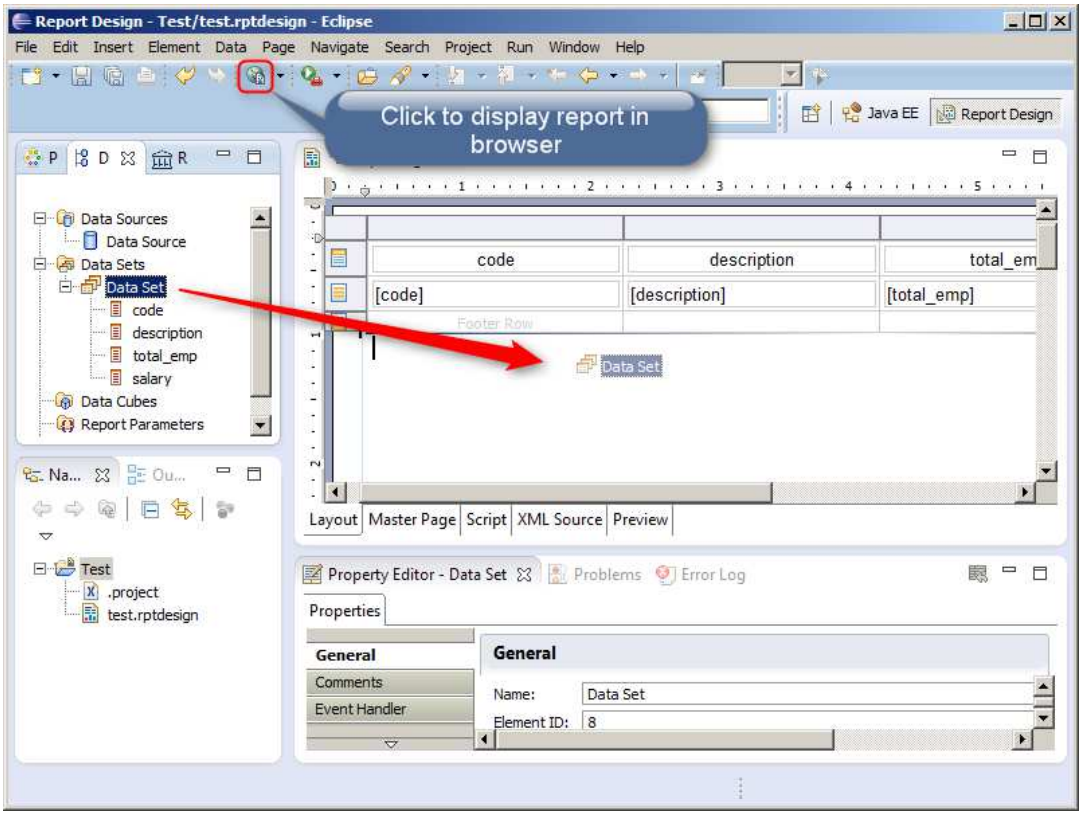


Next we need to create the Data Set that will retrieve data from the Hortonworks Sandbox's Hive Server.

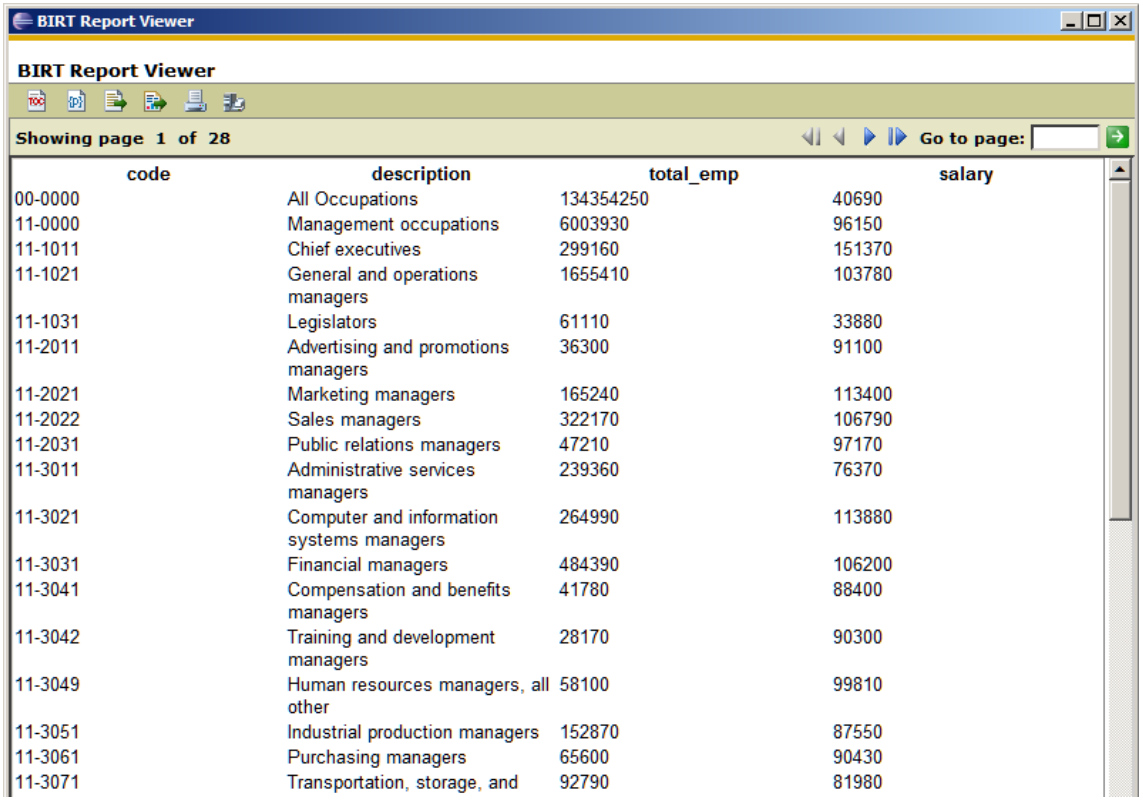
To create our Data Set, navigate to the **Data Explorer** tab, right click **Data Sets**, and select **New Data Set**. In the displayed dialog click **Next >**. In the next dialog enter "`select * from sample_07`" as the **Query Text** and click **Finish**. In the next displayed dialog click **OK**.

Now we'll simply display the data in a quick raw report.

Click and drag the Data Set you've just created and drop it on the report canvas as shown in the following image. Preview the report data in your browser by clicking the **Browser Preview** icon in the eclipse toolbar.



After clicking the **Browser Preview** icon you should see the following.



code	description	total_emp	salary
00-0000	All Occupations	134354250	40690
11-0000	Management occupations	6003930	96150
11-1011	Chief executives	299160	151370
11-1021	General and operations managers	1655410	103780
11-1031	Legislators	61110	33880
11-2011	Advertising and promotions managers	36300	91100
11-2021	Marketing managers	165240	113400
11-2022	Sales managers	322170	106790
11-2031	Public relations managers	47210	97170
11-3011	Administrative services managers	239360	76370
11-3021	Computer and information systems managers	264990	113880
11-3031	Financial managers	484390	106200
11-3041	Compensation and benefits managers	41780	88400
11-3042	Training and development managers	28170	90300
11-3049	Human resources managers, all other	58100	99810
11-3051	Industrial production managers	152870	87550
11-3061	Purchasing managers	65600	90430
11-3071	Transportation, storage, and	92790	81980

Now that you have successfully performed all of the steps in the tutorial, you have the foundation to complete the "BIRT Integration and Report" tutorial as well as developing your own BIRT reports which access the Hortonworks Sandbox.

Survey

Thank you for downloading and working through this tutorial. We're interested in your feedback. Please take a minute to answer a few questions in this survey so that we can continue to improve.

About Actuate - The BIRT Company™

Actuate founded and co-leads the BIRT open source project, which is used by more than 2.5 million developers around the globe and serves as the foundation of Actuate's commercial offerings. Applications built with BIRT deliver more business and consumer insights to more people than all BI companies combined - ensuring organizations are ready for the exponential growth of Big Data and the proliferation of touch devices. Actuate's BIRT empower developers to rapidly develop custom, BIRT-based business analytics and customer communications applications. Applications built with one BIRT design can access and integrate any data, including unstructured sources. Headquartered in Silicon Valley, Actuate has over 5,000 customers globally in a diverse range of business areas including financial services, technology and the government. Actuate is listed on NASDAQ under the symbol BIRT. For more information, visit www.actuate.com or engage with the BIRT community at www.birt-exchange.com.

About Hortonworks

Hortonworks develops, distributes and supports the only 100-percent open source distribution of Apache Hadoop explicitly architected, built and tested for enterprise grade deployments. Developed by the original architects, builders and operators of Hadoop, Hortonworks stewards the core and delivers the critical services required by the enterprise to reliably and effectively run Hadoop at scale. Our distribution, Hortonworks Data Platform, provides an open and stable foundation for enterprises and a growing ecosystem to build and deploy big data solutions. Hortonworks also provides unmatched technical support, training and certification programs. For more information, visit www.hortonworks.com. The Hortonworks Sandbox can be found at: www.hortonworks.com/Sandbox.