

HIVE/HCATALOG/PIG LABS

Overview: These labs are the first five tutorials in the Hortonworks Sandbox VM. You won't finish them all in the allotted hour, so do your best to get started and you may finish them at a later time if you wish. After this hour we'll be shutting down the Hortonworks Sandbox VM and starting up the Cloudera Quickstart VM for later labs. Fear not: your VirtualBox will not lose your progress!

Bonus: One thing not covered in the tutorials that's kind of fun is the WebHCat/Templeton REST service. Why two names? It used to be called Templeton (after the mouse in Charlotte's web) but someone changed the name to WebHCat. For backwards-compatibility's sake, the name Templeton is still all over URIs and APIs. Naming things is hard; changing names is harder.

From a browser, test that the service is running at:
<http://localhost:50111/templeton/v1/status>

You can get a list of databases used by HCatalog/Hive metastore:
<http://localhost:50111/templeton/v1/ddl/database?user.name=me>

You can get a list of tables on the default database:
<http://localhost:50111/templeton/v1/ddl/database/default/table?user.name=me>

When you make your new tables in the tutorials do they show up in the WebHCat REST API?

WebHCat can do a lot more than just this. You can actually run Hive and Pig scripts using this REST API!