How to run AsterixDB locally on your machine

1. Preparation

You need to check the Java Development Kit (JDK) installation on your machine. You must upgrade your JDK to the latest version (**1.8**) to run AsterixDB.

For Windows users:

a) Start the Command Prompt (How).

b) Execute ‘java -version’.

c) If you see ‘java version xxxx’ that means you have Java properly installed.

d) If you see ‘java is not recognized as xxxx’, that means you don’t have Java installed. Check this to install Java and try again.

For OS X users:

a) Start your terminal.

b) Execute ‘java -version’.

c) If you see ‘java version xxxx’ that means you have Java properly installed.

d) If you see ‘command not found java’, that means you don’t have Java installed. Check this to install Java and try again.

For Linux users:

a) Start your terminal.

b) Execute ‘java -version’ to make sure Java is properly installed.

c) If you didn’t install the Java environment, follow the link to install it.

d) Execute ‘echo $JAVA_HOME’.

e) If you see a blank line, that means your environment variable is not set properly. Execute the following command to set it:

```
export JAVA_HOME=PATH_TO_YOUR_JAVA_INSTALLATION
```

f) You can put this command into your configuration file (e.g., .bashrc) to set the path variable automatically every time you login.

g) Rerun the command in b) and d) to make sure everything is properly configured.

h) openJDK is also OK, but we highly recommend the Oracle JDK.

2. Download an AsterixDB sample cluster

Download the right version of AsterixDB [Link]. This link will redirect you to an Apache website. You may choose any of the mirrors to download. The filename is **asterix-server-0.9.4.1-binary-assembly.zip**. The SHA1sum of the binary release is

65e1baa60da0144a4ea2b6ceb2b6a8f456ed0cc6

You may compare this against the checksum of the downloaded file to make sure that you have pulled the right version [Windows, OSX, Linux]. (This step can be **skipped** if you are certain about your package.)

3. Start an AsterixDB sample cluster

After you download the package, move it to your preferred working directory and unzip it. Please check the download folder if you can’t find it. You should see the following structure once you’re done unzipping the package:

```
asterix-server-0.9.4.1-SNAPSHOT-binary-assembly
├── LICENSE
├── NOTICE
├── bin
├── etc
├── lib
├── opt
└── repo
```

The scripts used for starting a sample AsterixDB cluster reside in the **opt/local/bin** directory. You should see 4 scripts files under this (**opt/local/bin**) directory:

```
bin
├── start-sample-cluster.bat
├── start-sample-cluster.sh
├── stop-sample-cluster.bat
└── stop-sample-cluster.sh
```

For Windows Users:

a) [**Recommended**] Start the Windows Command Prompt and navigate to asterix-server-0.9.4.1-SNAPSHOT-binary-assembly/opt/local/bin under your working directory.

Execute: start-sample-cluster.**bat**

For example, if you downloaded this file and uncompressed it in Downloads folder, you may want to execute the following command. Here, XXX is your username.

```
cd c:\Users\XXX\Downloads\asterix-server-0.9.4.1-SNAPSHOT-binary-assembly\opt\local\bin
start-sample-cluster.bat
```

b) You can also navigate to this directory using your file browser and double click on start-sample-cluster.bat to start the cluster.

c) By executing the stop-sample-cluster script in the same way, you can shutdown the sample cluster.

Execute: stop-sample-cluster.**bat**.

For Linux/OSX Users:

a) Start your terminal.

b) Navigate to asterix-server-0.9.4.1-SNAPSHOT-binary-assembly/opt/local/bin under your working directory.

c) Execute start-sample-cluster.sh using the following command:

```
bash start-sample-cluster.sh
```

Whichever way you start the sample cluster, if everything works properly, you should see the following message printed.

```
INFO: Starting sample cluster... INFO: Waiting up to 30 seconds for cluster
127.0.0.1:19002 to be available. INFO: Cluster started and is ACTIVE.
```

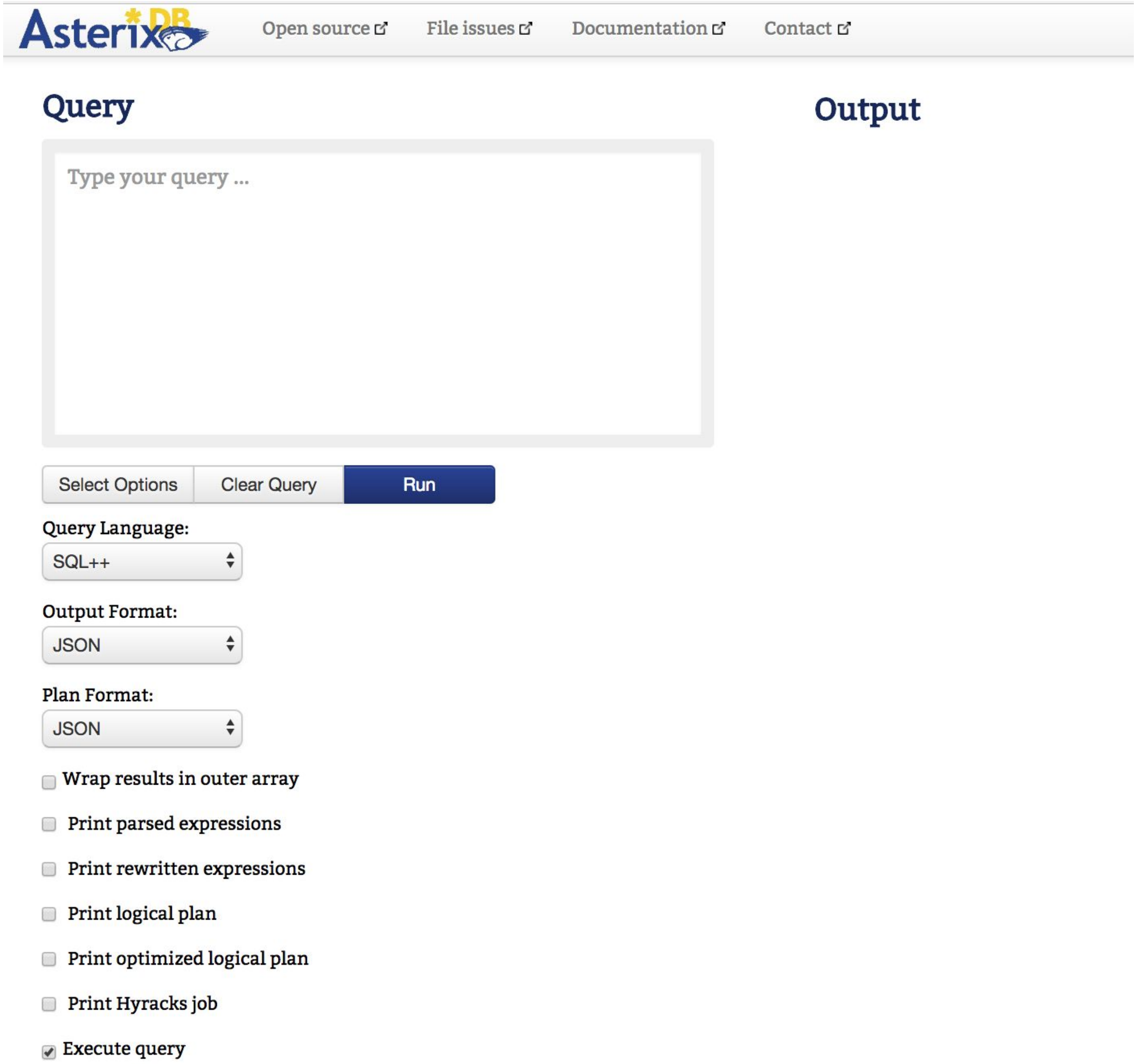
By executing the stop-sample-cluster script in the same way, you can shutdown the sample cluster.

```
bash stop-sample-cluster.sh
```

Note that AsterixDB is designed to run on parallel hardware; the sample cluster is a virtual cluster with two nodes.

4. Access the AsterixDB query interface

After you have started the local cluster, you can access the query interface from your browser at: http://localhost:19001. Our query interface looks like this:



Query Language: SQL++

Output Format: JSON

Plan Format: JSON

□ Wrap results in outer array

□ Print parsed expressions

□ Print rewritten expressions

□ Print logical plan

□ Print optimized logical plan

□ Print Hyracks job

☑ Execute query

Type the following query into the query box on the left and click the `Run` button to execute

```
SELECT * FROM Metadata.`Dataset`;
```

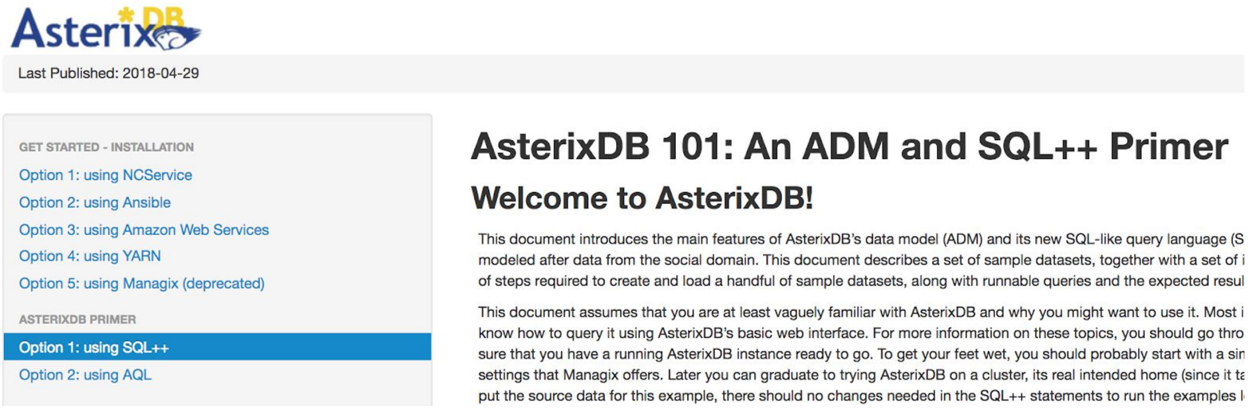
You will see the query result in the blank area on the right side. The default query language for AsterixDB is SQL++ (our newer query language).

5. Enjoy the ride!

If you are here in the instructions, now you have a mini-cluster running an AsterixDB instance on your machine. There are a lot more interesting features and examples to explore the Apache AsterixDB website. **Especially, you must read and execute all of the materials covered in The SQL++ Primer before starting the actual homework**. You may also wish to look at The SQL++ Query Language and Builtin Functions while doing the homework problems. We also uploaded the dataset in SQL++ Primer to the class web-page using Insert Statements. Please give them a try and enjoy!

1) AsterixDB 101: SQL++ Primer

https://ci.apache.org/projects/asterixdb/sqlpp/primer-sqlpp.html


Last Published: 2018-04-29
GET STARTED - INSTALLATION
Option 1: using NCService
Option 2: using Ansible
Option 3: using Amazon Web Services
Option 4: using YARN
Option 5: using Managix (deprecated)
ASTERIXDB PRIMER
Option 1: using SQL++
Option 2: using AQL

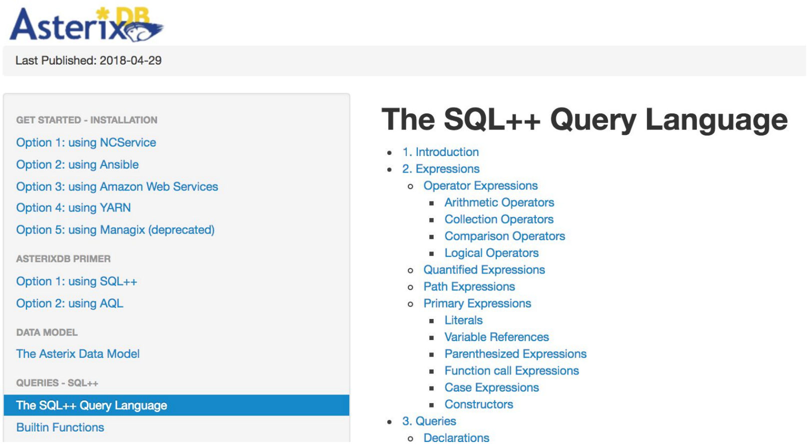
AsterixDB 101: An ADM and SQL++ Primer

Welcome to AsterixDB!

This document introduces the main features of AsterixDB's data model (ADM) and its new SQL-like query language (S
modeled after data from the social domain. This document describes a set of sample datasets, together with a set of i
of steps required to create and load a handful of sample datasets, along with runnable queries and the expected resul

This document assumes that you are at least vaguely familiar with AsterixDB and why you might want to use it. Most i
know how to query it using AsterixDB's basic web interface. For more information on these topics, you should go thro
sure that you have a running AsterixDB instance ready to go. To get your feet wet, you should probably start with a sin
settings that Managix offers. Later you can graduate to trying AsterixDB on a cluster, its real intended home (since it ta
put the source data for this example, there should no changes needed in the SQL++ statements to run the examples i

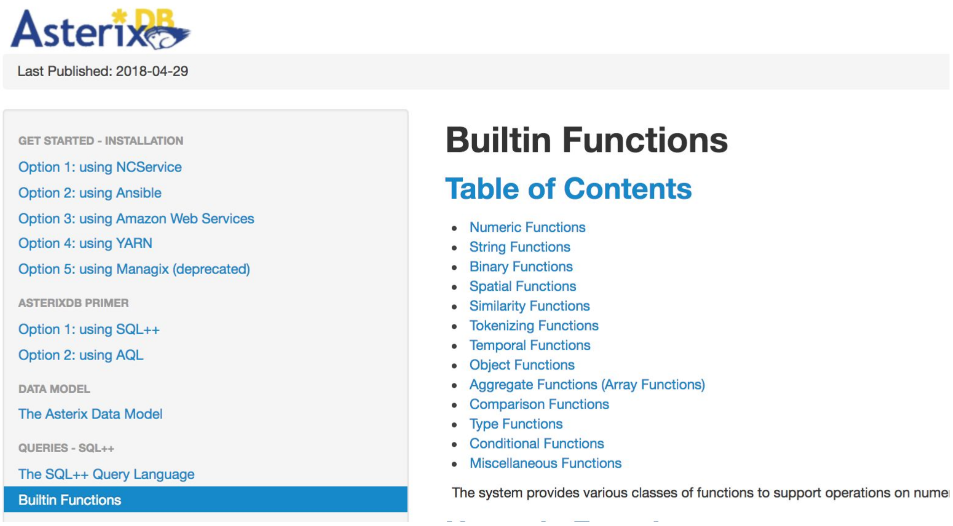
2) The SQL++ Query Language
https://ci.apache.org/projects/asterixdb/sqlpp/manual.html



The SQL++ Query Language

3) Builtin Functions

https://ci.apache.org/projects/asterixdb/sqlpp/builtins.html



Last Published: 2018-04-29

GET STARTED - INSTALLATION

Option 1: using NCService

Option 2: using Ansible

Option 3: using Amazon Web Services

Option 4: using YARN

Option 5: using Managix (deprecated)

ASTERIXDB PRIMER

Option 1: using SQL++

Option 2: using AQL

DATA MODEL

The Asterix Data Model

QUERIES - SQL++

The SQL++ Query Language

Builtin Functions

Builtin Functions

Table of Contents

The system provides various classes of functions to support operations on nume