

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 ensure that your JDK is sufficiently recent (version **11** or later) in order to run AsterixDB. To do the required check:

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
```

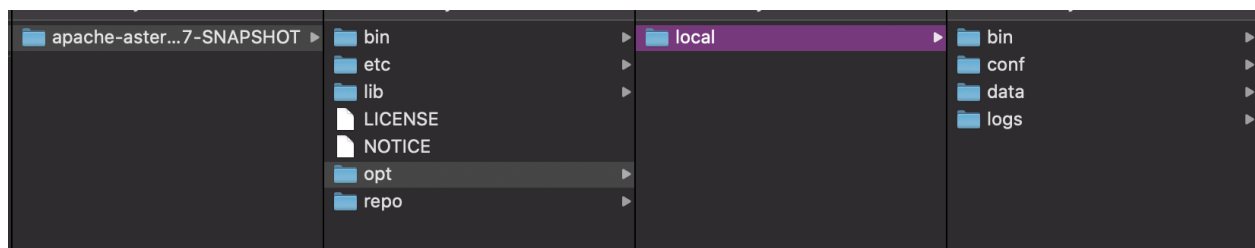
- f) You can put this command into your configuration file (e.g., .bashrc) to set the path variable automatically every time you login.
- g) Rerun the command in b) and d) to make sure everything is properly configured. openJDK and Oracle JDK are both OK to use.

2. Download AsterixDB

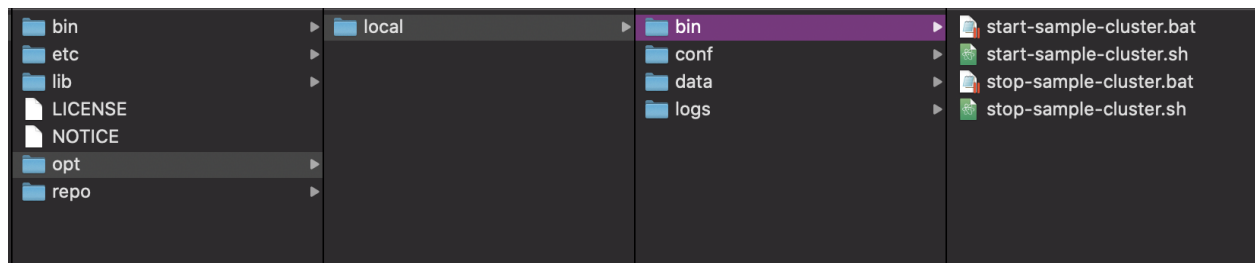
Download AsterixDB (SNAPSHOT of version 0.9.7) from the link here: [\[Link\]](#). This Drive link is a zip file of the AsterixDB instance you will be working on. **Do not download from other sources!**

3. Start AsterixDB

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 have unzipped the package:



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



For Windows Users:

a) **[Recommended]** Start the Windows Command Prompt and navigate to `apache-asterixdb-0.9.7-SNAPSHOT/opt/local/bin` under your working directory. Execute the command: `start-sample-cluster.bat`

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

```
cd c:\Users\XXX\Downloads\apache-asterixdb-0.9.6-SNAPSHOT\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 the command: `stop-sample-cluster.bat`.

For Linux/OSX Users:

- a) Start your terminal.
- b) Navigate to `apache-asterixdb-0.9.7-SNAPSHOT/opt/local/bin` under your working directory.
- c) Execute `start-sample-cluster.sh` using the following command:

```
./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 90 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.

```
./stop-sample-cluster.sh
```

Note that AsterixDB is designed to run on parallel hardware; however, the sample cluster for this homework has just one node.

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:19006>. Its query interface looks like this:

AsterixDB

WEBSITE FILE ISSUES DOCUMENTATION CONTACT GITHUB

QUERY INPUT (1/1)

Default PLAN FORMAT OUTPUT FORMAT QUERY HISTORY

JSON JSON

1

WARNINGS(0)

CLEAR EXPLAIN

METADATA INSPECTOR

DATAVERSES

☐ Default

☐ Metadata

☐ swoosh_schema

☐ swoosh_schemaless

☐ swoosh_test

REFRESH

DATASETS

DATATYPES

INDEX

USER DEFINED FUNCTIONS

Type the following query into the query box and click the `Run` icon (which looks like a play button) to execute it:

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

You will see the query result in the area underneath the query box.

5. Enjoy the ride!

If you are here in the instructions, now you will have an AsterixDB instance running on your machine. There are many more interesting features and examples to explore the Apache AsterixDB website. **NOTE: You must read and execute all of the materials covered in The SQL++ Primer before starting the actual homework.** Do not attempt the homework or ask questions about SQL++ on Piazza before you have finished that step! You may also wish to refer to [The SQL++ Query Language](#) and [Builtin Functions](#) while doing the homework problems. We have also attached the [dataset](#) in the [SQL++ Primer](#) to the class wiki page as a script 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 NCServe](#)
- [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 (SQL++). This document describes a set of sample datasets, together with a set of steps required to create and load a handful of sample datasets, along with runnable queries and the expected results.

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

2) The SQL++ Query Language

<https://ci.apache.org/projects/asterixdb/sqlpp/manual.html>



Last Published: 2018-04-29

GET STARTED - INSTALLATION

- [Option 1: using NCServe](#)
- [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](#)

The SQL++ Query Language

- 1. Introduction
- 2. Expressions
 - Operator Expressions
 - Arithmetic Operators
 - Collection Operators
 - Comparison Operators
 - Logical Operators
 - Quantified Expressions
 - Path Expressions
 - Primary Expressions
 - Literals
 - Variable References
 - Parenthesized Expressions
 - Function call Expressions
 - Case Expressions
 - Constructors
- 3. Queries
 - Declarations

3) Builtin Functions

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

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

- Numeric Functions
- String Functions
- Binary Functions
- Spatial Functions
- Similarity Functions
- Tokenizing Functions
- Temporal Functions
- Object Functions
- Aggregate Functions (Array Functions)
- Comparison Functions
- Type Functions
- Conditional Functions
- Miscellaneous Functions

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

3) SQL++ for SQL Users (Don Chamberlin)

https://asterixdb.apache.org/files/SQL_Book.pdf

Syllabus

Topic	Reading
Databases and DB Systems	Ch. 1
Entity-Relationship (E-R) Data Model	Ch. 2.1-2.5, 2.8
Relational Data Model	Ch. 3.1-3.2
E-R to Relational Translation	3.5
Relational Design Theory	Ch. 19.1-19.6, 20.8
<i>Midterm Exam 1</i>	<i>Fri, Oct 30</i> (during lecture time)
Relational Algebra	Ch. 4.1-4.2
Relational Calculus	Ch. 4.3-4.4
SQL Basics (SPJ and Nested Queries)	Ch. 3.4, 5.1-5.3
SQL Analytics (Aggregation, Nulls, and Outer Joins)	Ch. 5.4-5.6
Advanced SQL Goodies (Constraints, Triggers, Views, and Security)	Ch. 3.3, 3.6, 5.7-5.9, 21.1-21.3, 21.7
<i>Midterm Exam 2</i>	<i>Mon, Nov 23</i> (during lecture time)
Tree-Based Indexing	Ch. 9.1, 8.1-8.3, 10.1-10.2
Hash-Based Indexing	Ch. 10.3-10.8, 11.1
Physical DB Design	Ch. 8.5, 20.1-20.7
Semistructured Data Management (<i>a.k.a.</i> NoSQL)	⇒ AsterixDB SQL++ Prime, ⇒ Couchbase SQL++ Book
Basics of Transactions	Ch. 16 and Lecture Notes
<i>Endterm Exam</i>	<i>Wed, Dec 16</i> (during final exam time)