

# Getting Started with Basic Queries in Hive

---



**Janani Ravi**

CO-FOUNDER, LOONYYCORN

[www.loonycorn.com](http://www.loonycorn.com)

# Overview

**Set up Hive on your local system**

**Understand the basic building blocks of the Hive CLI**

**Run simple queries in Hive**

**Understand the basic Hive data types**

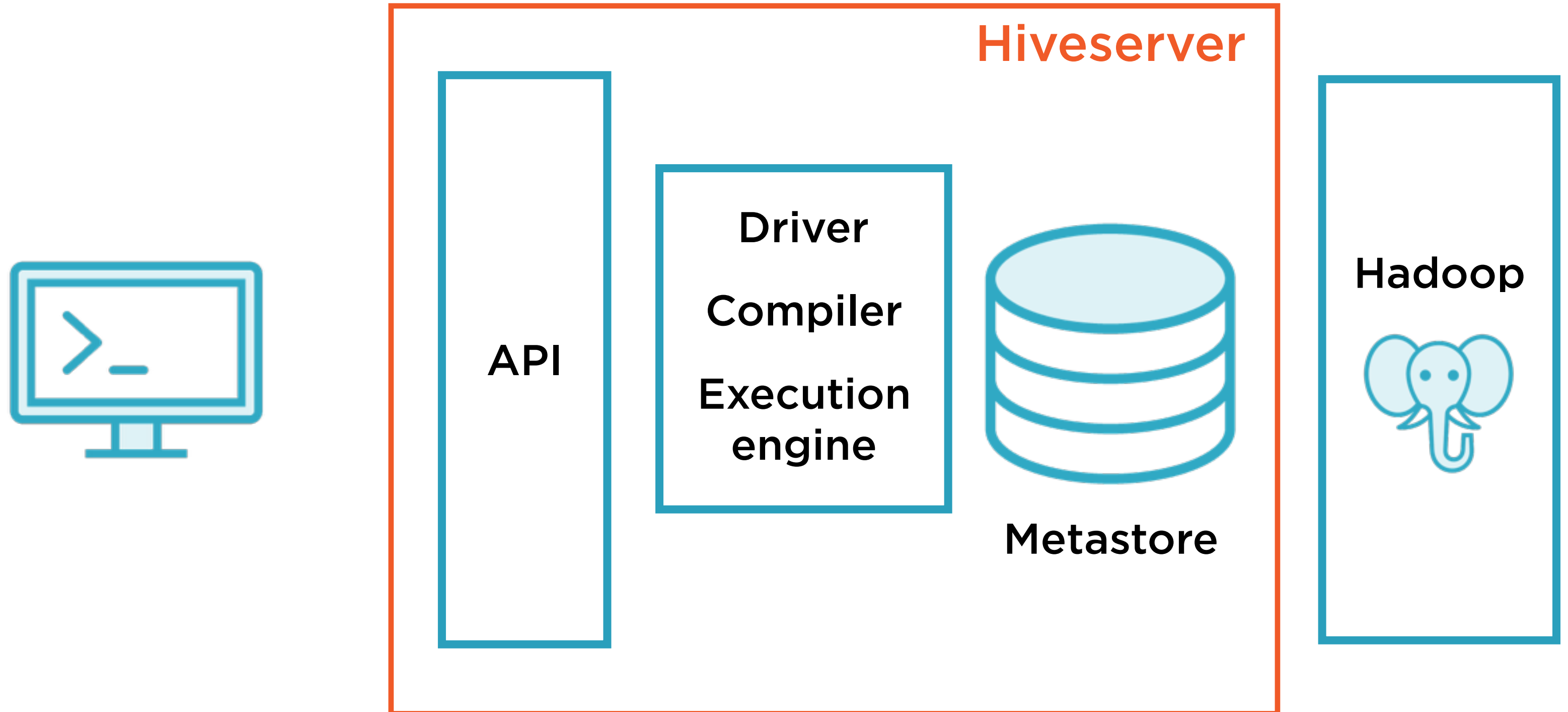
# Basic Hive Building Blocks

---

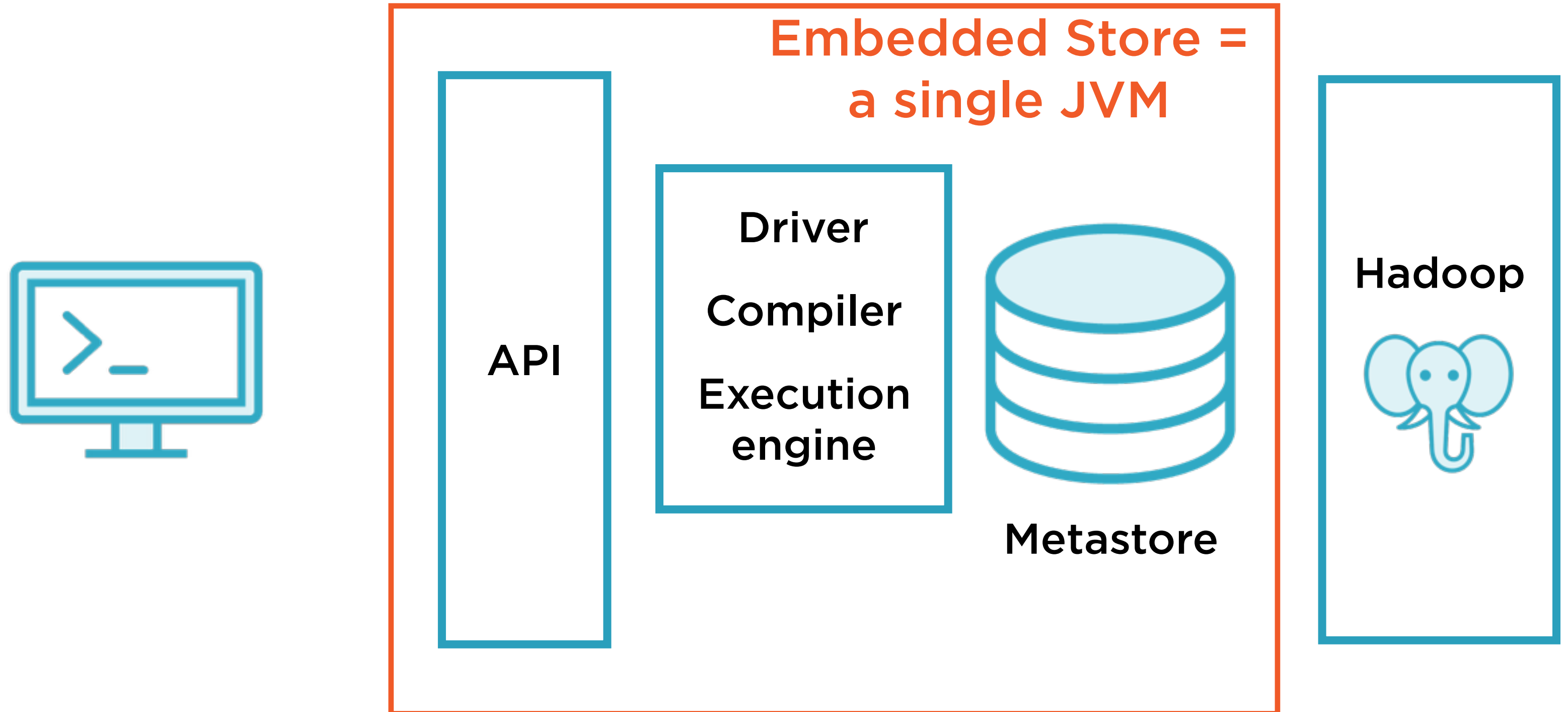
# Basic Hive Building Blocks



# Basic Hive Building Blocks



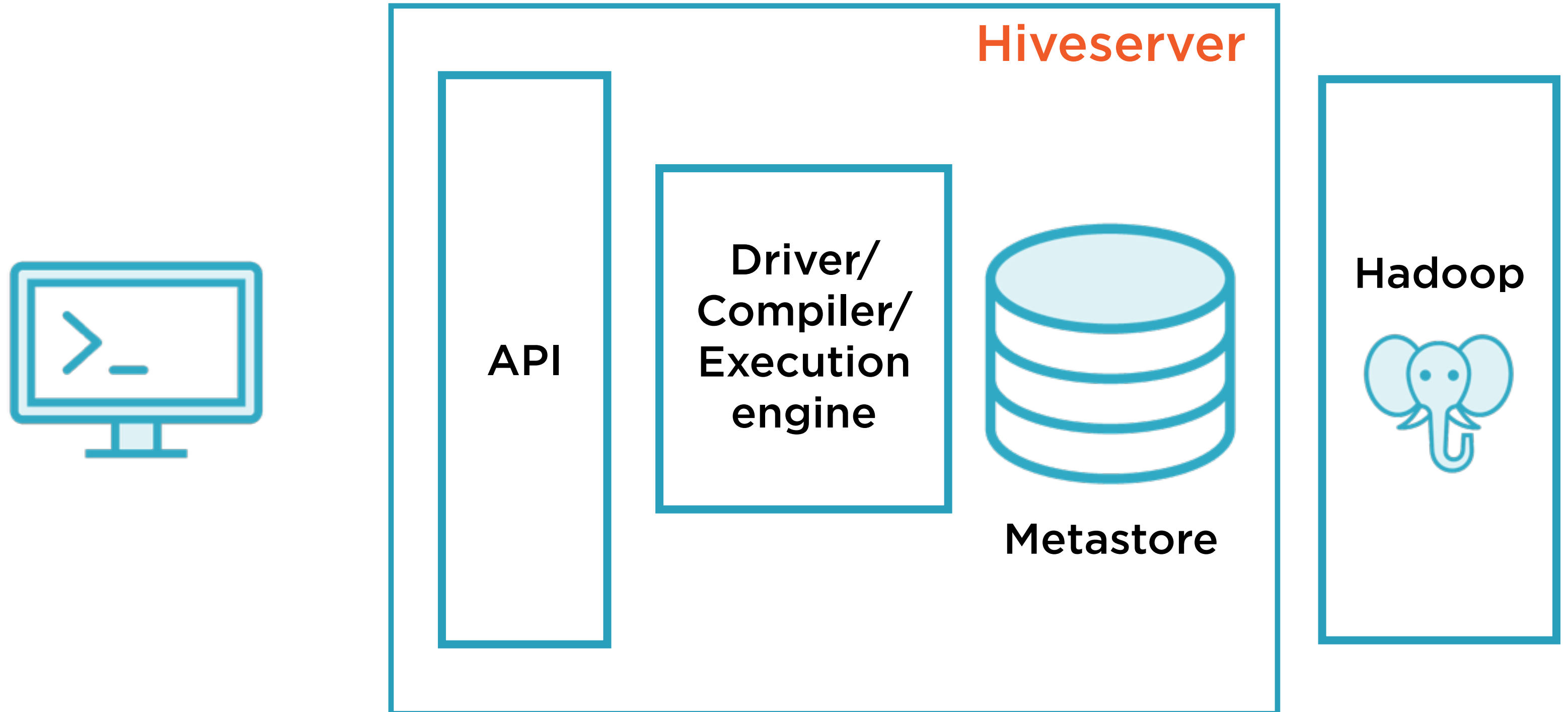
# Basic Hive Building Blocks



Demo

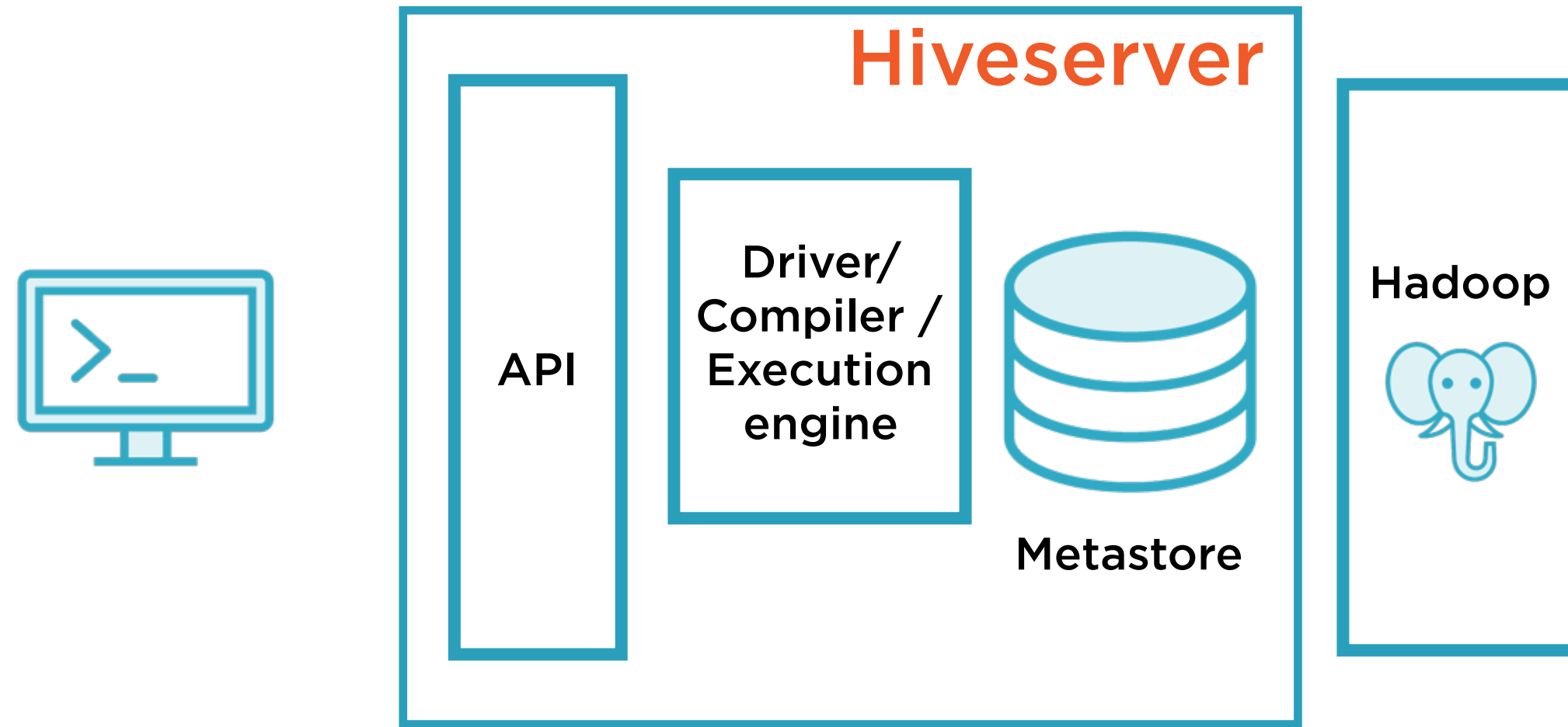
**Install and set up Hive on your local machine**

# Basic Hive Building Blocks



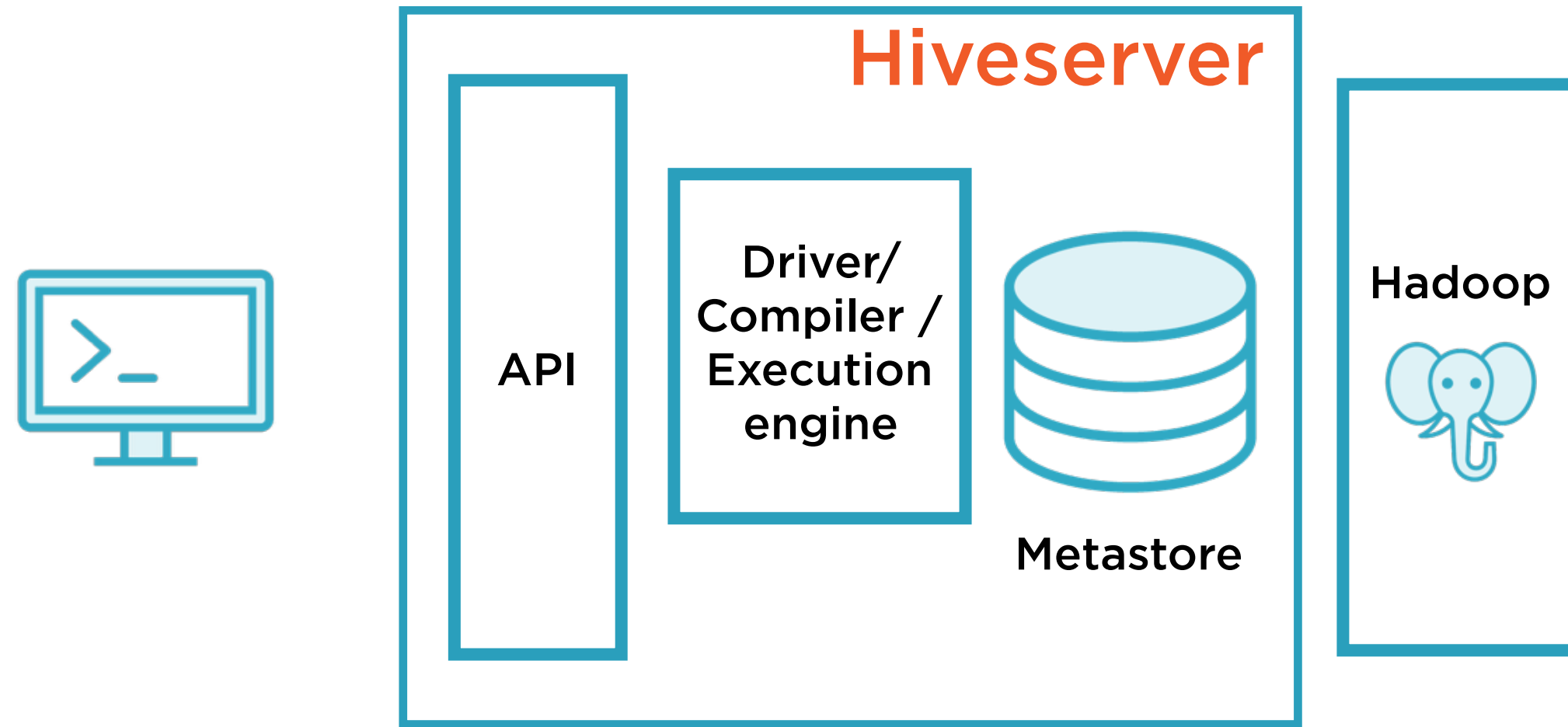


# Basic Hive Building Blocks



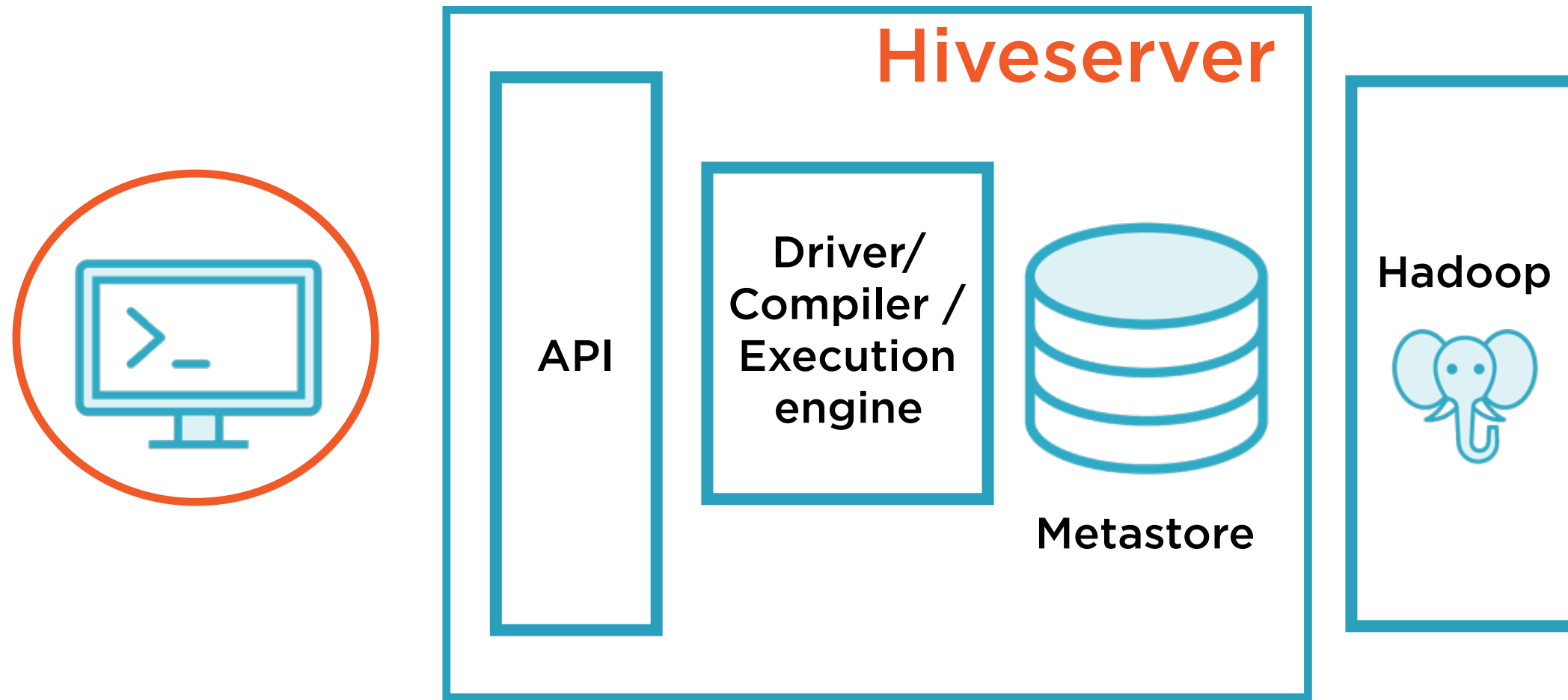
The latest versions of Hive run a re-written server called **Hiveserver2**

# Basic Hive Building Blocks



Hiveserver2 supports **concurrent** clients and better **authentication** and **authorization**

# Basic Hive Building Blocks



**Beeline is a command line interface which works with this new server**

# Beeline vs. Hive



## Hive CLI

- directly accesses the Hive metastore and driver
- does not go through the Hiveserver2 API

## Beeline CLI

- more secure as it uses the Hiveserver2 API
- has better authentication and authorization

# Demo

**Run the Beeline command line interface  
in embedded mode**

# Demo

**Simple queries in HiveQL using a single table**

# Demo

**One time execution from the terminal**

Demo

**Joins using multiple tables**



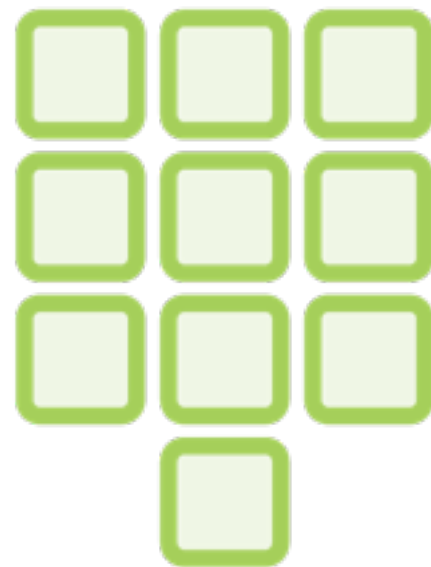
# Hive Data Types

---

# Hive Data Types



**Boolean**



**Numeric**



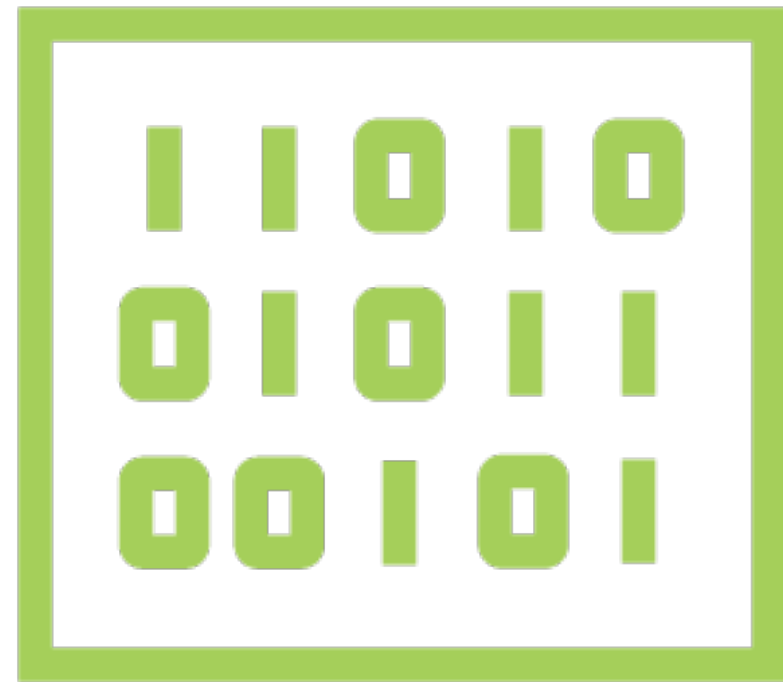
**String**



**Timestamp**

**Primitive data types**

Boolean

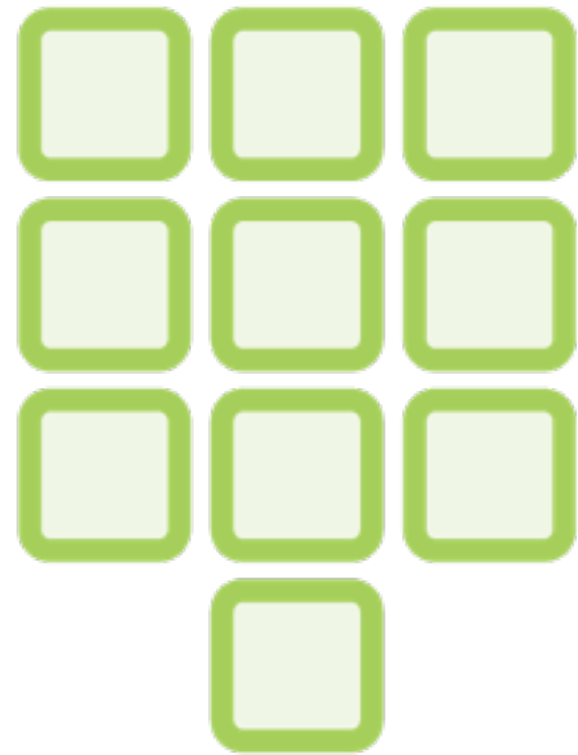


true or false

yes/no

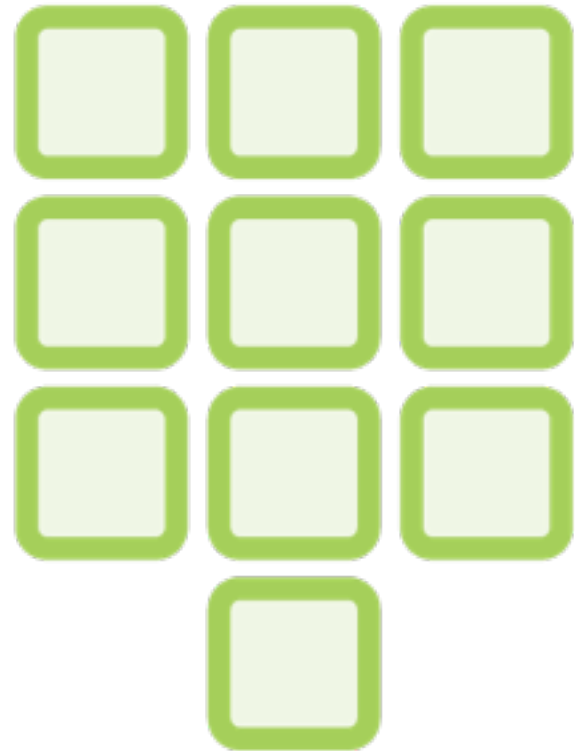
questions

Numeric



# Integers or Decimals

# Integers

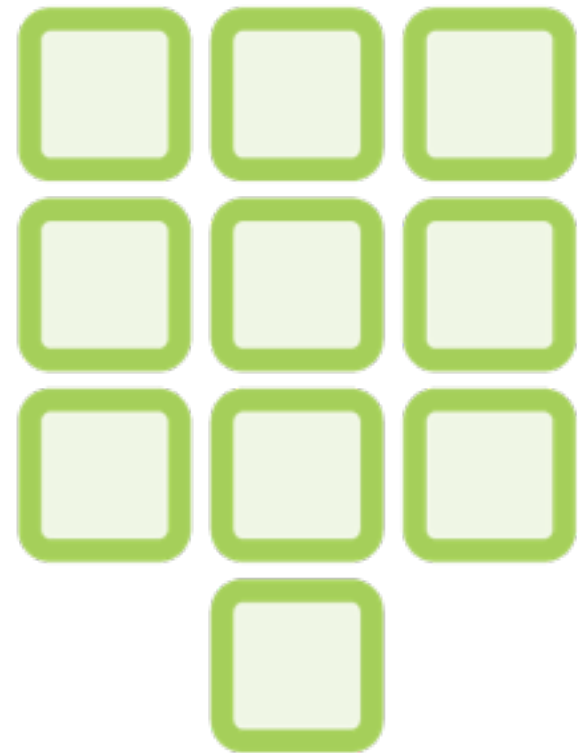


**Tinyint:** 1 byte, range -128 to 128

**Smallint:** 2 bytes, range  $-2^{15}$  to  $2^{15} - 1$

**Int:** 4 bytes, range  $-2^{31}$  to  $2^{31} - 1$

**Bigint:** 8 bytes, range  $-2^{63}$  to  $2^{63} - 1$



## Decimals

**Float:** 4 bytes

**Double:** 8 bytes

**Decimal:** Arbitrary precision

- `dec(10, 2)`



# String

**String:** Unbounded, variable length character string

**Char:** Fixed length character string

**Varchar:** Bounded, variable length character string

# Timestamp



**Integers:** Unix timestamp in nanoseconds

**Floats:** Unix timestamp in seconds with decimal precision

**String:** JDBC compliant "YYYY-MM-DD HH:MM:SS.ffffffffff"



# Dates



**Values of the form "YYYY-MM-DD"**

# Summary

**Set up Hive on your machine**

**Run Hive queries using the Beeline interface**

**Understood primitive data types in Hive**