



APACHE PHOENIX

SQL for HBase

What is Phoenix?

- A SQL driver for HBase that supports transactions
- Fast, low-latency - OLTP support
- Originally developed by Salesforce, then open-sourced
- Exposes a JDBC connector for HBase
- Supports secondary indices and user-defined functions
- Integrates with MapReduce, Spark, Hive, Pig, and Flume

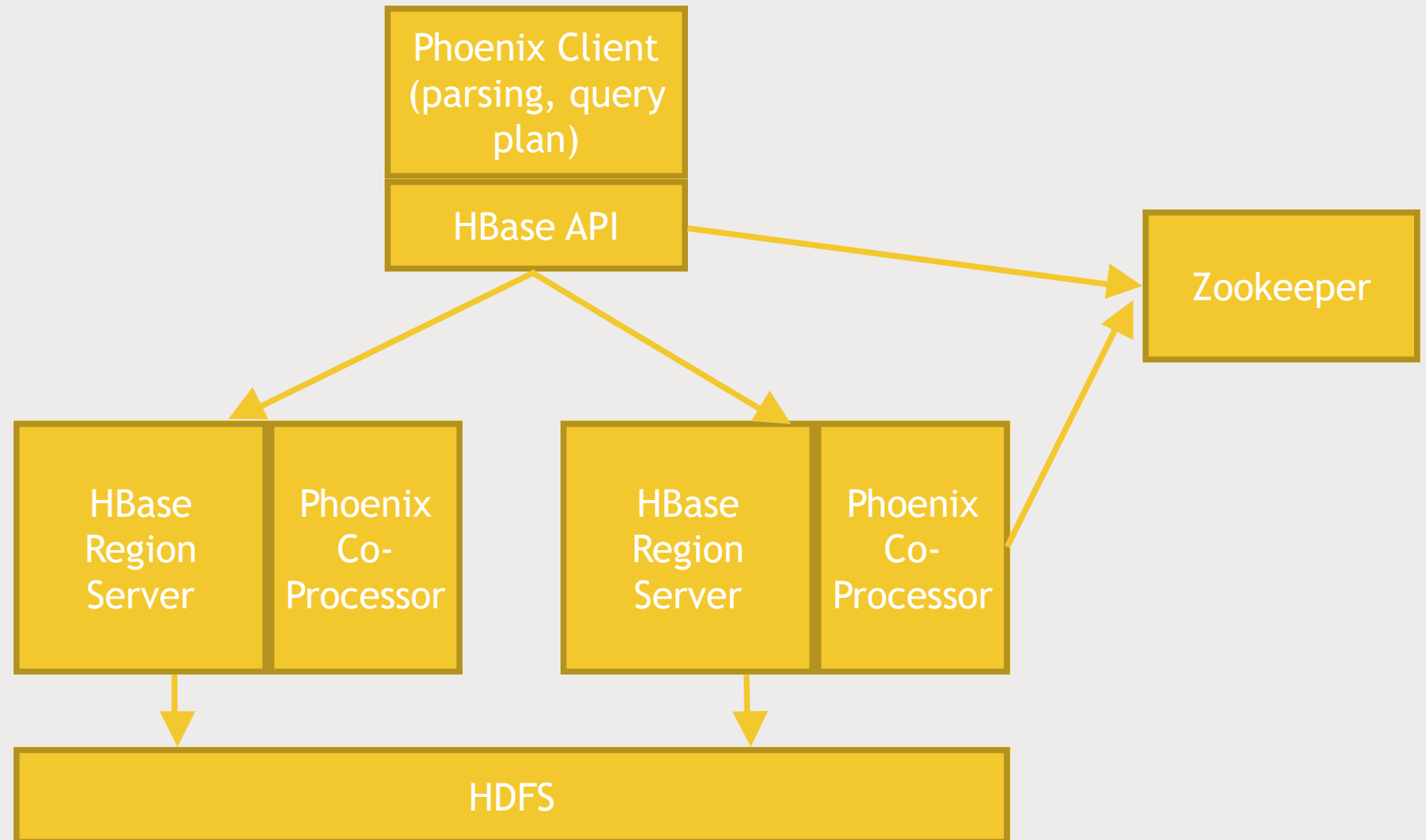


Why Phoenix?



- It's really fast. You probably won't pay a performance cost from having this extra layer on top of Hbase.
- Why Phoenix and not Drill?
 - *Well, choose the right tool for the job.*
- Why Phoenix and not HBase's native clients?
 - *Your apps, and analysts, may find SQL easier to work with.*
 - *Phoenix can do the work of optimizing more complex queries for you*
 - But remember HBase is still fundamentally non-relational!

Phoenix architecture



Using Phoenix

- Command-Line Interface (CLI)
- Phoenix API for Java
- JDBC Driver (thick client)
- Phoenix Query Server (PQS) (thin client)
 - *Intended to eventually enable non-JVM access*
- JAR's for MapReduce, Hive, Pig, Flume, Spark



Let's Play

- Install Phoenix on our Hortonworks Sandbox
- Mess around with the CLI
- Set up a users table for MovieLens
- Store and load data to it through the Pig integration

One more cool Phoenix picture

