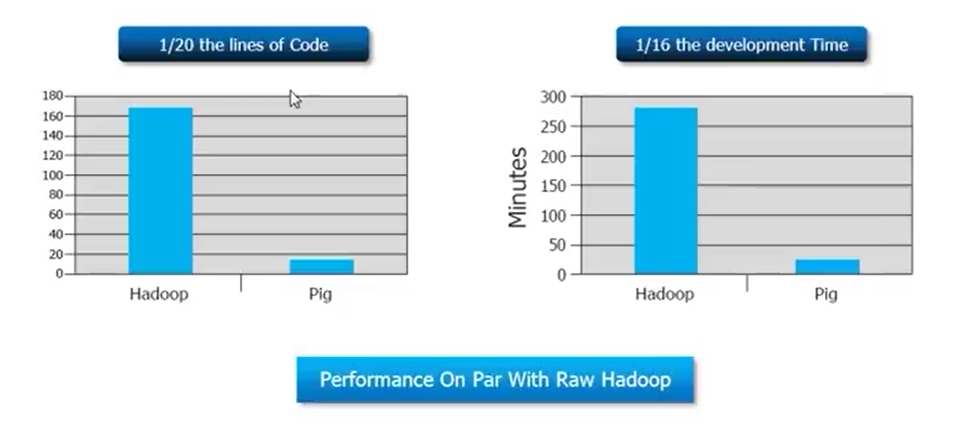
**PIG**

**NEED OF PIG**

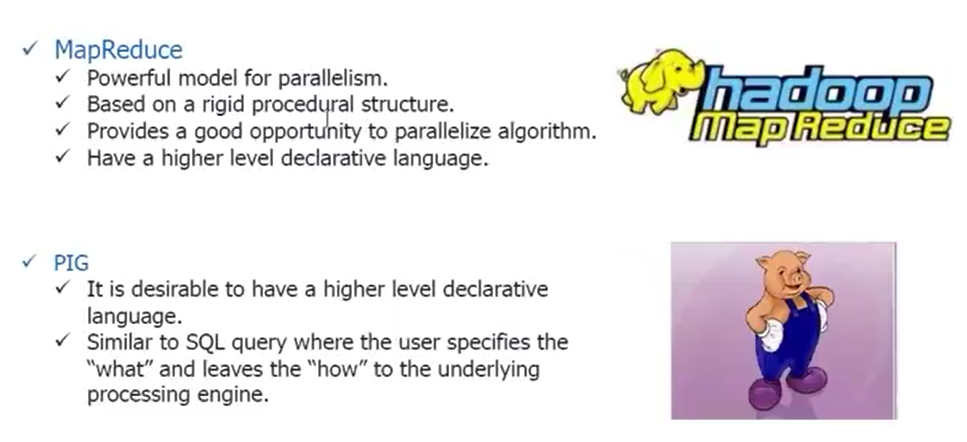
* Very less time required for development
* 10 Lines of PIG are almost equal to around 200 lines of codes for Java
* Built in operators like Join, Group, Filter, Sort and many more. We don’t have to write a code for the same.
* Default operators are always there.

**WHY WAS PIG CREATED**

* PIG was actually created by Yahoo.
* This was actually created because an ad-hoc way of creating and executing map-reduce jobs on very large data sets.
* Very much suitable for Non Java resources and can be developed very rapidly
* PIG not only reduces the lines of code but also it reduces the development time drastically



**Why PIG when MapReduce is already there**



PIG is a dataflow language at the top of Hadoop and makes it possible to create complex jobs to process large volumes of data quickly and efficiently. Most of the use cases involve following:-

**Case 1 :- Time sensitive data loads**

That means we need to process any set of data after every 2,3or 4 hours periodically.

**Case 2 :- Processing many data sources**

Something ETL kind of processes, that means you must be getting from various data sources, mixing them and then processing them and then provide the final output

**Case 3 :- Analytical insight through sampling**

Analyzing data patterns etc.

**What is PIG**

PIG is an open source high level dataflow system. It provides a simple language for queries and data manipulation.