HBase Local Cluster

Karthik Kumar Viswanathan
"guilt"

Cluster in a Process

- HBase should ideally be run on top of a dedicated HDFS cluster

- But for development, you may not want that.

  https://github.com/guilt/hbase-local

- Minimize time to HBase on most laptops

- Also, I run it on my Mac.

What does HBase Need?

- ZooKeeper

- DFS

- Running a DFS requires multiple Hadoop Daemons – or you could use MiniDFSCluster

- Running ZooKeeper embedded is good, but it is recommended that you run ZooKeeper separately.

Hadoop Dependencies

- Bring it from Maven

- Uses HortonWorks' repository (was tested on Cloudera's as well)

- Hadoop 1 or 2 – with minor changes.

- Includes AppAssembler, so you can package these separately and run it on a common machine.

A Demo, Perhaps.

What I Learnt

- Hadoop is awesome!

- It's open source, and there is no one-shoe-fits-all

- Developer productivity is a key factor in enabling Hadoop. I got a chance to do it during the hacknight. (HADOOP-9635)