1) HADOOP IN LAYMENS TERM-

LAYMENS TERM stands for, Instead of one big person solving, as a society or group of people are stronger. In layman terms instead of one big computer solving, multiple small boxes will solve the problem.

Hadoop is an open source software that makes doing mapreduce type programming easier. Once you setup a hadoop cluster over the 30 machines, you can give it any program and data and backround work and give you the result, there is no need to install,load,and copy program over different machines.

Consider a example, I am a project manager and a new project is assigned to me which is half done and need to do it within almost no time. then what will i do? i will read out and learn all previously done work and divide the work into parts and individually we all will learn about it and come to conclusion. Hadoop does the same. You are a master(manager) in Hadoop who assigns a task to each of your slave nodes And they process individually and give you a result. The master node may or may not give a newer task from these. But will finally return a result. Hadoop is one such platform that stores and processes big data captured by any biggies(amazon,flipkart) by tracking you while you're browsing and shopping ,and provide suggestions. like we see on goibibo.com if you book a flight ticket,it ask you to book room or lodge.

2)

Apache Hadoop is an open source software framework that allows large sets of data to be processed using commodity hardware. Hadoopruns on large cluster of nodes that are linked to create wide newtwork.(attachment assign2.5.3)

like in previous question , we took example of goibibo.com, Apache is a web-server that provide a service like goibibo accessible over the World, while Hadoop is a platform that stores and processes big data captured by goibibo by tracking users while you're browsing or booking.

The apache framework has  
1.management and scheduling- here ambani comes into picture it is a RESTful API which provides easy to use web user interface for Hadoop management.

1. coordination- Zookeeper is the king of coordination and provides simple, fast, reliable and ordered operational services for a Hadoop cluster.

3.workflow and scheduling- Oozie is a workflow scheduler where the workflows are expressed as Graphs.

1. data integration- Sqoop component is used for importing data from external sources into related Hadoop components like HDFS. It can also be used for exporting data from Hadoop o other external structure.

5.Query and NoSql database- Hives makes querying faster through indexing. HBase supports random reads and also batch computations using MapReduce. With HBase NoSQL database enterprise can create large tables with millions of rows and columns on hardware machine.

6.Scripting- Apache Pig is a convenient tools developed by Yahoo for analysing huge data sets efficiently and easily. it is optimized, extensible and easy to use

7.HDFS and Map Reduce-

3) Reason to learn Big Data- 1. It is flexible. More data systems can be added, edited or deleted when required. 2. It is cost-effective and practical. More storage units can be added by procuring readily-available storage from vendors. 3. It is open source, providing ample flexibility for corporations to customize it the way they want for effective use. 4.High Demand of skilled workers. 5.You will have variety job titles to choose 6. Increasing pay for data analytics professionals 7.Big data analytics is vast and spreded widely 8. Data analytics is taking over faster than expected