CSE 6242 A / CS 4803 DVA

Feb 28, 2013

Scaling Up

Hadoop

Duen Horng (Polo) Chau
Georgia Tech

Some lectures are partly based on materials by
Professors Guy Lebanon, Jeffrey Heer, John Stasko, Christos Faloutsos, Le Song

What You've Learned So Far...

How to collect, clean, integrate and store data

Some techniques, tools, and applications (both computation and visualization)

- Classification
- Clustering
- Dimensionality Reduction
- Graph centrality, algorithms

Easily work for **"large" data** (e.g., gigabytes), even on a **single commodity computer**

But what if you data is **really** large?

Really big, as in...

- Petabytes (PB, about 1000 times of terabytes)
- Or beyond: exabyte, zettabyte, etc.

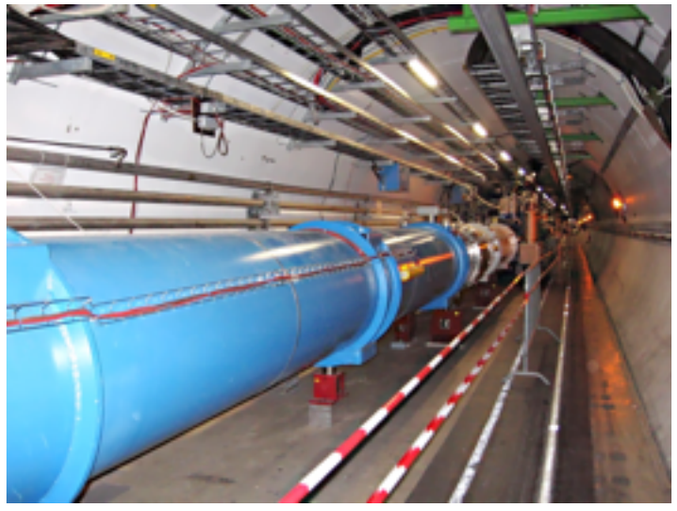
Do we *really* need to deal with such scale?

- Yes!

Big Data is Quite Common...

Google processed **24 PB / day** (2009)

Facebook's add **0.5 PB / day** to its data warehouses



CERN generated **200 PB** of data from "Higgs boson" experiments

Avatar's 3D effects took **1 PB** to store

So, think **BIG**!

http://www.theregister.co.uk/2012/11/09/facebook_open_sources_corona/
http://thenextweb.com/2010/01/01/avatar-takes-1-petabyte-storage-space-equivalent-32-year-long-mp3/
http://dl.acm.org/citation.cfm?doid=1327452.1327492

How to analyze such large datasets?

First thing, how to **store** them?

Single machine? 4TB drive is out

Cluster of machines?

- How many machines?
- Need to worry about machine and drive failure. **Really?**
- Need data backup, redundancy, recovery, etc.

Failure Trends in a Large Disk Drive Population
http://static.googleusercontent.com/external_content/untrusted_dlcp/research.google.com/en/us/archive/disk_failures.pdf

How to analyze such large datasets?

First thing, how to **store** them?

Single machine? 4TB drive is out

Cluster of machines?

- How many machines?
- Need to worry about machine and drive failure. **Really?**
- Need data backup, redundancy, recovery, etc.

3% of 100,000 hard drives fail within **first 3 months**

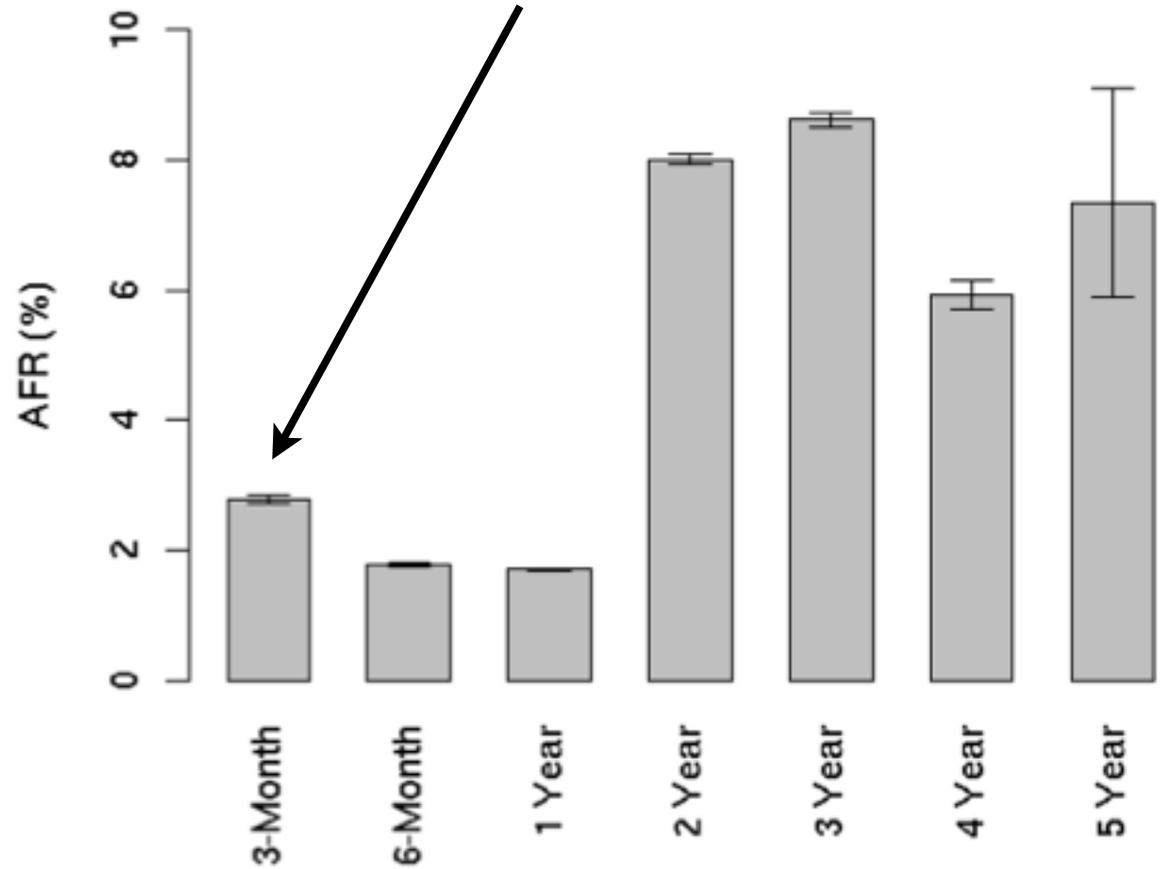


Figure 2: Annualized failure rates broken down by age groups

Failure Trends in a Large Disk Drive Population
http://static.googleusercontent.com/external_content/untrusted_dlcp/research.google.com/en/us/archive/disk_failures.pdf

How to analyze such large datasets?

How to analyze them?

- What software libraries to use?
- What programming languages to learn?
- Or more generally, what framework to use?

Open-source software for reliable, scalable, distributed computing

Written in Java

Scale to thousands of machines

- Linear scalability: if you have 2 machines, your job runs twice as fast

Uses simple programming model (MapReduce)

Fault tolerant (HDFS)

- Can recover from machine/disk failure (no need to restart computation)

Why learn Hadoop?

Fortune 500 companies use it

Many research groups/projects use it

Strong community support, and favored/backed my major companies, e.g., IBM, Google, Yahoo, Microsoft, etc.

It's **free**, open-source

Lost cost to set up (works on commodity machines)

Will be an "essential skill", like SQL

http://strataconf.com/strata2012/public/schedule/detail/22497

Elephant in the room

Hadoop created by Doug Cutting and Michael Cafarella while at Yahoo

Hadoop named after Doug’s son’s toy elephant

How does Hadoop scales up computation?

Uses master-slave architecture, and a simple computation model called **MapReduce** (popularized by Google's paper)

Simple explanation

1. Divide data and computation into smaller pieces; each machine works on one piece
2. Combine results to produce final results

MapReduce: Simplified Data Processing on Large Clusters
http://static.usenix.org/event/osdi04/tech/full_papers/dean/dean.pdf

How does Hadoop scales up computation?

More technically...

1. **Map phase**
   Master node **divides** data and computation into smaller pieces; each machine ("mapper") works on one piece independently in parallel

2. **Shuffle phase** (automatically done for you)
   Master **sorts and moves** results to "reducers"

3. **Reduce phase**
   Machines ("reducers") **combines** results independently in parallel

An example
Find words' frequencies among text documents

Input

- "Apple Orange Mango Orange Grapes Plum"
- "Apple Plum Mango Apple Apple Plum"

Output

- Apple, 4
  Grapes, 1
  Mango, 2
  Orange, 2
  Plum, 3

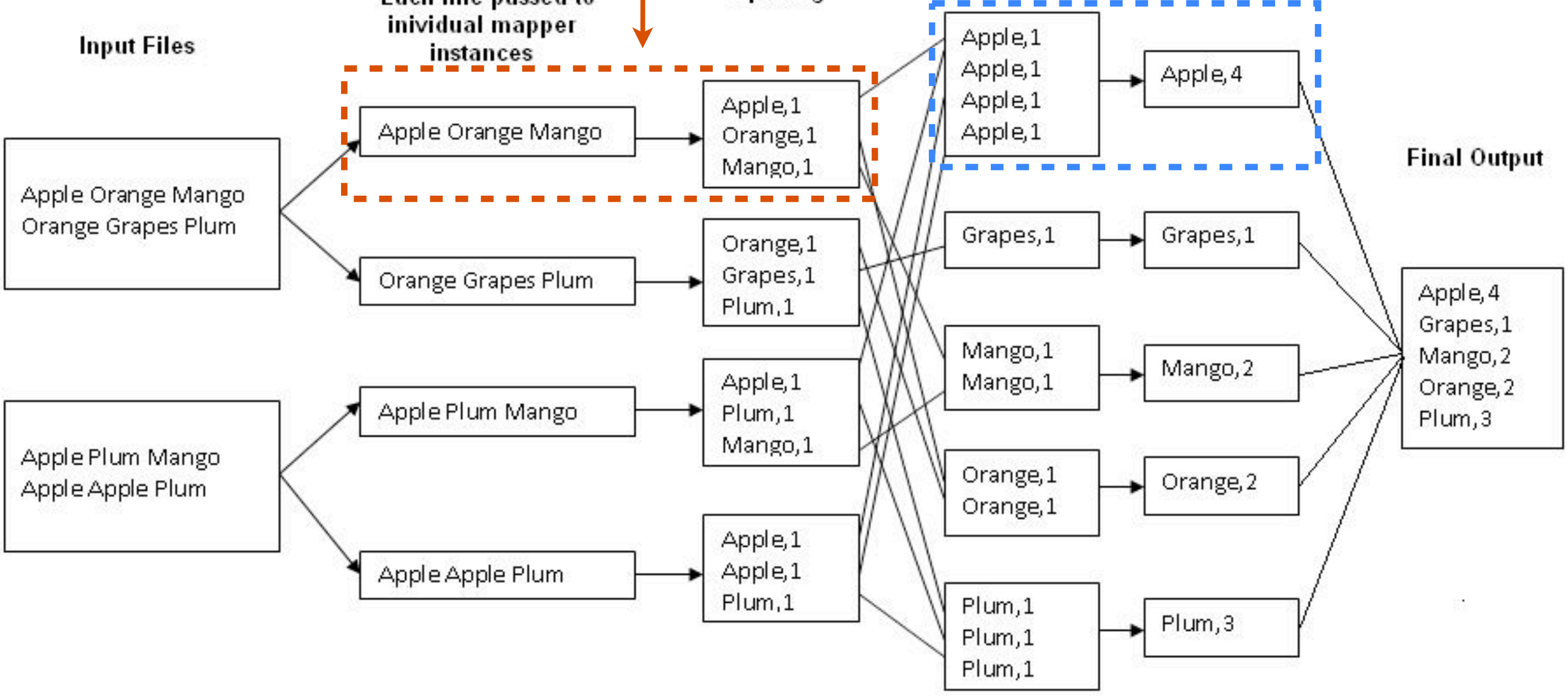
http://kickstarthadoop.blogspot.com/2011/04/word-count-hadoop-map-reduce-example.html

Each machine (mapper) outputs a **key-value pair**

Pairs sorted by key (automatically done)

Each machine (reducer) combines pairs into one

Master divides the data (each machine gets one line)

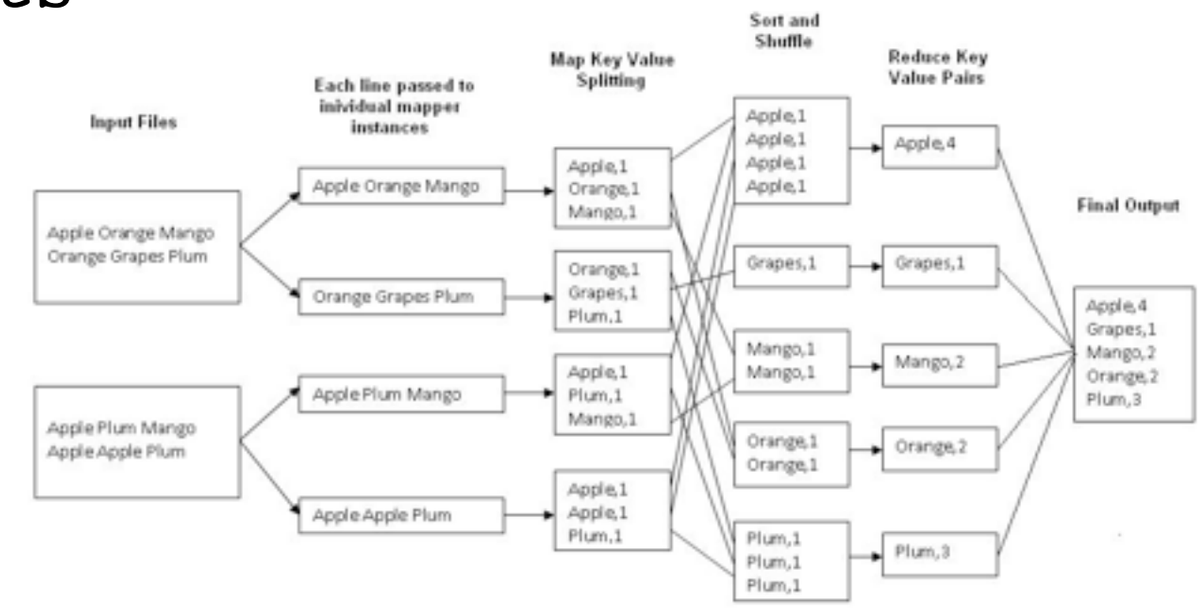


One machine can be both a mapper and a reducer

How to implement this?

```
map(String key, String value):
    // key: document id
    // value: document contents
    for each word w in value:
      emit(w, "1");
```

```
reduce(String key, Iterator values):
    // key: a word
    // values: a list of counts
    int result = 0;
    for each v in values:
      result += ParseInt(v);
    Emit(AsString(result));
```



What can you use Hadoop for?

As a “swiss knife”

- For many types of analyses/tasks
- But not all of them
- Next time will talk about machine learning and data mining libraries built on top of Hadoop
- Also tools that make it easier to write MapReduce program (Pig), or to query results (Hive)

What if a machine dies?

Replace it!

- “map” and “reduce” jobs can be redistributed to other machines

Hadoop’s HDFS (Hadoop File System) enables this

HDFS: Hadoop File System

A distribute file system

built on top of OS's existing file system to provide redundancy and distribution

HSDF hides complexity of distributed storage and redundancy from the programmer

In short, you don't need to worry much about this!

How to try Hadoop?

Hadoop can run on a single machine (e.g., your laptop)

- Takes < 30min from setup to running

Or a “home-brew” cluster

- Research groups often connect retired computers as a small cluster

Amazon EC2 (Amazon Elastic Compute Cloud)

- You only pay for what you use, e.g, compute time, storage
- You will use it in HW3

http://aws.amazon.com/ec2/

Next time

Pig

HBase

Hive

Mahout

Pegasus

...