Thinking In Map Reduce

Alan Gates

Yahoo!

About Me

- Pig Committer
- Member of Hadoop PMC
- Architect for Pig development group at Yahoo!

A Little Motivation

A Little Motivation

- You have terabytes of data, you could…
  - process it on your desktop, but that's slow
  - buy a parallel database system, but that's expensive
  - write your own parallel processing system, but that's hard

A Little Motivation

- You have terabytes of data, you could…
  - process it on your desktop, but that's slow
  - buy a parallel database system, but that's expensive
  - write your own parallel processing system, but that's hard

A Little Motivation

- You have terabytes of data, you could…
  - process it on your desktop, but that's slow
  - buy a parallel database system, but that's expensive
  - write your own parallel processing system, but that's hard

Map-Reduce gives you a simple programming paradigm while solving the hard parts of parallelism

Hadoop

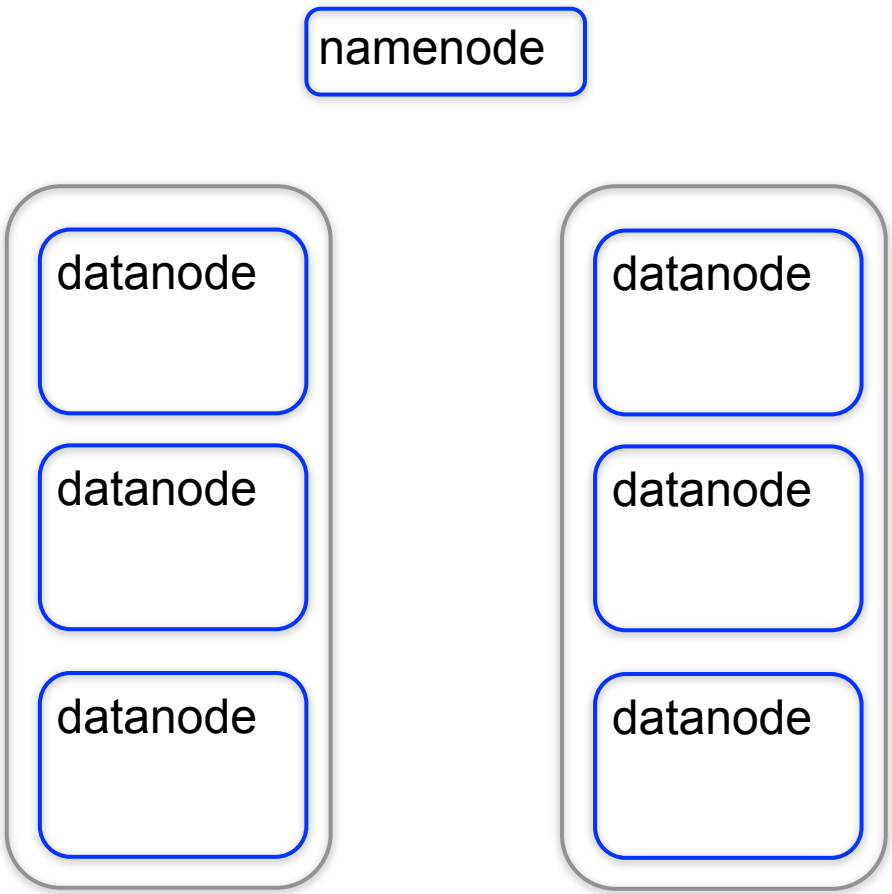
- Hadoop is an Apache open source project
- It provides
  - HDFS, a distributed file system
  - Map-Reduce, a simple parallel processing paradigm
  - Languages for data processing, Pig Latin and Hive SQL
  - And a whole lot more
- Based heavily off of Map Reduce and Google File System papers published by Google

Hadoop Distributed File System

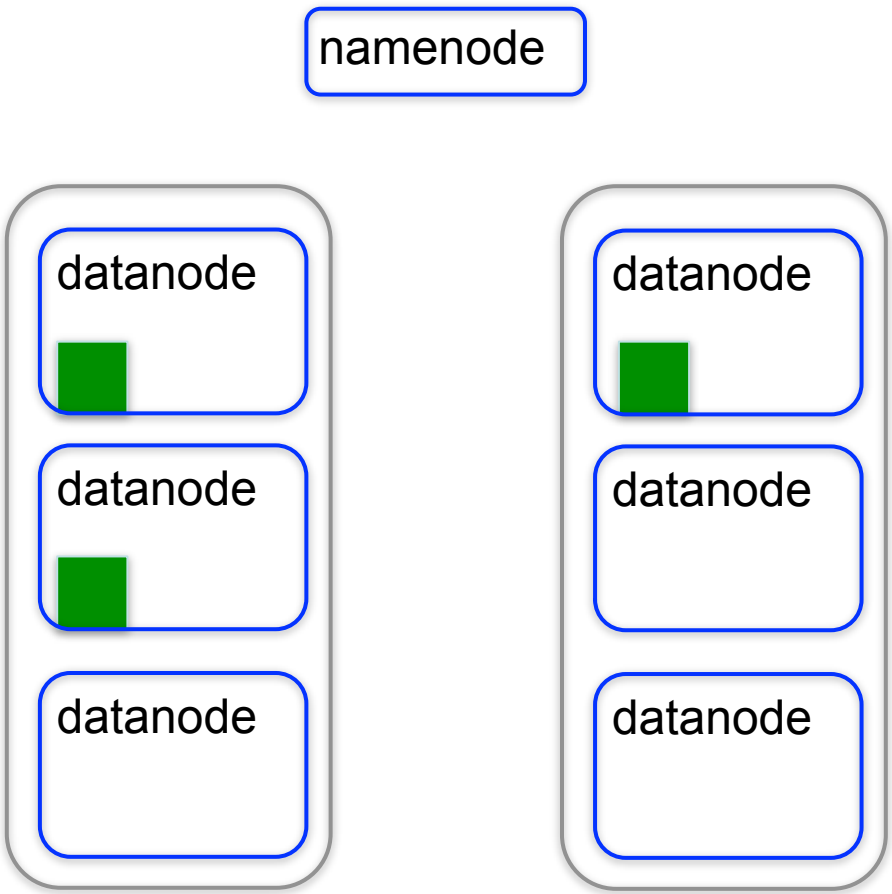
- Presents single filesystem for entire cluster, can be petabytes in size
  - Managed by a *namenode*
  - Each machine in the cluster is a *datanode*
  - Files are write once, read many
  - Optimized for streaming large files
- File system interface in Java, C, and command line
- Files are broken into blocks and distributed across datanodes

HDFS Files

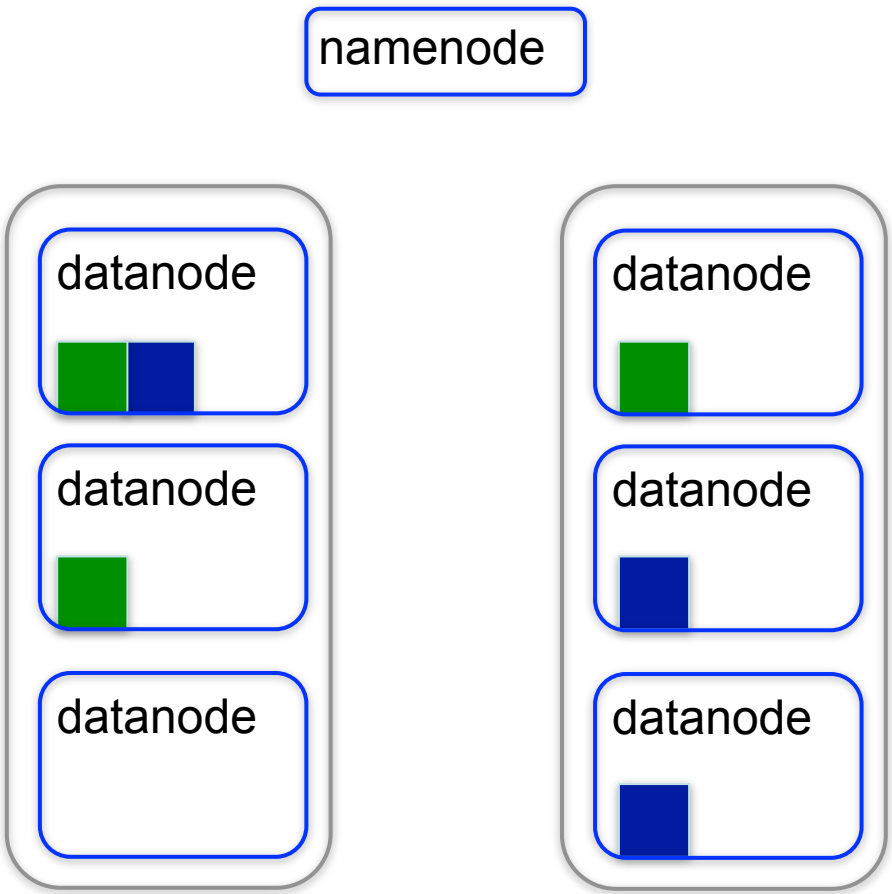
HDFS Files


namenode
datanode
datanode
datanode
datanode
datanode
datanode

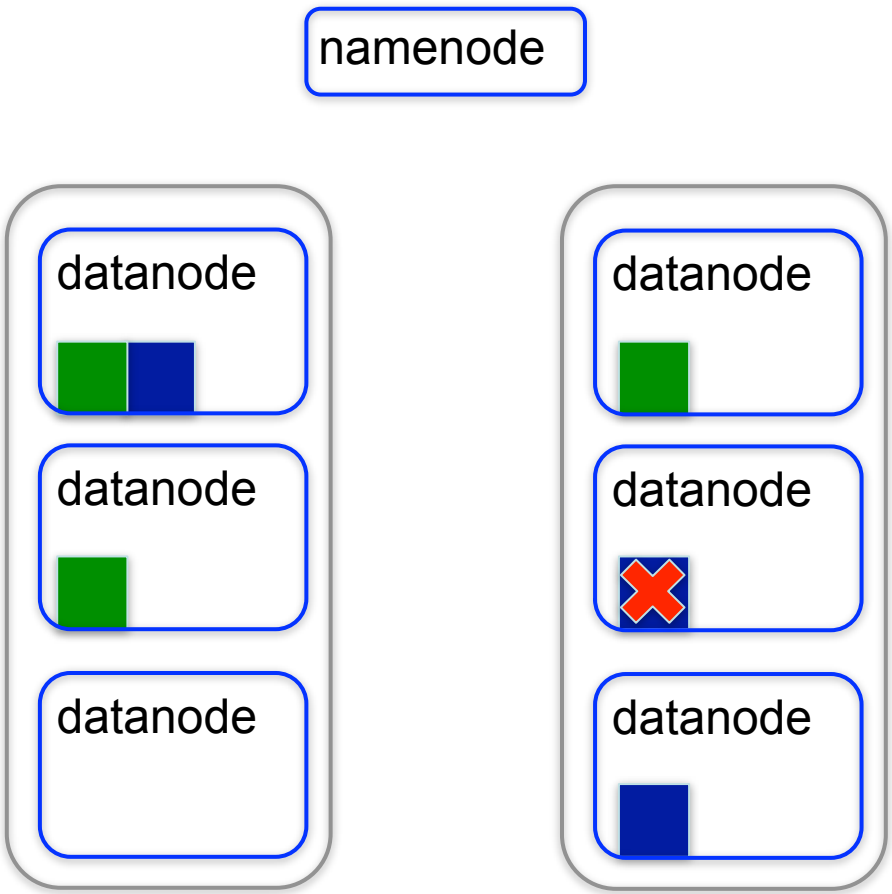
HDFS Files


namenode
datanode
datanode
datanode
datanode
datanode
datanode

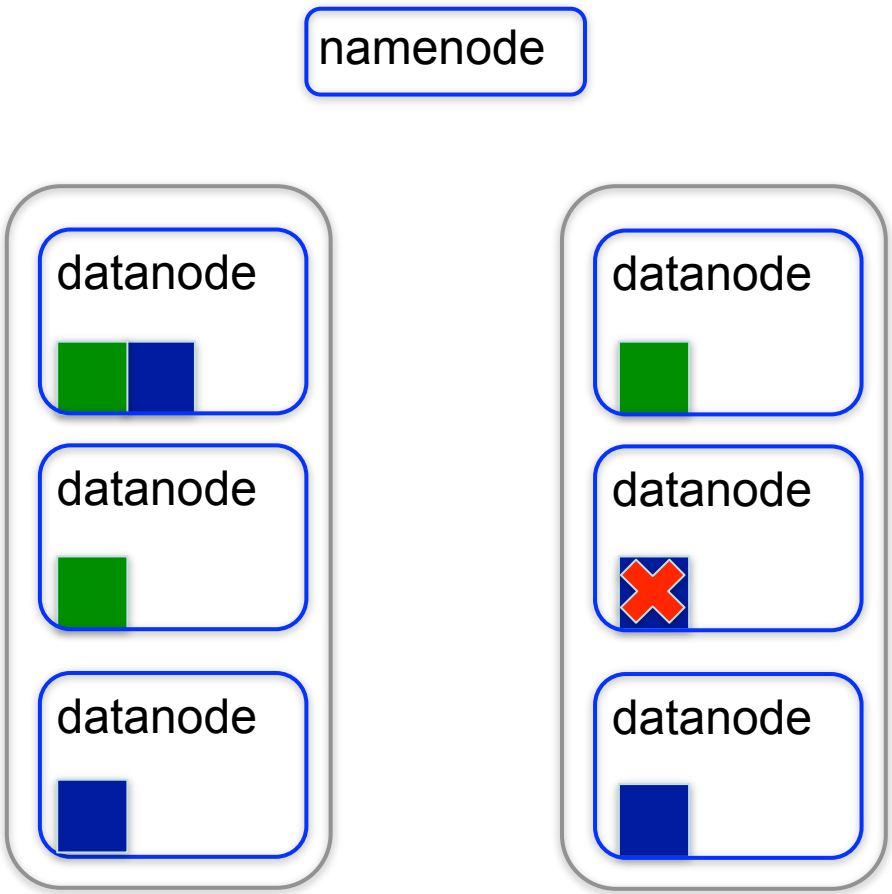
HDFS Files


namenode
datanode
datanode
datanode
datanode
datanode
datanode

HDFS Files


namenode
datanode
datanode
datanode
datanode
datanode
datanode

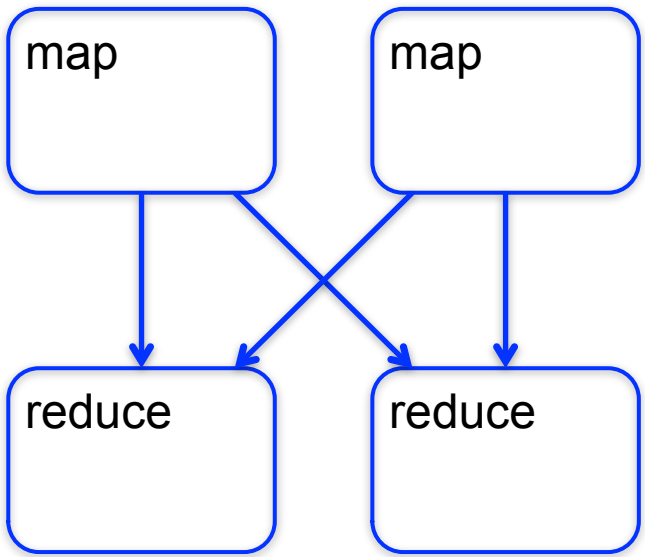
HDFS Files


namenode
datanode
datanode
datanode
datanode
datanode
datanode

Map-Reduce

- Computation is moved to the data
- A simple yet powerful programming model
  - Map: every record handled individually
  - Shuffle: records collected by key
  - Reduce: key and iterator of all associated values
- User provides:
  - input and output (usually files)
  - map Java function
  - key to aggregate on
  - reduce Java function
- Opportunities for more control: partitioning, sorting, partial aggregations, etc.

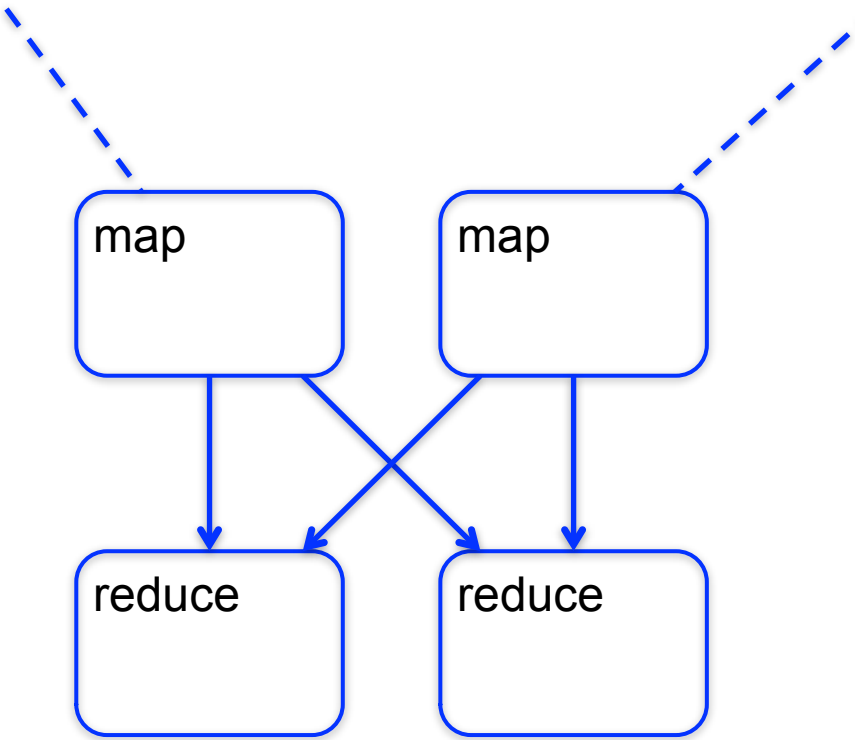
Map Reduce Illustrated



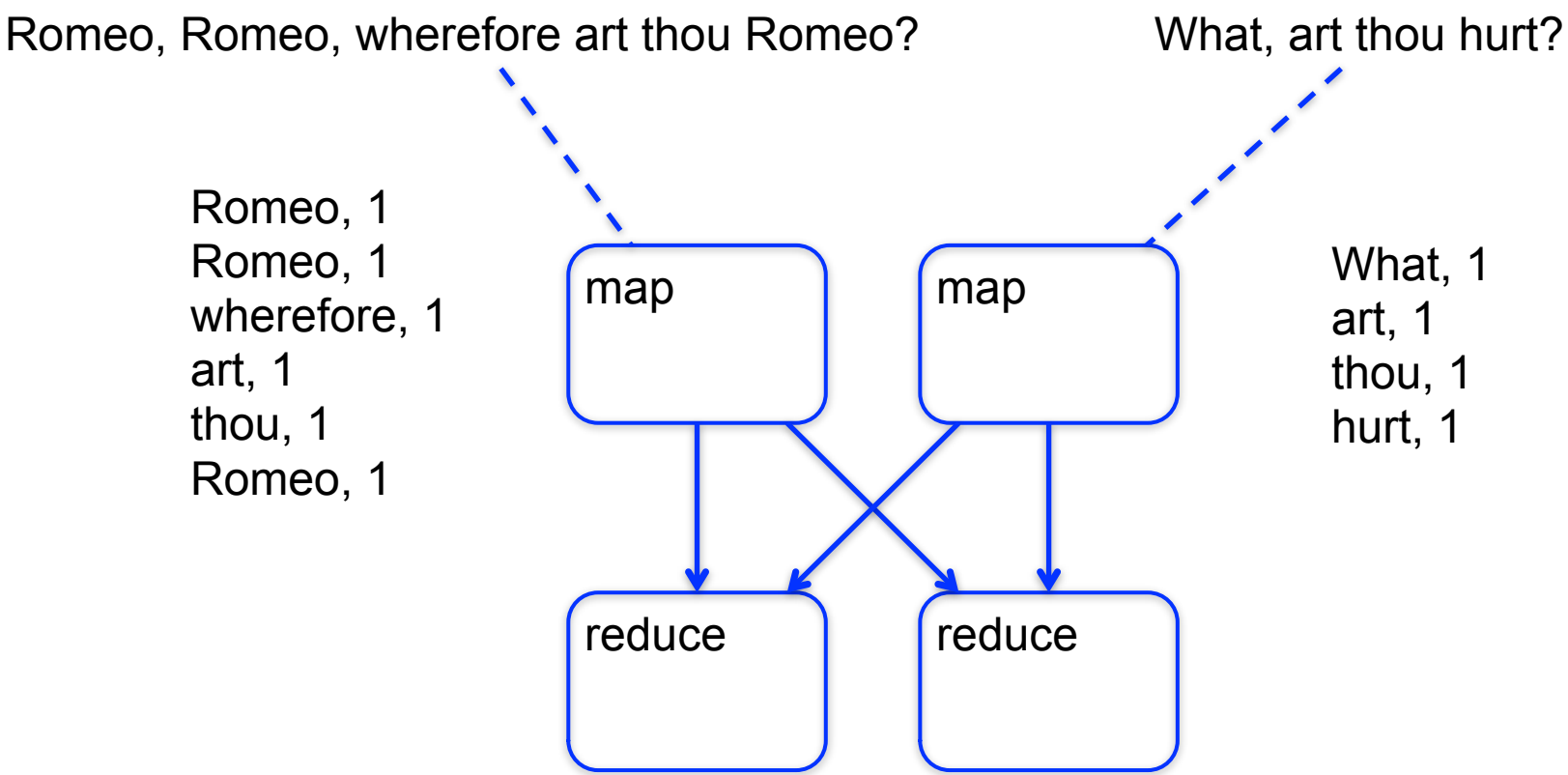
Map Reduce Illustrated

Romeo, Romeo, wherefore art thou Romeo?

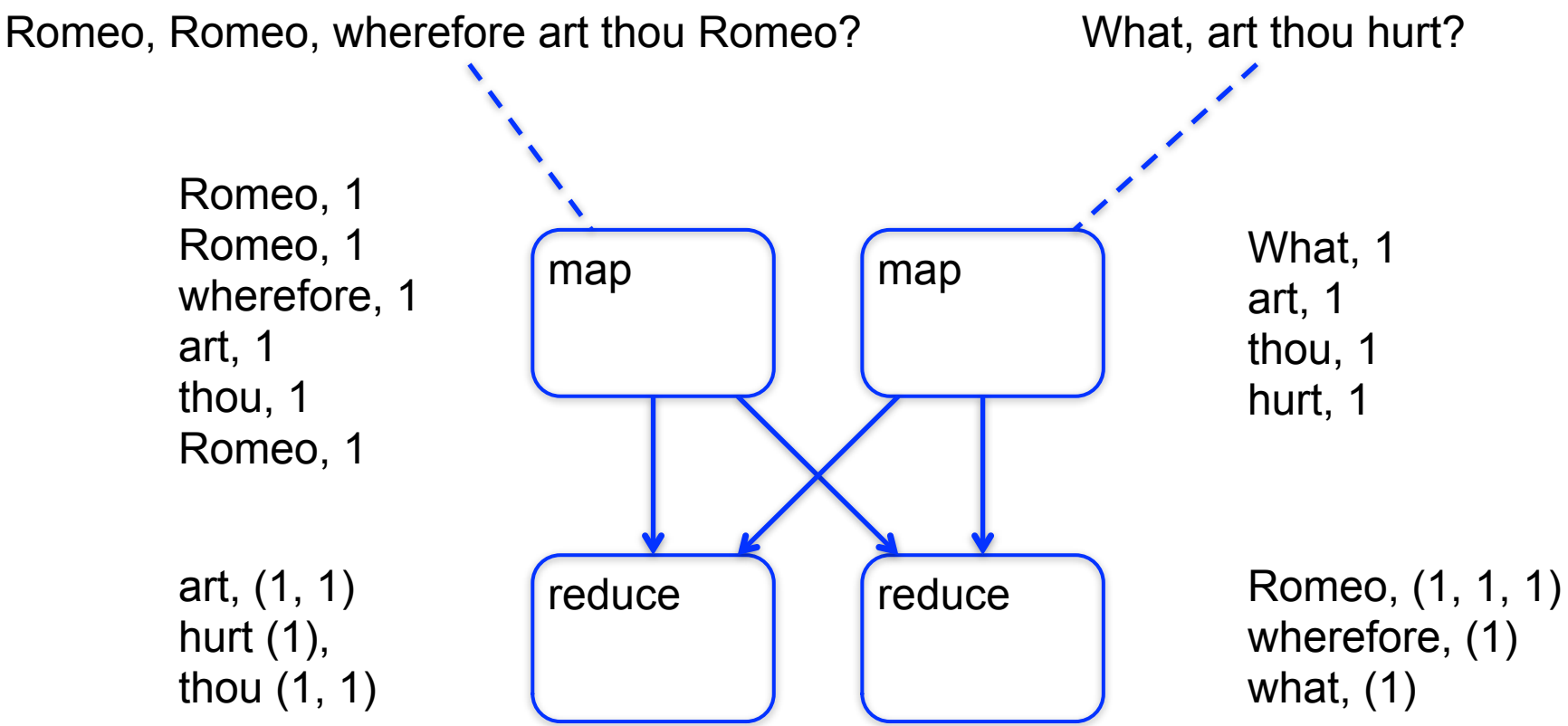
What, art thou hurt?



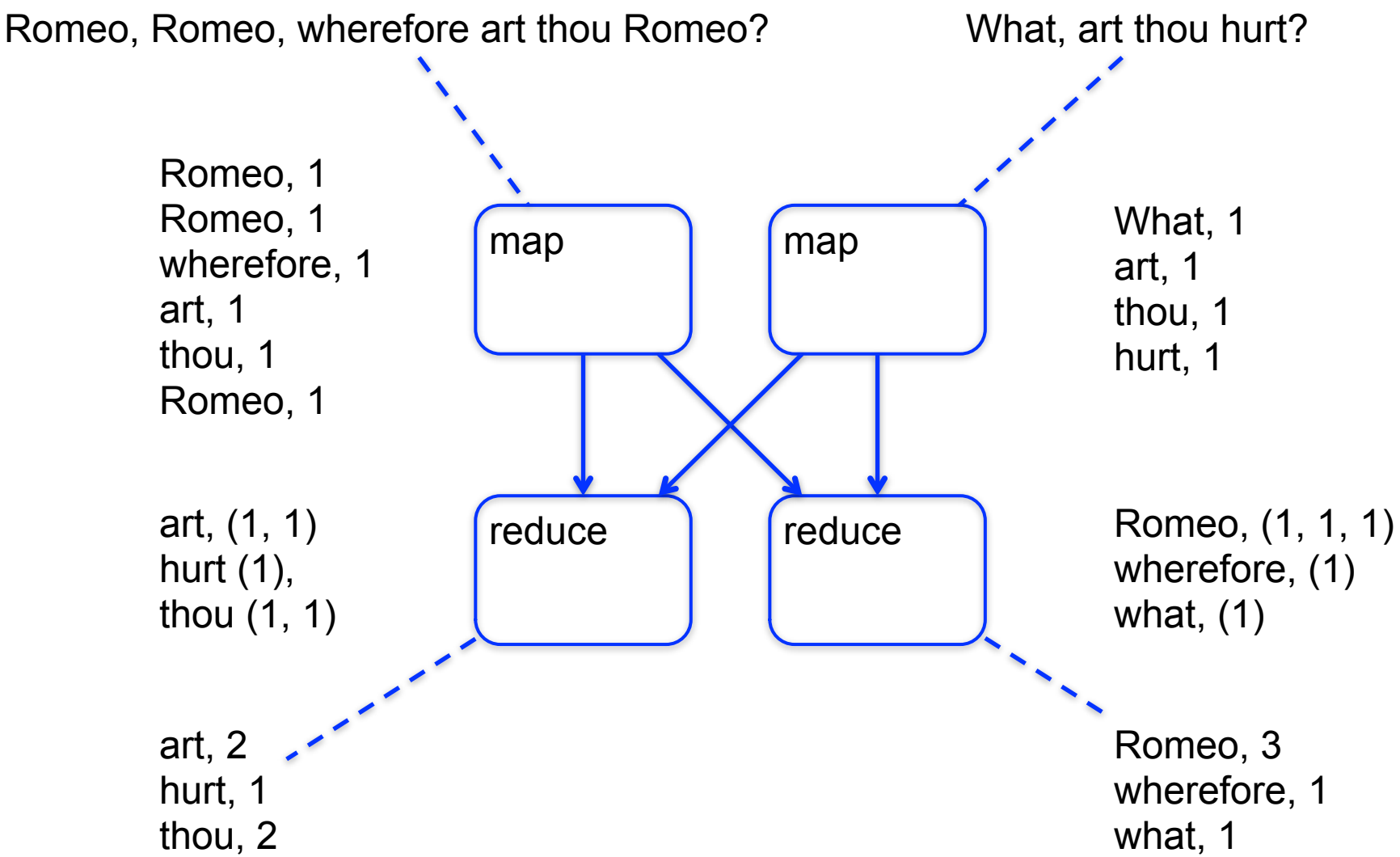
Map Reduce Illustrated



Map Reduce Illustrated



Map Reduce Illustrated



Making Parallelism Simple

- Sequential reads = good read speeds
- In large cluster failures are guaranteed; Map Reduce handles retries
- Good fit for batch processing applications that need to touch all your data:
  - data mining
  - model tuning
- Bad fit for applications that need to find one particular record
- Bad fit for applications that need to communicate between processes; oriented around independent units of work

But I Don't Want to Write in Java!

- We'll talk about Pig and Hive, higher level languages for Hadoop, in a bit.
- If you want to write in your favorite language (say, Ruby), there's streaming. User provides:
  - executables for map and reduce in language of her choice
  - input and output files
  - list of files to send to execution machine (executables, libraries, etc.)
- Not quite as fast as standard Map-Reduce, since data has to be serialized and deserialized and passed between processes.

Pig: A Parallel Data Processing Language

- Map Reduce is powerful, but
  - writing Java for everything is cumbersome
  - Hadoop does not provide many common data operations such as join
- Pig provides a data flow language (Pig Latin) with standard relational operations (join, group, sort, etc.)
- Allows users to plug in their code wherever they want

```
A = load 'romeoandjuliet' as line;
B = foreach A flatten(TOKENIZE(line)) as word;
C = group B by word;
D = foreach C generate group, COUNT(B);
```

Hive: Data Warehousing On Hadoop

- Provides data warehousing capabilities on Hadoop
- Uses Hive QL, a SQL variant, for processing
- Enables ad hoc querying and analysis of large data sets
- Allows plug in of custom map and reduce code to support more sophisticated analysis not easily doable in Hive QL.
- Focused on batch processing, not fast lookup queries
- No updates, insert of new partitions only

HBase: Big Table for Hadoop

- Provides quick lookup and update for very large data set
- Model is a columnar store where every record has a key and some number of values
- Aims to support billions of rows of millions of columns each with lookups of individual rows at web speed
- Based off Google's Big Table paper
- Oriented around point queries, not for large scans or batch jobs

Also Starring

- ZooKeeper - Centralized service for
  - maintaining shared configuration
  - distributed synchronization
  - group services

- Avro - Serialization system, for use in storing data and network communications

- Chukwa – Open source data collection system, includes toolkit for displaying, monitoring, and analyzing collected data

Who Is Using Hadoop?

- Amazon: to build product search indices
- Facebook: processing of web logs, via both Map-Reduce and Hive
- IBM and Google: making large compute clusters available to higher ed and research organizations
- Microsoft: use HBase for natural language search (as part of Powerset)
- New York Times: large scale image conversions
- Twitter: use Pig for web log processing
- Yahoo: use Map Reduce and Pig for web log processing, data model training, web map construction, and much, much more
- Many universities for teaching parallel and large data systems
- And many more, see them all at http://wiki.apache.org/hadoop/PoweredBy

Getting Your Own Cluster

- You can download Hadoop from http://hadoop.apache.org/
- There are two maintained distributions that I am aware of:
  - Yahoo, the version tested (at large scale) and used at Yahoo: http://developer.yahoo.com/hadoop/distribution/
  - Cloudera provides a distribution that includes HDFS, MR, Pig, Hive, and HBase: http://www.cloudera.com/hadoop
- Amazon Elastic Map Reduce service allows you to use Hadoop with no hardware investment: http://aws.amazon.com/elasticmapreduce/

Next Steps

- Read on the online documentation: http://hadoop.apache.org/
- On line tutorials
  - From Yahoo, http://developer.yahoo.com/hadoop/tutorial/
  - From Cloudera, http://www.cloudera.com/hadoop-training
- Several books available, search at your favorite bookstore
- Join the mailing lists

Credits

- I stole slides used in previous Hadoop presentations from both Owen O'Malley (Map-Reduce architect for Yahoo and VP of Apache for Hadoop) and Tom White (Hadoop PMC member, Cloudera employee, and author of *Hadoop, the Definitive Guide*)

Questions

Questions