Apache Hadoop

Large scale data processing

Speaker: Isabel Drost

Isabel Drost

Nighttime:

Co-Founder Apache Mahout.
Organizer of Berlin Hadoop Get Together.

Daytime:

Software developer

Hello FOSDEM visitors!

Hello FOSDEM visitors!

How many know Hadoop?

Hello FOSDEM visitors!

How many Hadoop users?

How many nodes?

Hello FOSDEM visitors!

Hello FOSDEM visitors!

Zookeeper?

Hello FOSDEM visitors!

Hive?

HBase?

Hello FOSDEM visitors!

Hello FOSDEM visitors!

Pig?

Hello FOSDEM visitors!

Lucene?

Solr?

Hello FOSDEM visitors!

Hello FOSDEM visitors!

Mahout?

Agenda

- Collecting and storing data.
- Analysing data.
- Tour of Hadoop.
- Hadoop ecosystem.

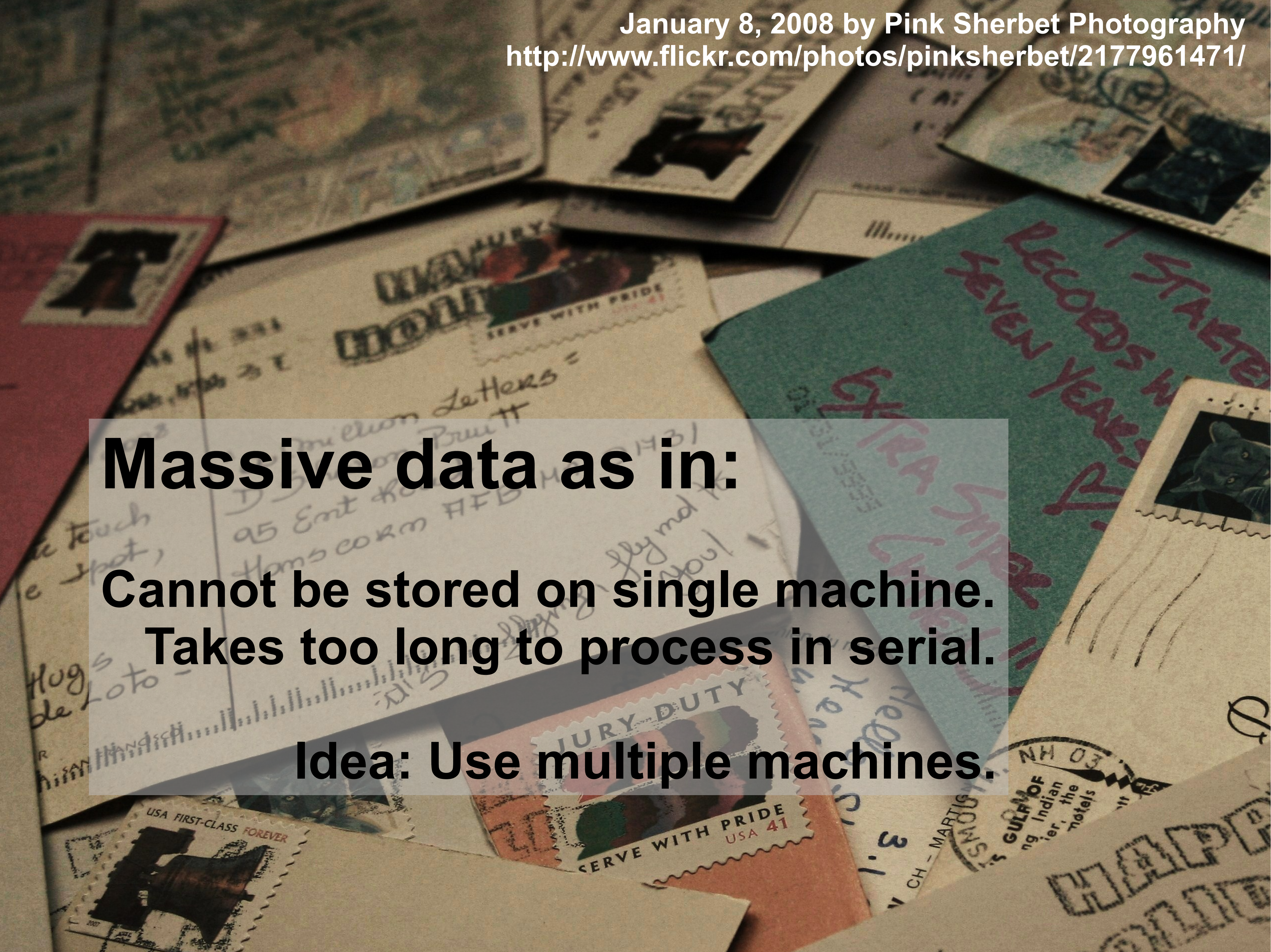
Collecting and storing data.

By Lab2112, http://www.flickr.com/photos/lab2112/462388595/

Data storage options

- Structured, relational.
  - Customer data.
  - Bug database.

By bareform, http://www.flickr.com/photos/bareform/2483573213/



January 8, 2008 by Pink Sherbet Photography
http://www.flickr.com/photos/pinksherbet/2177961471/

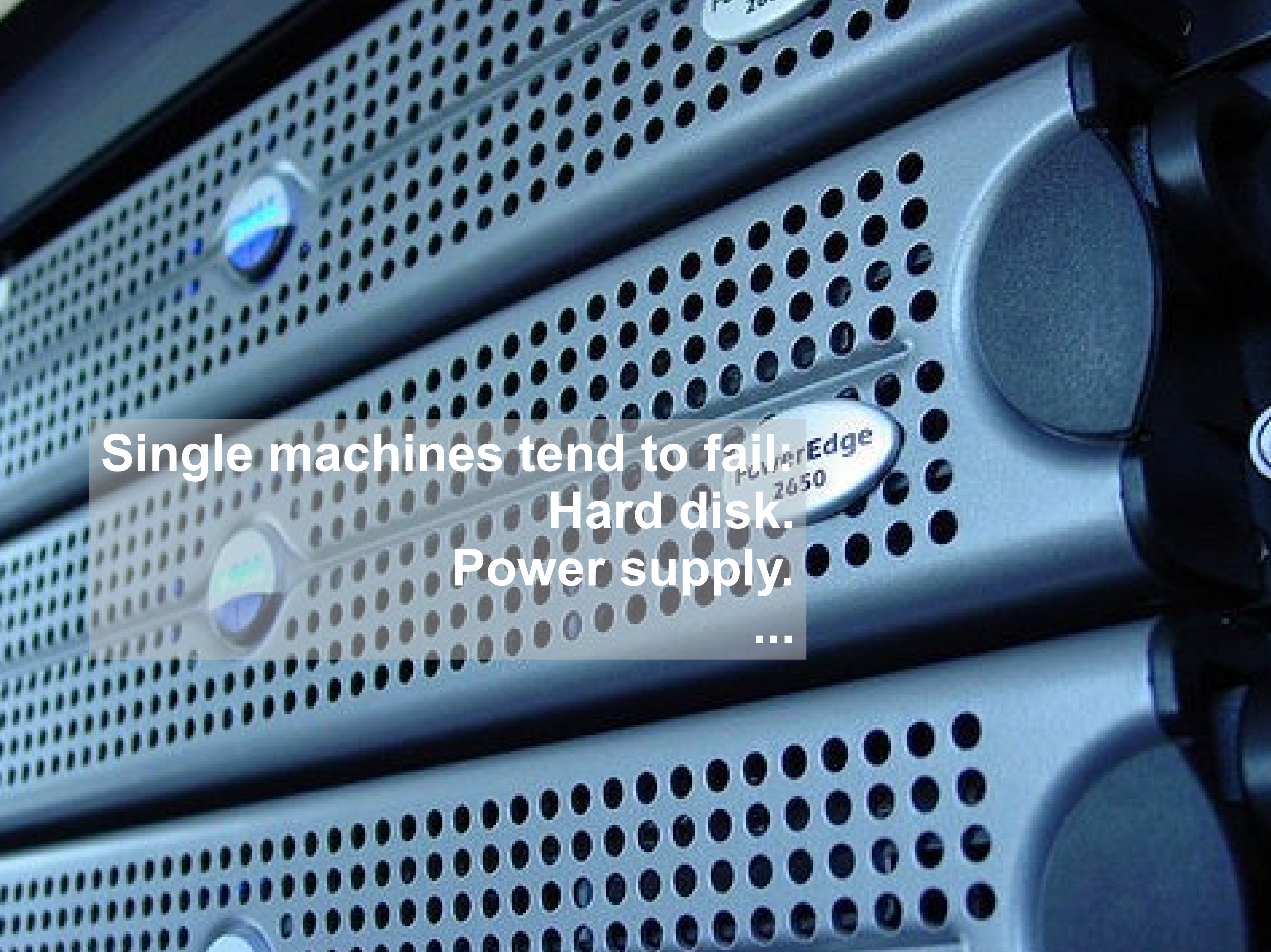
Massive data as in:

Cannot be stored on single machine.

Takes too long to process in serial.

Idea: Use multiple machines.

Challenges when scaling out.



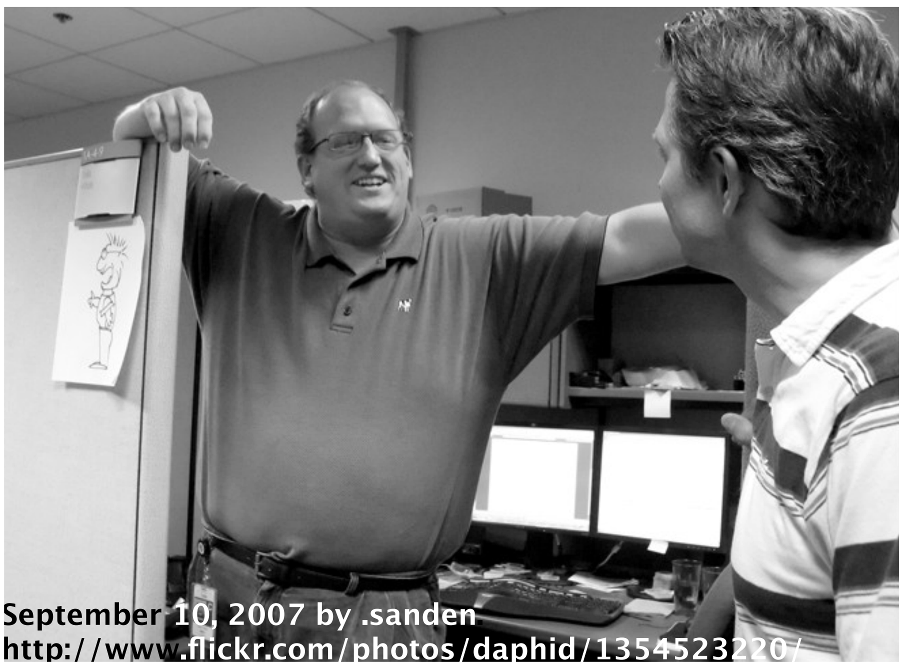
Single machines tend to fail:
Hard disk.
Power supply.
...


More machines – increased failure probability.
January 11, 2007, skreuzer
http://www.flickr.com/photos/skreuzer/354316053/

Requirements

- Built-in backup.
- Built-in failover.

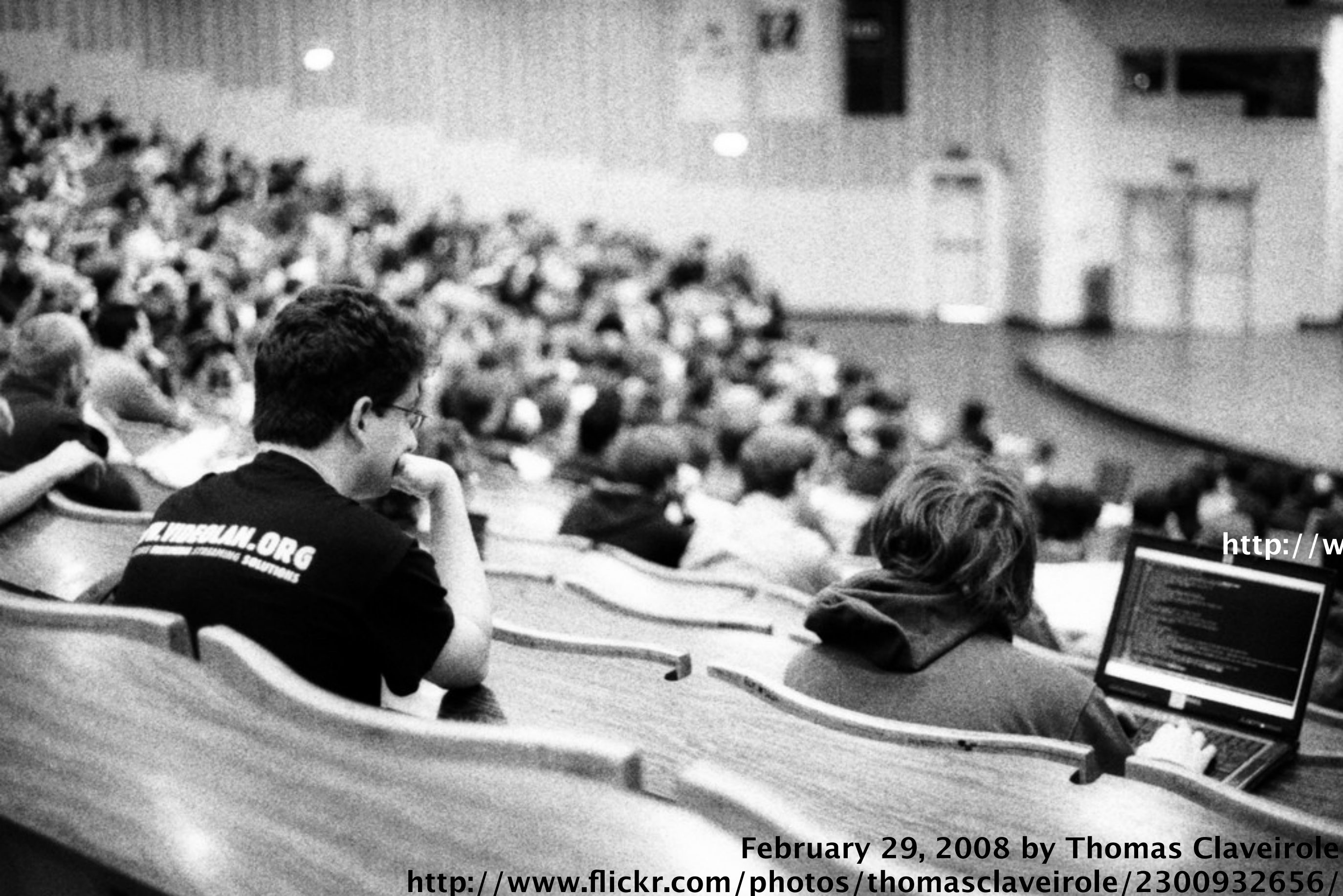
Typical developer


September 10, 2007 by .sanden
http://www.flickr.com/photos/daphid/1354523220/

- Has never dealt with large (petabytes) amount of data.
- Has no thorough understanding of parallel programming.
- Has no time to make software production ready.

Requirements

- Built-in backup.
- Built-in failover.
- Easy to use.
- Parallel on rails.



February 29, 2008 by Thomas Claveirole
http://www.flickr.com/photos/thomasclaveirole/2300932656/

http://www.flickr.com/photos/jaaronfarr/3384940437/
March 25, 2009 by jaaron

http://www.flickr.com/photos
/jaaronfarr/3385756482/
March 25, 2009 by jaaron

May 1, 2007 by danny angus
http://www.flickr.com/photos/killerbees/479864437/

Requirements

- Built-in backup.
- Built-in failover.

- Easy to use.
- Parallel on rails.
- Active development.

http://www.flickr.com/photos/cspowers/282944734/ by cspowers on October 29, 2006

Requirements

- Built-in backup.
- Built-in failover.
- Easy to administrate.
- Single system.
- Easy to use.
- Parallel on rails.
- Active development.

Easy distributed programming.

Well known in industry and research.

Scales well beyond 1000 nodes.

Some history.

Feb '03 first Map Reduce library @ Google

Oct '03 GFS Paper

Dec '04 Map Reduce paper

Dec '05 Doug reports that nutch uses map reduce

Feb '06 Hadoop moves out of nutch

Apr '07 Y! running Hadoop on 1000 node cluster

Jan '08 Hadoop made an Apache Top Level Project

Petabyte sorting benchmark

Bytes	Nodes
500,000,000,000	1406
1,000,000,000,000	1460
100,000,000,000,000	3452
1,000,000,000,000,000	3658

Replication	Time
1	59 seconds
1	62 seconds
2	173 minutes
2	975 minutes

Per node: 2 quad core Xeons @ 2.5ghz, 4 SATA disks, 8G RAM (upgraded to 16GB before petabyte sort), 1 gigabit ethernet.

Per Rack: 40 nodes, 8 gigabit ethernet uplinks.

Hadoop assumptions

Assumptions:

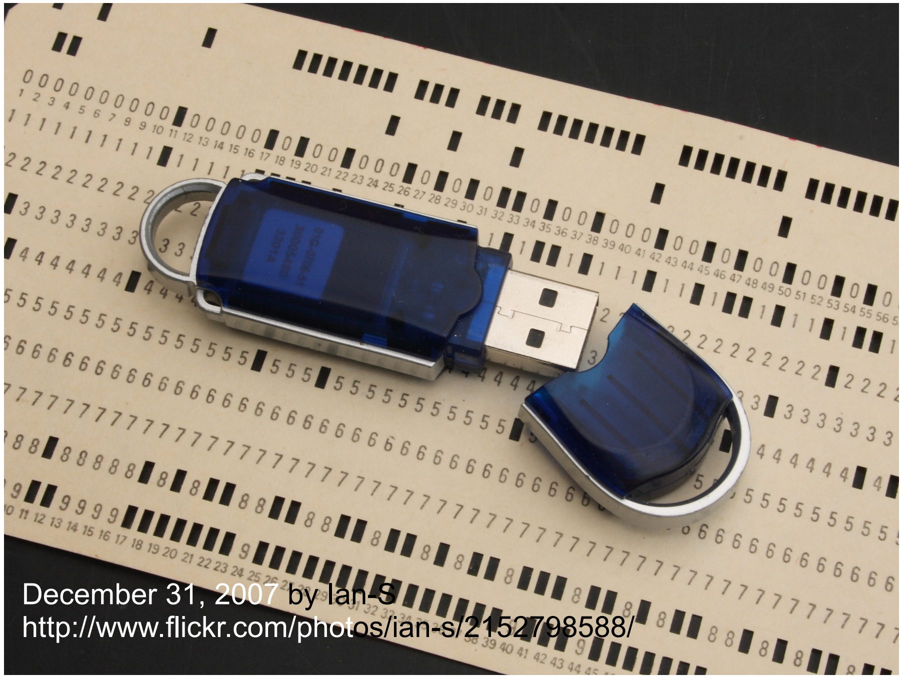
Data to process does not fit on one node.
Each node is commodity hardware.
Failure happens.

Ideas:

Distribute filesystem.
Built in replication.
Automatic failover in case of failure.

Assumptions:

Distributed computation is easy.
Moving computation is cheap.
Moving data is expensive.


December 31, 2007 by Ian-S
http://www.flickr.com/photos/ian-s/2152798588/

Ideas:

Move computation to data.
Write software that is easy to distribute.

Assumptions:

Systems run on spinning hard disks.
Disk seek >> disk scan.

August 27, 2007 by Stuart Bryant
http://www.flickr.com/photos/stuartbryant/1249649001/

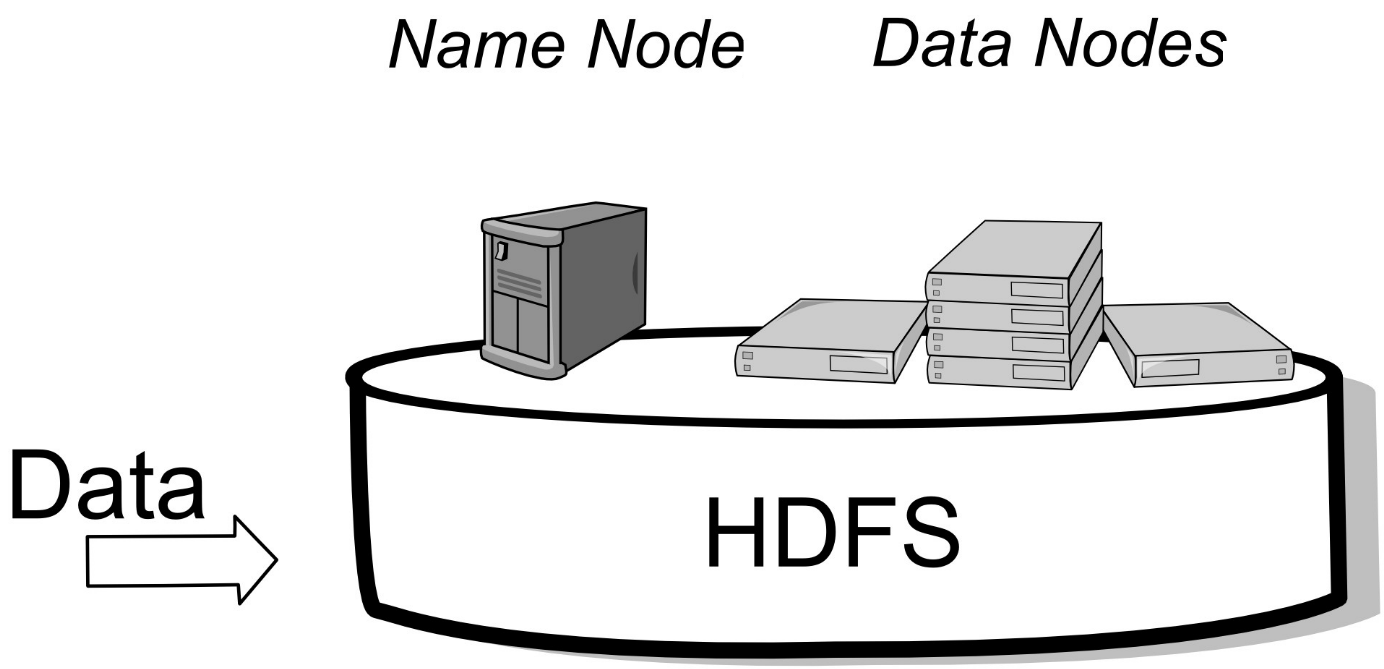
Ideas:

Improve support for large files.
File system API makes scanning easy.

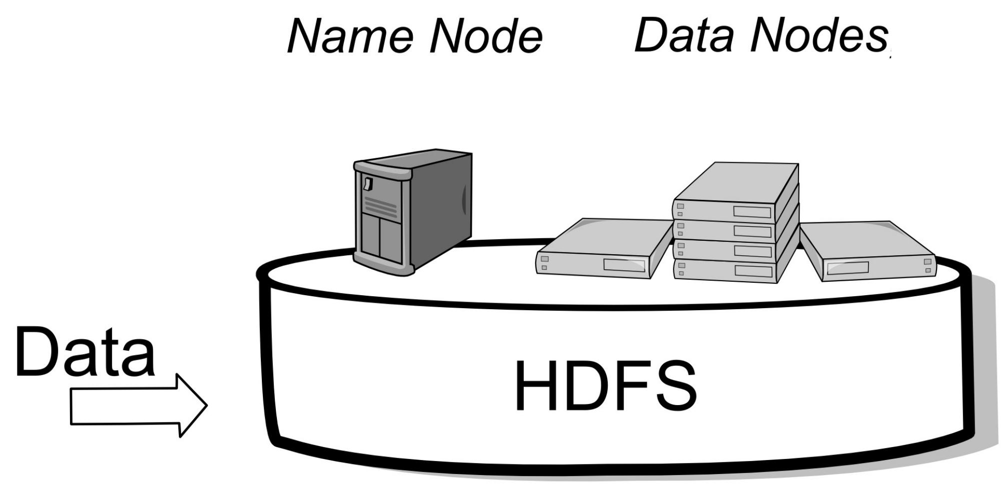
hadoop
HDFS

hadoop
Map Reduce

HDFS building blocks



(Graphics: Thanks to Thilo.)



NameNode

- Stores file meta data.
- In memory.
- Block-node mapping.

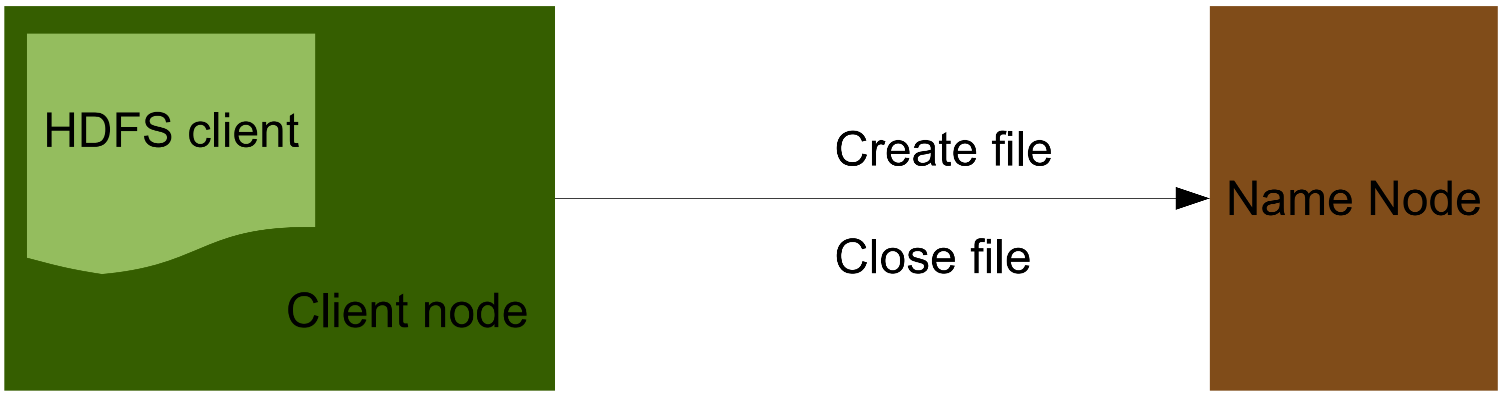
DataNode

- Stores file contents.
- On disk.
- Block-Id to disk.

Anatomy of a file write

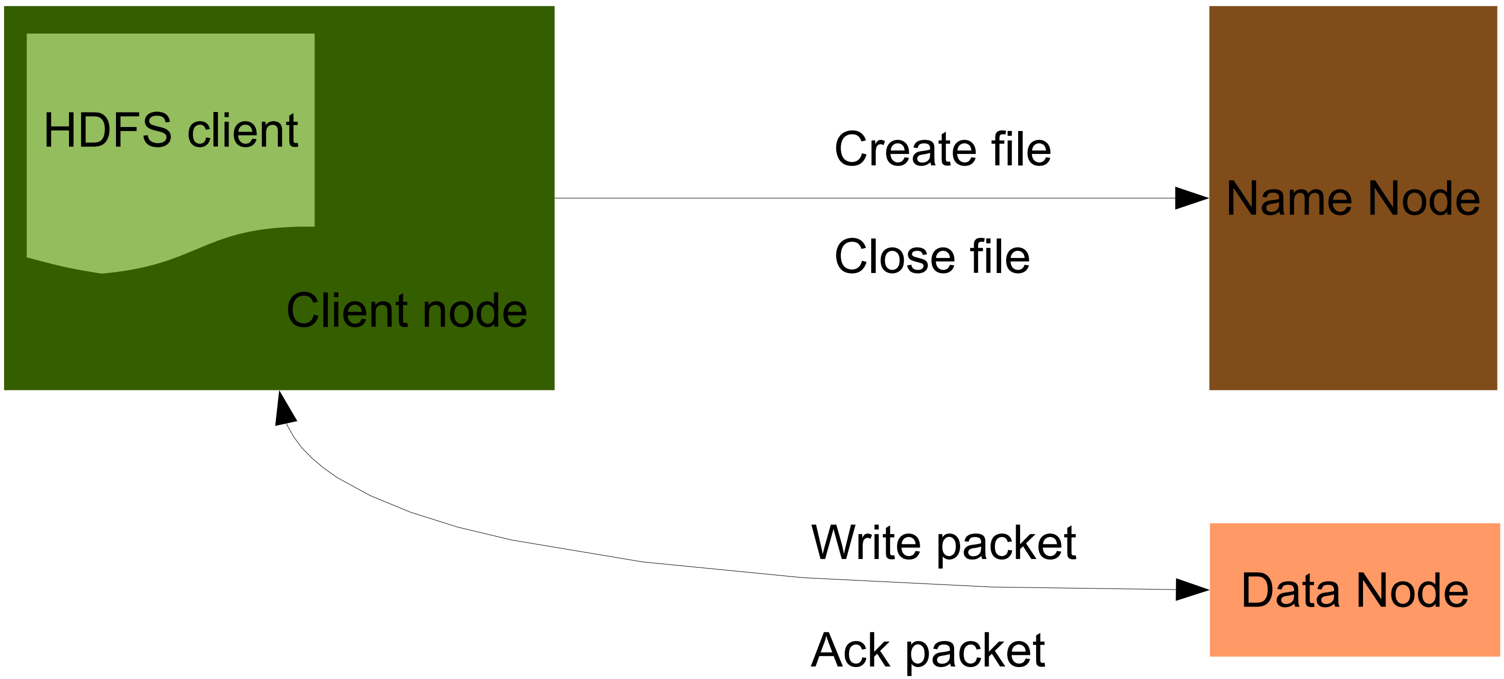
Slide inspired by: “Hadoop – The definitive guide”, Tom White, O'Reilly

Anatomy of a file write



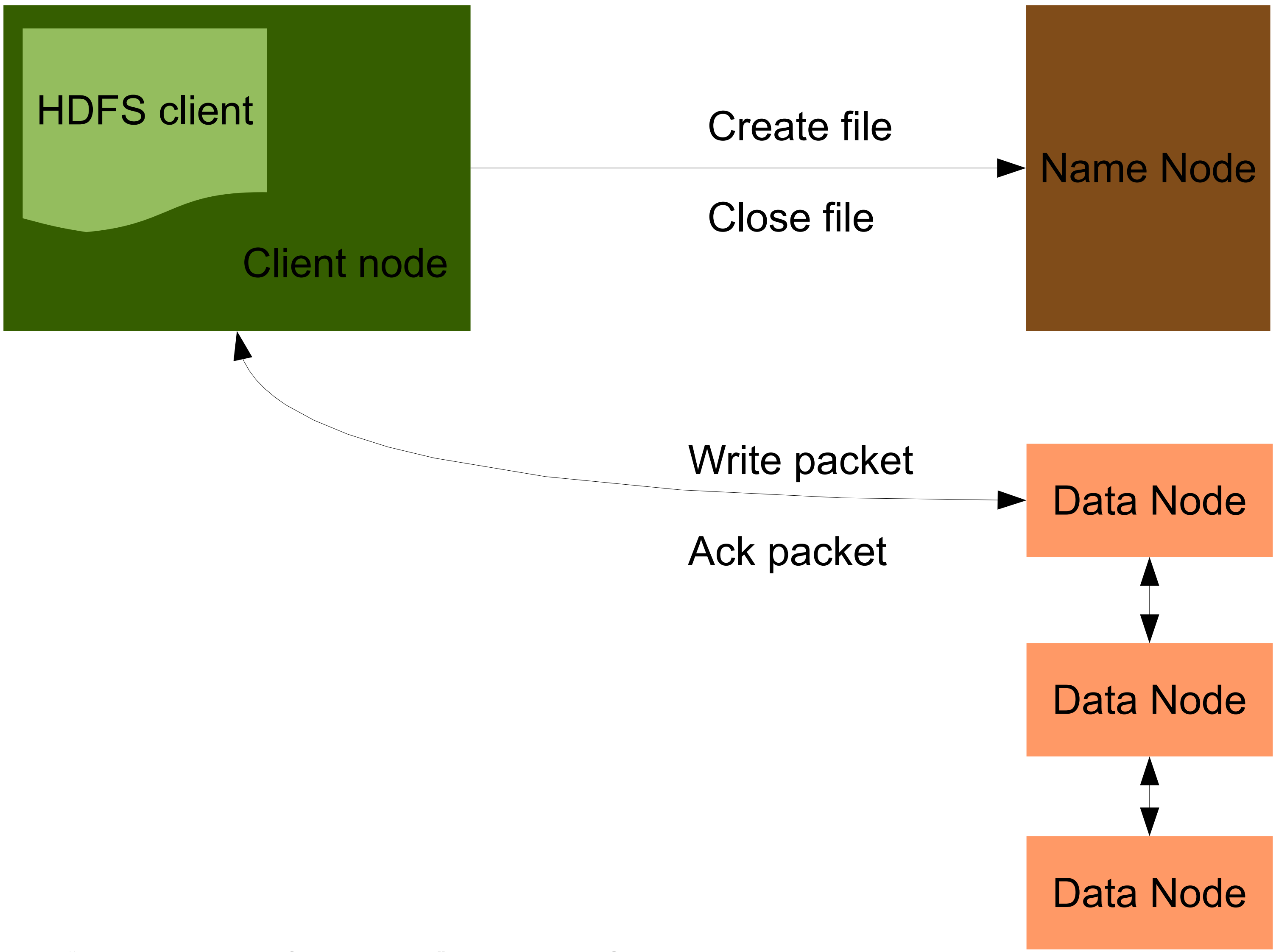
Slide inspired by: “Hadoop – The definitive guide”, Tom White, O'Reilly

Anatomy of a file write



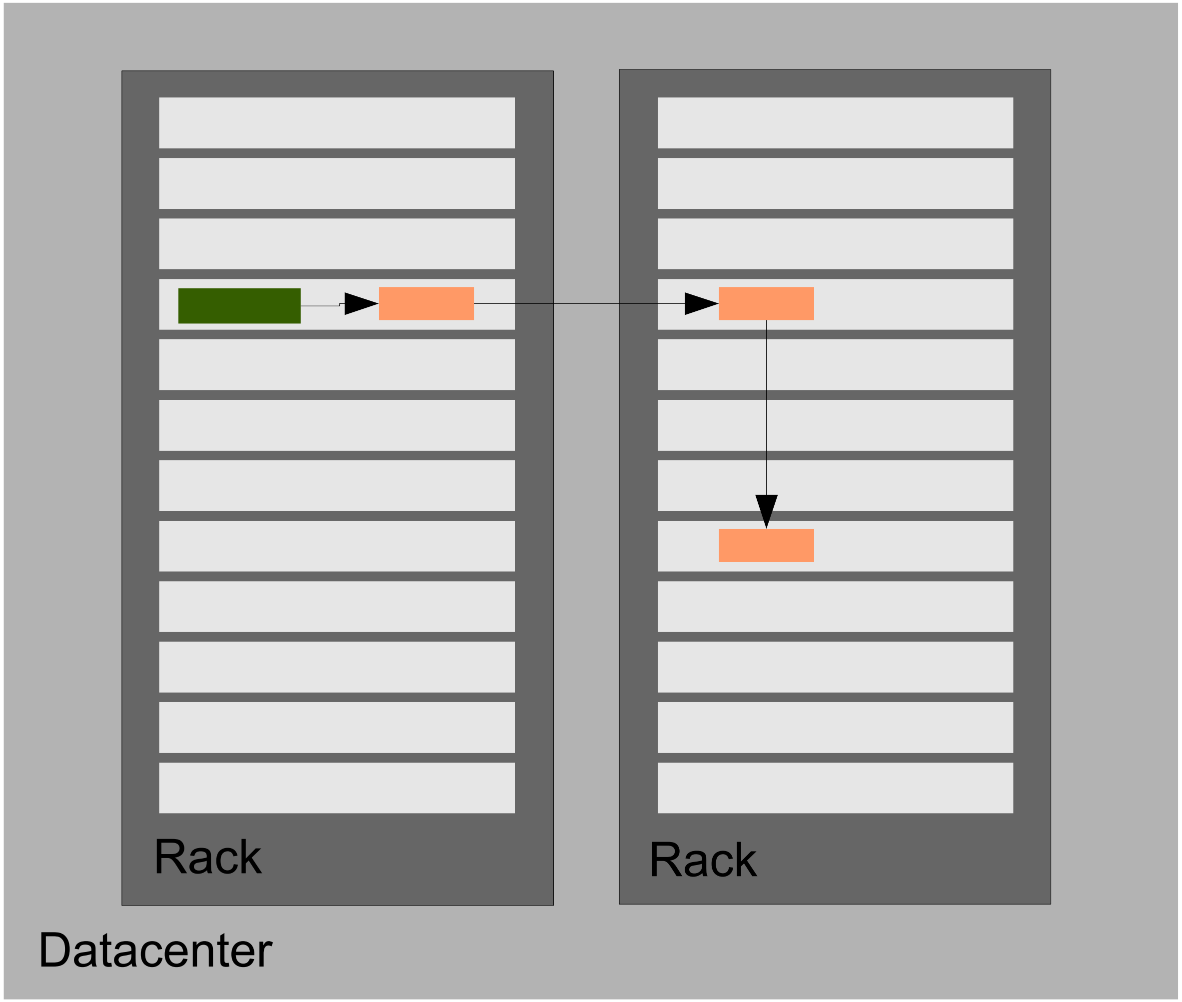
Slide inspired by: “Hadoop – The definitive guide”, Tom White, O'Reilly

Anatomy of a file write



Slide inspired by: “Hadoop – The definitive guide”, Tom White, O'Reilly

HDFS Replication Strategy

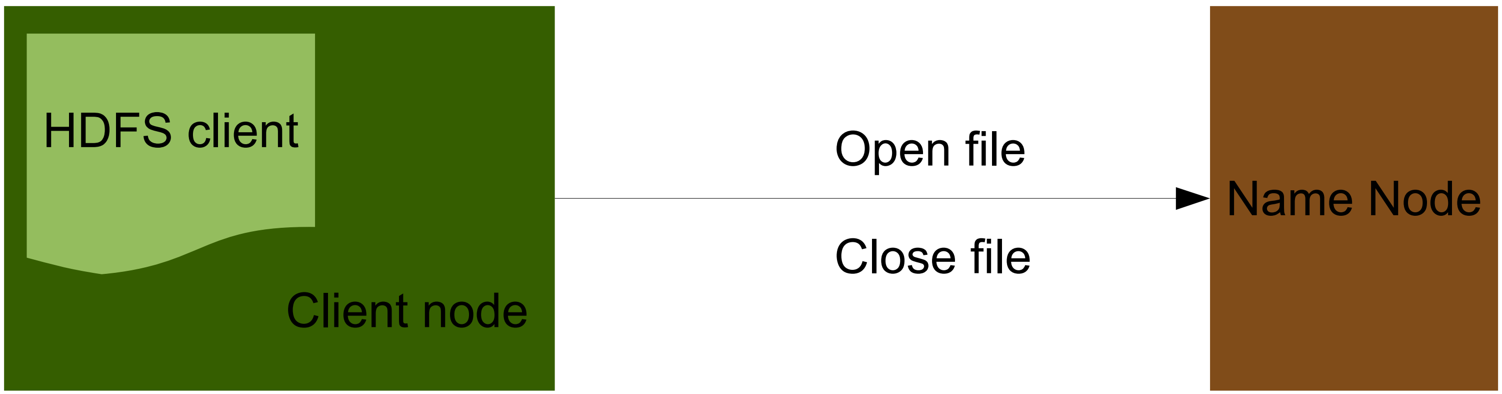


Slide inspired by: “Hadoop – The definitive guide”, Tom White, O'Reilly

Anatomy of a file read

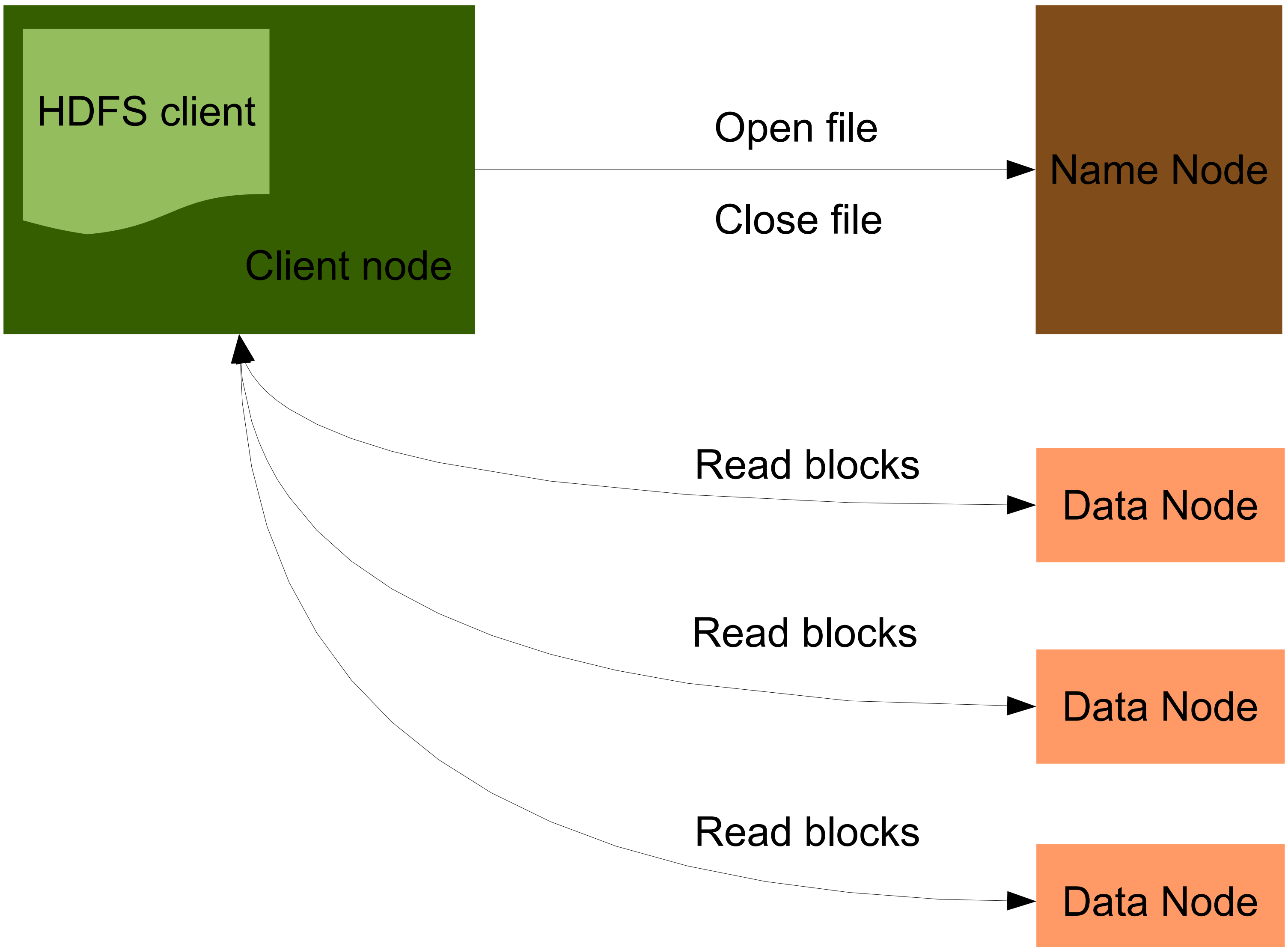
Slide inspired by: “Hadoop – The definitive guide”, Tom White, O'Reilly

Anatomy of a file read



Slide inspired by: “Hadoop – The definitive guide”, Tom White, O'Reilly

Anatomy of a file read



Slide inspired by: “Hadoop – The definitive guide”, Tom White, O'Reilly

Analyse and understand your data.

Map/Reduce by example

```
<?xml version="1.0" encoding="UTF-8"?>
<opml version="1.0" >
 <head>
  <text></text>
 </head>
 <body>
  <outline htmlUrl="http://eventseer.net" title="EventSeer - A Digital Library of Call for Papers" useC
alDefault" version="RSS" type="rss" xmlUrl="http://eventseer.net/feeds/main/rss.xml" id="312053548" tex
tseer.net" />
  <outline isOpen="false" id="669809145" text="Silent" >
   <outline htmlUrl="http://www.theserverside.com" title="TheServerSide.com: Patterns" useCustomFetchIn
ersion="RSS" type="rss" xmlUrl="http://www.theserverside.com/rss/theserverside-j2eepatterns-rss2.xml" i
taining up-to-date news, discussions, patterns, resources, and media" />
   <outline htmlUrl="http://chadwa.wordpress.com" title="Chad's Search Blog" useCustomFetchInterval="fa
S" type="rss" xmlUrl="http://chadwa.wordpress.com/feed/" id="545368194" text="Chad's Search Blog" descr
" />
   <outline htmlUrl="http://www.find23.net/Site/Blog/Blog.html" title="My Blog" useCustomFetchInterval=
"RSS" type="rss" xmlUrl="http://www.find23.net/Site/Blog/rss.xml" id="1620106192" text="My Blog" descri
   <outline htmlUrl="http://emotion.inrialpes.fr/~dangauthier/blog" title="Yet Another Machine Learning
eMode="globalDefault" version="RSS" type="rss" xmlUrl="http://emotion.inrialpes.fr/~dangauthier/blog/fe
g" />
   <outline htmlUrl="http://ml.typepad.com/machine_learning_thoughts/" title="Machine Learning Thoughts
="globalDefault" version="RSS" type="rss" xmlUrl="http://ml.typepad.com/machine_learning_thoughts/rss.x
etical and practical aspects of Machine Learning." />
   <outline htmlUrl="http://yaroslavvb.blogspot.com/" title="Machine Learning, etc" useCustomFetchInter
ion="RSS" type="rss" xmlUrl="http://yaroslavvb.blogspot.com/feeds/posts/default" id="805998569" text="M
   <outline htmlUrl="http://ptufts.blogspot.com/" title="Pinhead's Progress" useCustomFetchInterval="fa
S" type="rss" xmlUrl="http://ptufts.blogspot.com/feeds/posts/default" id="1019393988" text="Pinhead's P
   <outline htmlUrl="http://resnotebook.blogspot.com/" title="Misc Research Stuff" useCustomFetchInterv
on="RSS" type="rss" xmlUrl="http://resnotebook.blogspot.com/feeds/posts/default" id="216193226" text="M
   <outline htmlUrl="http://absolutely-regular.blogspot.com/" title="Absolutely Regular" useCustomFetch
 version="RSS" type="rss" xmlUrl="http://absolutely-regular.blogspot.com/feeds/posts/default" id="17850
   <outline htmlUrl="http://atomai.blogspot.com/" title="Data Mining, Analytics and Artificial Intellig
Mode="globalDefault" version="RSS" type="rss" xmlUrl="http://atomai.blogspot.com/feeds/posts/default" i
nt in data mining, artificial intelligence, analytics, intelligent agents, semiconductors, distributing
siness Objects, Oracle, Intel, AMD, or Pentaho. Heuristic, Six Sigma, or CMM. Contractor or in-house. H
ail.com" />
```

```
isabel@h1349259:~$ more data/feeds.opml | grep -o "http://[0-9A-Za-z\-_\.]*" | s
ort | uniq --count | sort | tail
      3 http://agbs.kyb.tuebingen.mpg.de
      3 http://irgupf.com
      3 http://jeffsutherland.com
      4 http://ml.typepad.com
      4 http://weblogs.java.net
      4 http://www.gridvm.org
      4 http://yaroslavvb.blogspot.com
      5 http://feeds.feedburner.com
      6 http://blogsearch.google.com
     10 http://arxiv.org
```

```
pattern=”http://[0-9A-Za-z\-_\.]*”

grep -o "$pattern" feeds.opml     | sort          | uniq --count
```

pattern=”http://[0-9A-Za-z\-_\.]*”

grep -o "$pattern" feeds.opml

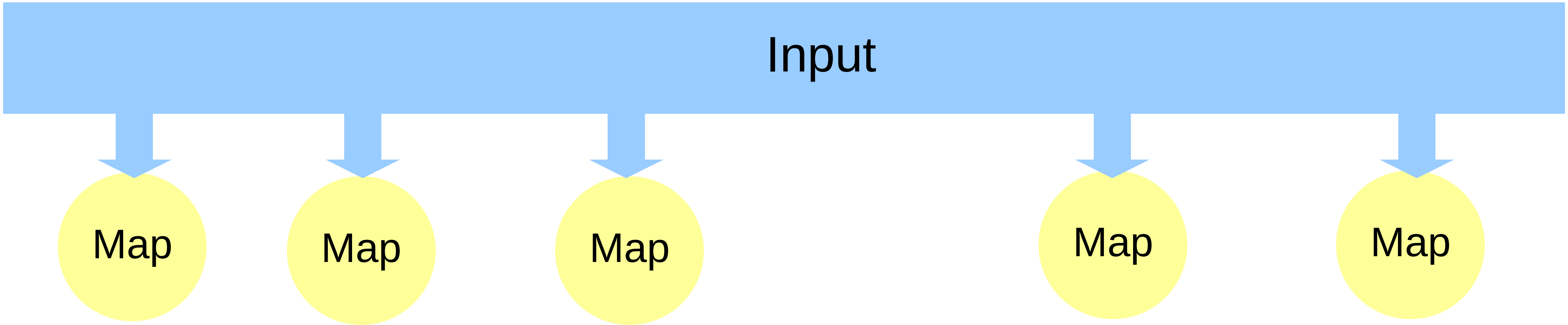
M A P

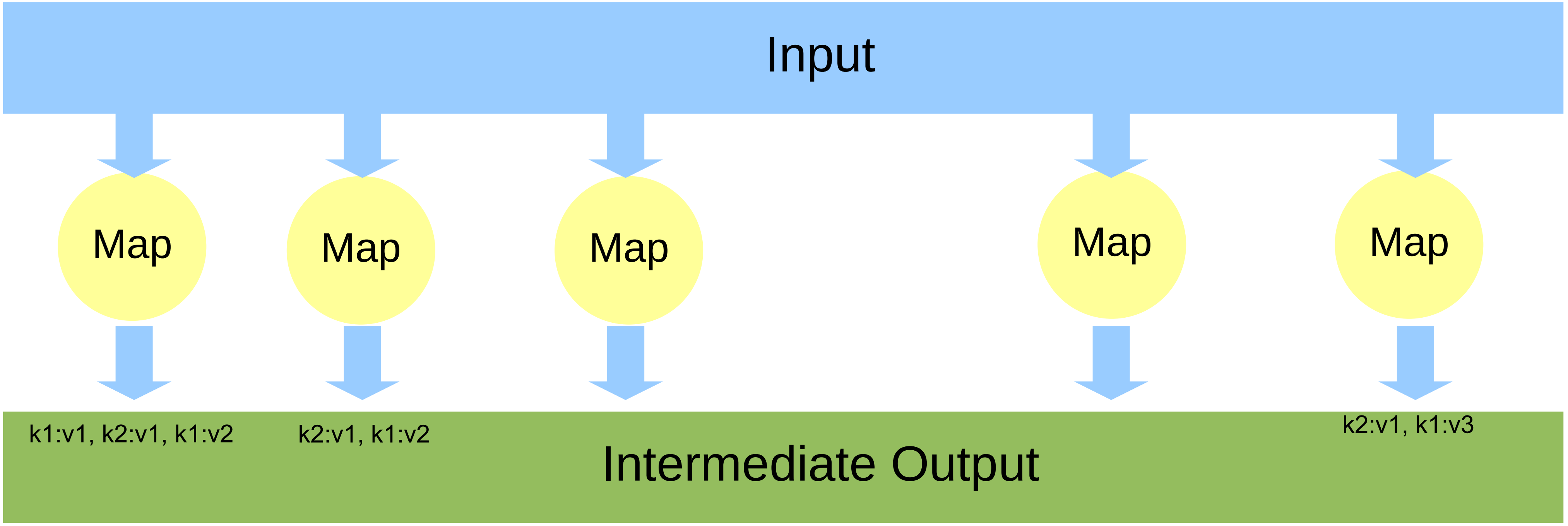
| sort

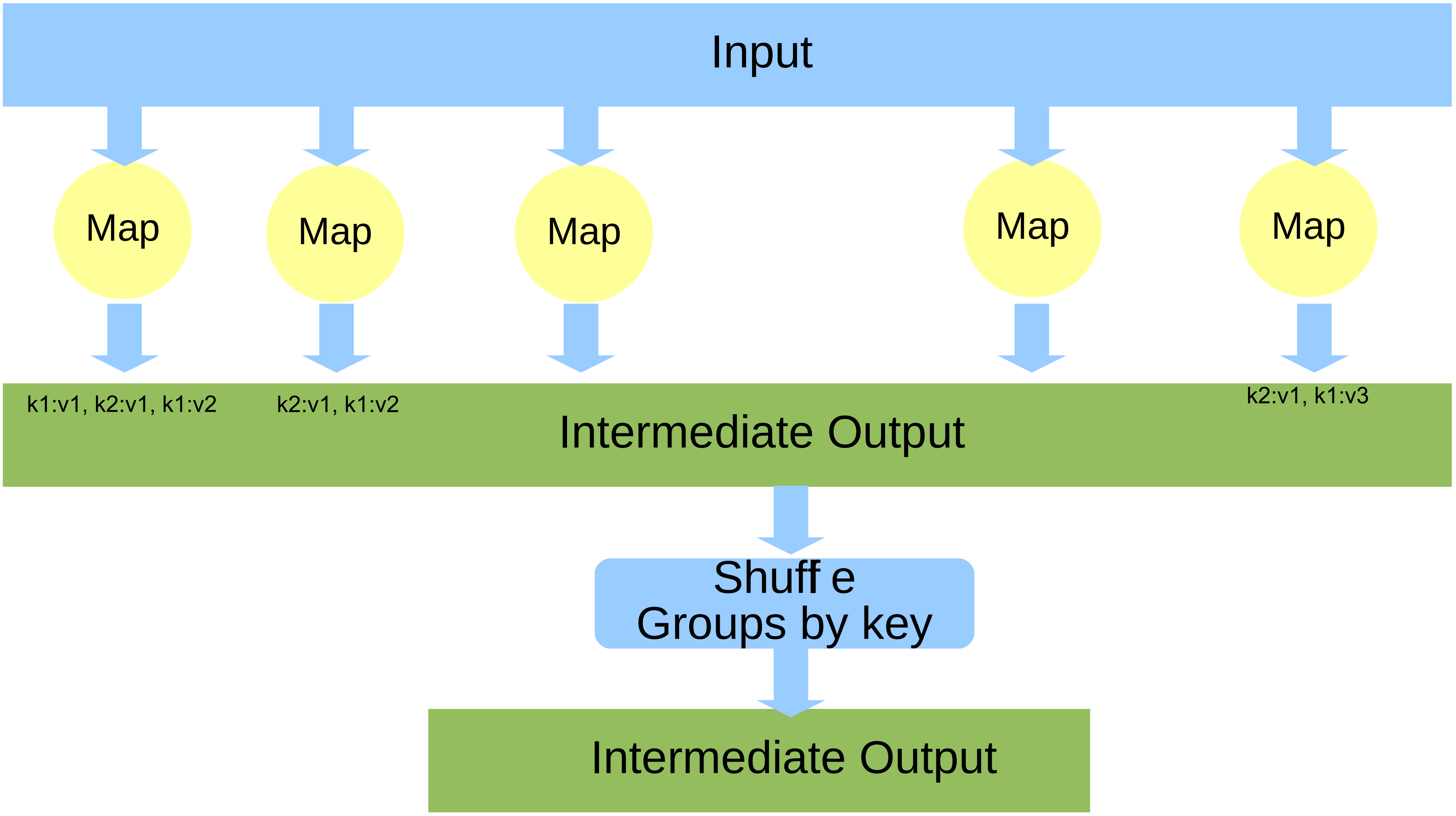
| SHUFFLE

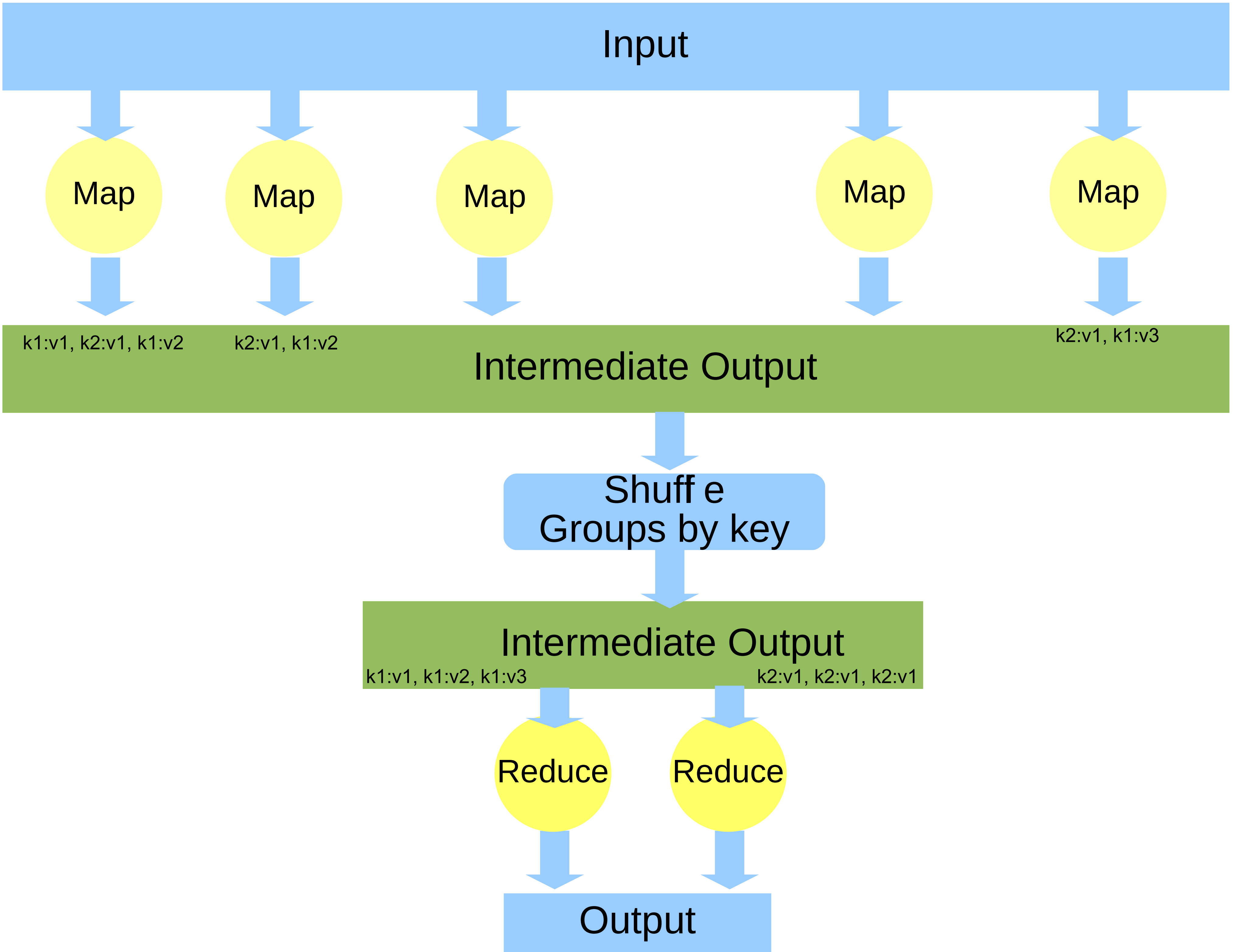
| uniq --count

| R E D U C E

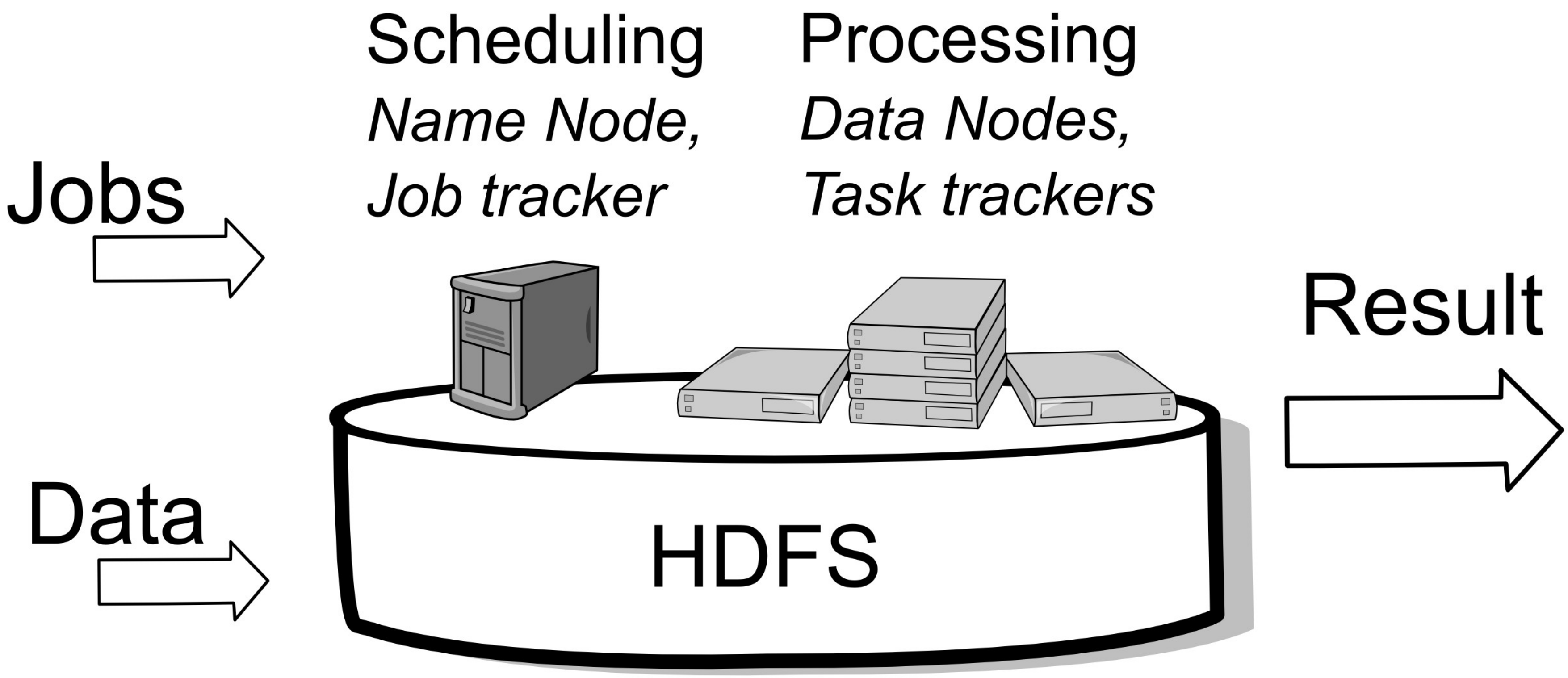

Input
Map
Map
Map
Map
Map


Input
Map
Map
Map
Map
Map
k1:v1, k2:v1, k1:v2
k2:v1, k1:v2
k2:v1, k1:v3
Intermediate Output


Input
Map
Map
Map
Map
Map
k1:v1, k2:v1, k1:v2
k2:v1, k1:v2
k2:v1, k1:v3
Intermediate Output
Shuff e
Groups by key
Intermediate Output

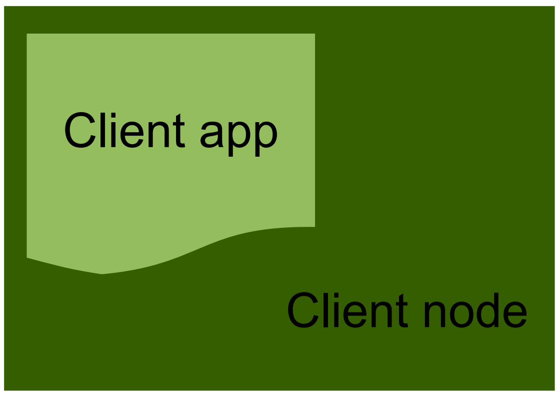

Input
Map
Map
Map
Map
Map
k1:v1, k2:v1, k1:v2
k2:v1, k1:v2
k2:v1, k1:v3
Intermediate Output
Shuff e
Groups by key
Intermediate Output
k1:v1, k1:v2, k1:v3
k2:v1, k2:v1, k2:v1
Reduce
Reduce
Output

```java
private IntWritable one = new IntWritable(1);
private Text hostname = new Text();



public void map(K key, V value, Context context) {
  String line = value.toString();
  StringTokenizer tokenizer = new StringTokenizer(line);
  while (tokenizer.hasMoreTokens()) {
    hostname.set(getHostname(tokenizer.nextToken()));
    context.write(hostname, one);
  }
}



public void reduce(K2 key, Iterable<V2> values,
  OutputCollector<K3, V2> output) {
  int sum = 0;
  while (values.hasNext()) {
    sum += values.next().get();
  }
  output.collect(key, new IntWritable(sum));
}
```



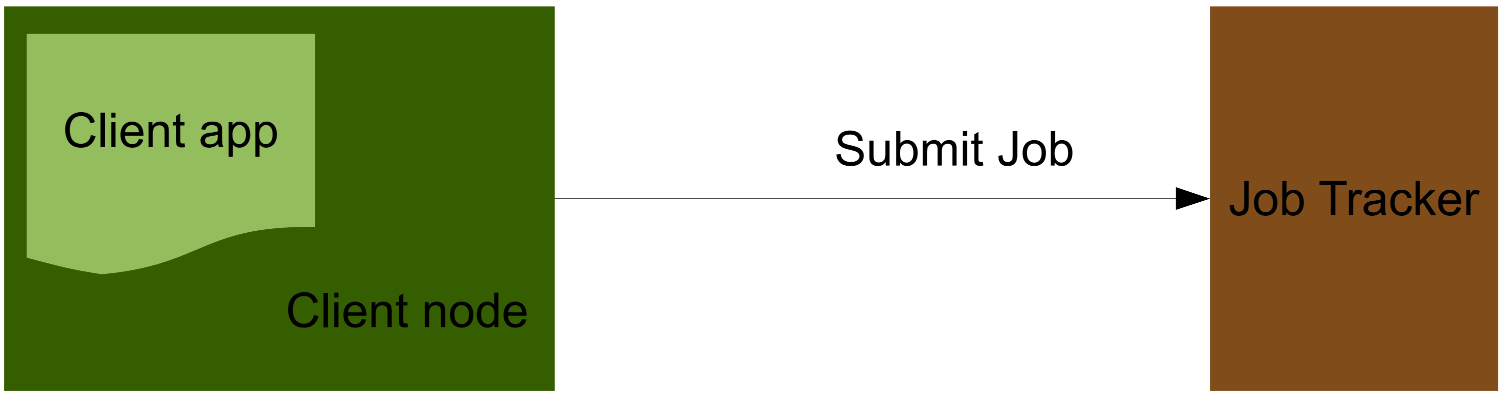
(Graphics: Thanks to Thilo.)

Anatomy of a map/reduce job



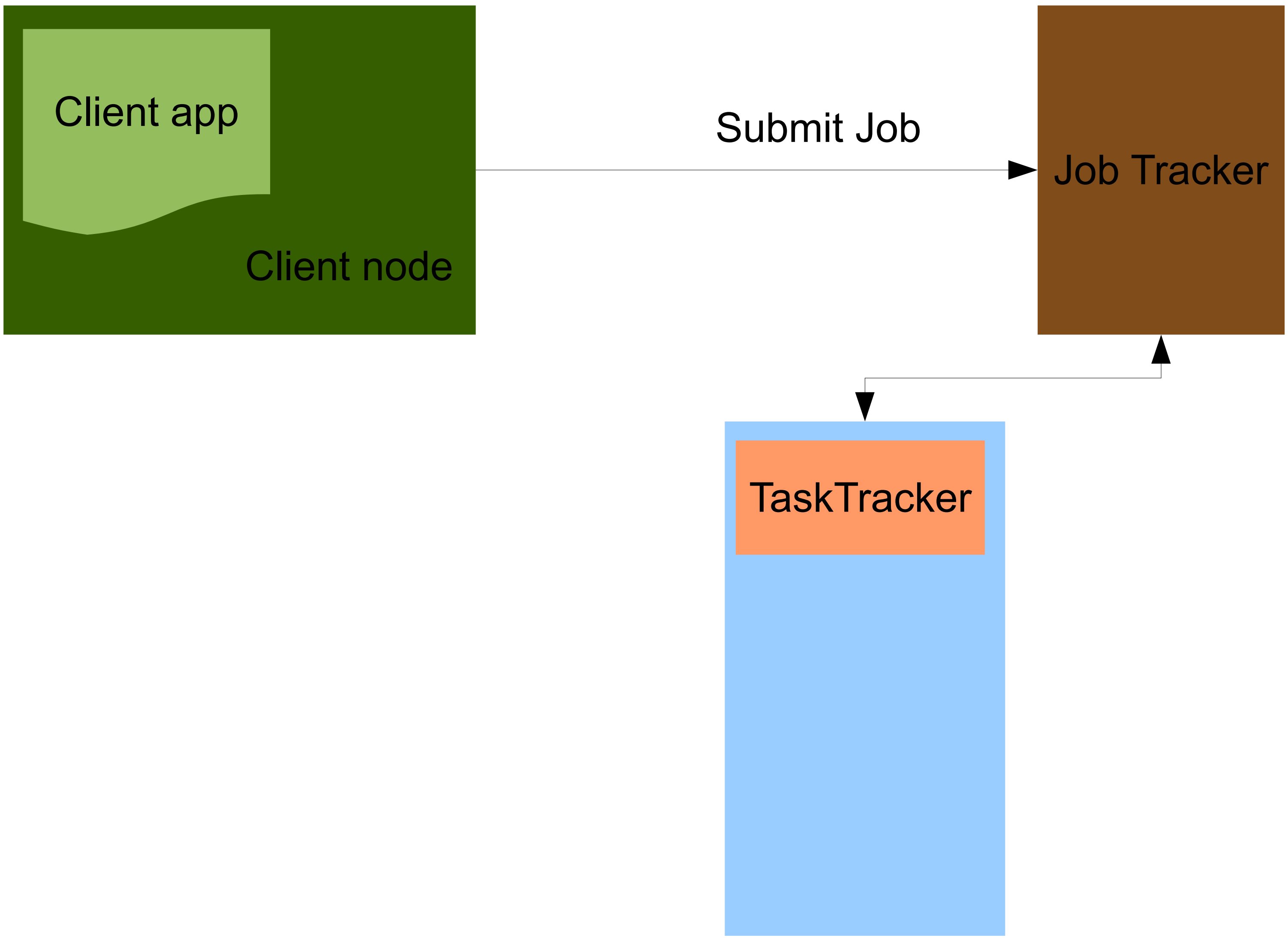
Slide inspired by: “Hadoop – The definitive guide”, Tom White, O'Reilly

Anatomy of a map/reduce job



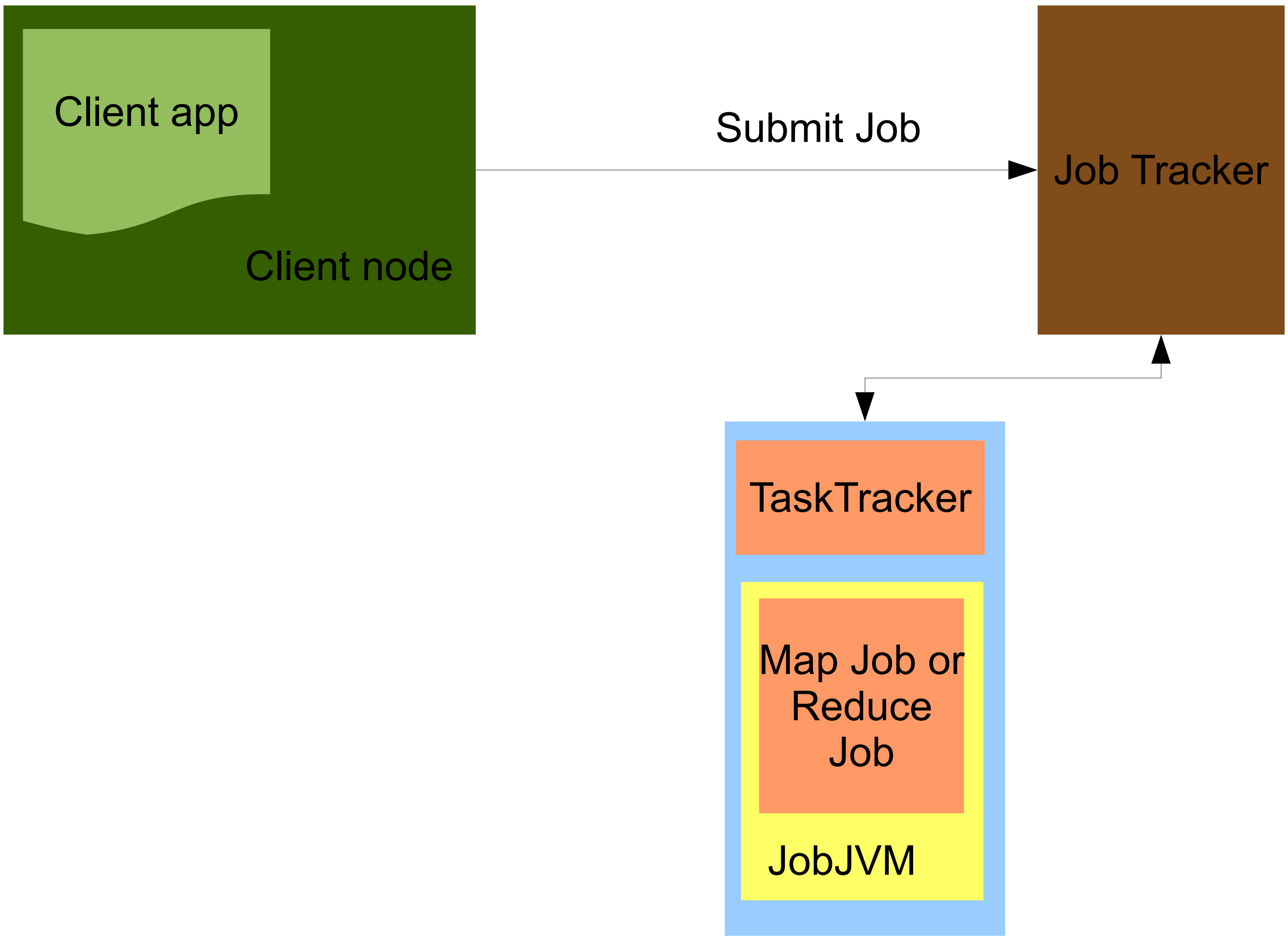
Slide inspired by: "Hadoop – The definitive guide", Tom White, O'Reilly

Anatomy of a map/reduce job



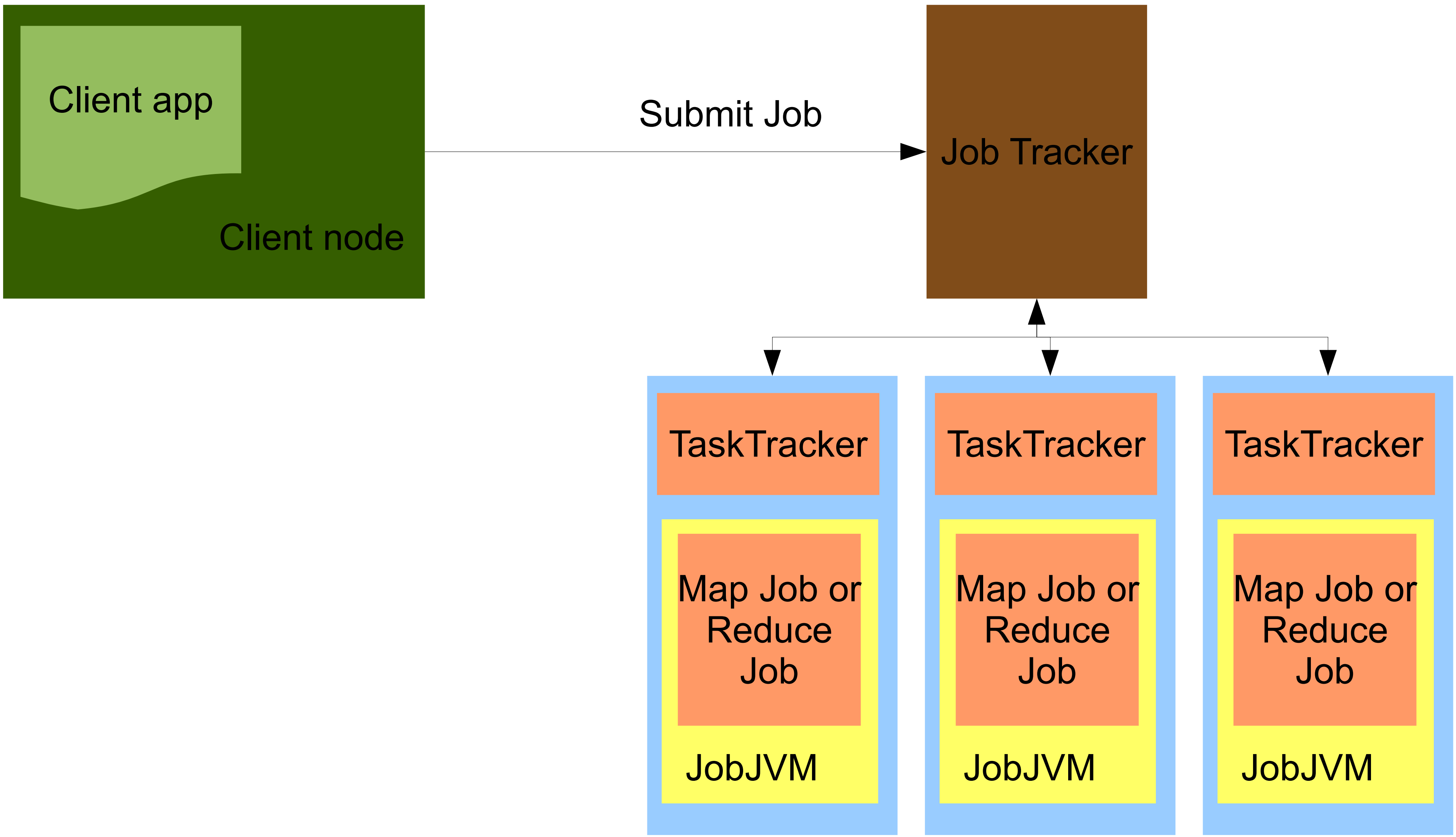
Slide inspired by: “Hadoop – The definitive guide”, Tom White, O'Reilly

Anatomy of a map/reduce job



Slide inspired by: “Hadoop – The definitive guide”, Tom White, O'Reilly

Anatomy of a map/reduce job



Slide inspired by: “Hadoop – The definitive guide”, Tom White, O'Reilly

Requirements to get started

March 14, 2009 by Artful Magpie
http://www.flickr.com/photos/kmtucker/3355551036/

Amazon Elastic Compute Cloud (Amazon EC2)

Amazon EC2 Details

- **EC2 Overview**
- FAQs
- Amazon EC2 SLA
- EC2 Instance Types

Amazon Elastic Compute Cloud (Amazon EC2) is a web service that provides resizable compute capacity in the cloud. It is designed to make web-scale computing easier for developers.

Amazon EC2's simple web service interface allows you to obtain and configure capacity with minimal friction. It provides you with complete control of your computing resources and lets you run on Amazon's proven computing environment. Amazon EC2 reduces the time required to obtain

amazon web services™

Sign in to the AWS Management Console | Create an AWS Account

AWS | Products | Developers | Community | Support | Account

Products & Services

Amazon Elastic Compute Cloud (Amazon EC2)

Amazon EC2 Details

- **EC2 Overview**
- FAQs
- Amazon EC2 SLA
- EC2 Instance Types

Amazon Elastic Compute Cloud (Amazon EC2) is a web service that provides resizable compute capacity in the cloud. It is designed to make web-scale computing easier for developers.

Amazon EC2's simple web service interface allows you to obtain and configure capacity with minimal friction. It provides you with complete control of your computing resources and lets you run on Amazon's proven computing environment. Amazon EC2 reduces the time required to obtain

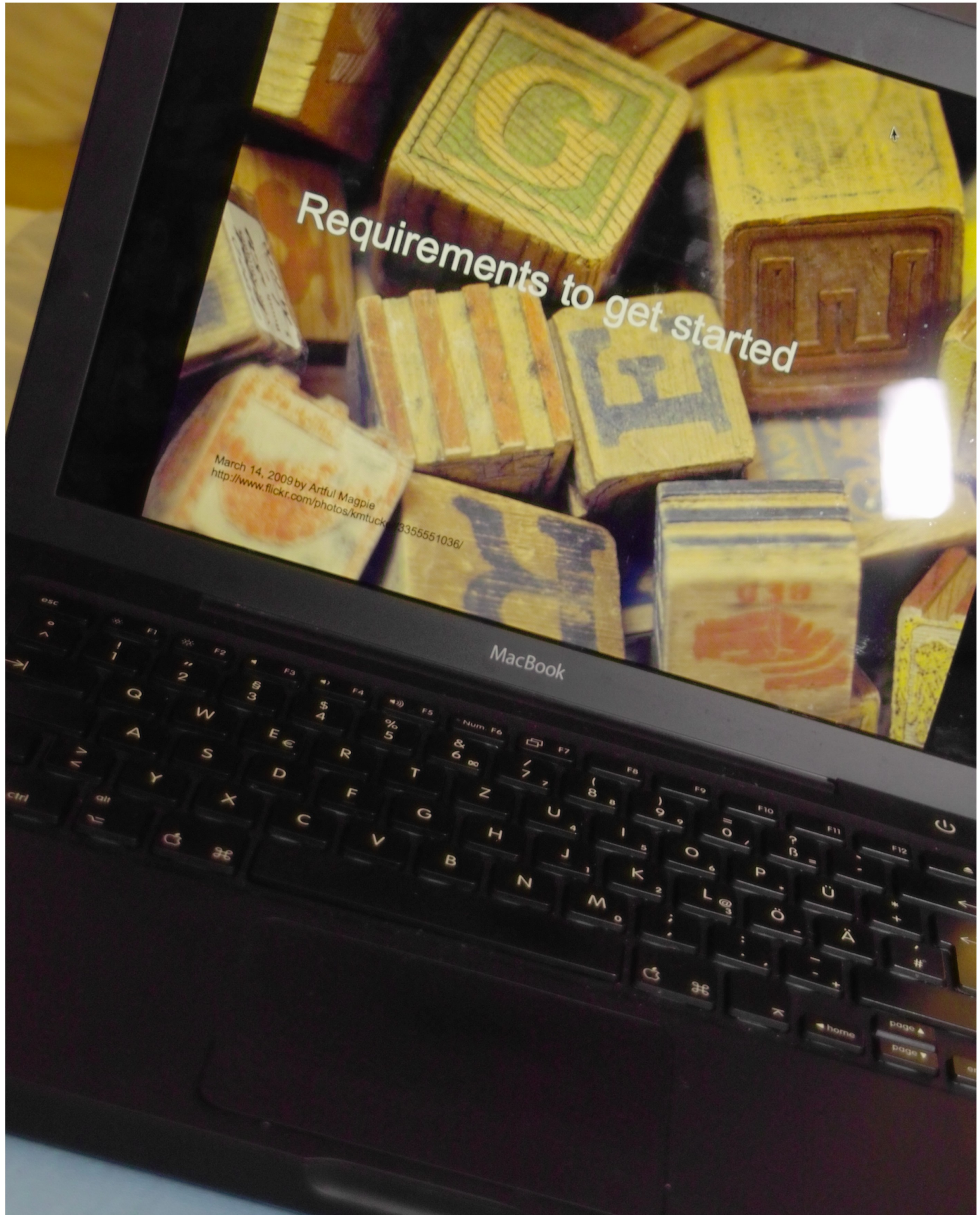
Sign Up For Amazon EC2

Amazon Elastic MapReduce

Amazon Elastic MapReduce is a web service that enables businesses, researchers, data analysts, and developers to easily and cost-effectively process vast amounts of data. It utilizes a hosted Hadoop framework running on the web-scale infrastructure of Amazon Elastic Compute Cloud (Amazon EC2) and Amazon Simple Storage Service (Amazon S3).

Using Amazon Elastic MapReduce, you can instantly provision as much or as little capacity as you like to perform data-intensive tasks for applications such as web indexing, data mining, log file analysis, machine learning, financial

(Thanks to Thilo for helping set up the cluster, Thanks to packet and masq for two of the three machines.)


Requirements to get started
March 14, 2009 by Artful Magpie
MacBook

http://www.flickr.com/photos/87106931@N00/3835231300/
By mhobl

Up next.

- In 0.21:
  - append/sync in HDFS
  - more advanced task schedulers
- In 0.22:
  - security
  - avro-based rpc for cross-version rpc compatibility
  - symbolic links
  - federated NameNodes


Web | Images | Video | Local | Shopping | more
hadoop
Search
Options
Customize
YAHOO!
Also try: apache hadoop, hadoop yahoo, hadoop api, More...
Welcome to Apache Hadoop Core!
Scalable: Hadoop can reliably store and process petabytes. ... Reliable: Hadoop automatically maintains multiple copies of data and ...
hadoop.apache.org/core - Cached
Welcome to Apache Hadoop!
The Apache Hadoop project develops open-source software for reliable, scalable, distributed computing.
hadoop.apache.org - Cached
Hadoop - Wikipedia, the free encyclopedia
Architecture | Prominent... | Hadoop on... | Hadoop with...
Apache Hadoop is a free Java software framework that supports data intensive distributed applications. It enables applications to work with thousands of nodes and petabytes of data. Hadoop was inspired by Google's MapReduce...
facebook
Email
Facebook helps you connect and share with the people in your life.
Sign Up
last.fm
Music Videos Radio Events Charts
Music
Search
Last.fm recommends music, videos and concerts based on what you listen to.
For example, people who like Justin Timberlake also like *NSYNC, Timbaland and The Pussycat Dolls. Who do you like?
Type an artist
Go
Or try Alexisonfire, Röyksopp, Ensiferum or Everything but the Girl.
New!
Home About A9 Product Search Clickriver OpenSearch Jobs (we're hiring)
A9
Innovations in Search Technologies™
A9.com helps people find what they want on the world's leading e-commerce sites.
Give it a try
Search
OpenSearch Product search
krugle
SELECT
Open Source Code
Open Source Projects
SCM Comments
[Clear Filters] [Advanced Search]
Language:
Found in:
Project:
hadoop
Search
All
Any area
Results
Code Search for hadoop
Code Files 1-10 (out of about 979 matching files)
SFHadoopException.java | SmartFrog | LGPL-2.1
34 * extract information from Hadoop classes (and helper libraries)
37 public class SFHadoopException extends SmartFrogExce
38 public static final String CONFIGURATION = "configur
39 public static final String SMARTFROG_DUMP_CONF = "sm
HadoopUtils.java | SmartFrog | LGPL-2.1
27 * Created 28-May-2008 15:22:20
30 public class HadoopUtils {
32 private HadoopUtils() {
HadoopClusterTest.java | SmartFrog | LGPL-2.1
25 * Created 05-Jan-2009 16:12:47
The New York Times
Friday, June 26, 2009 Last Update: 7:17 AM ET
deepdyve BETA
Widgets New! | Login
for Research
Enter a sentence, or cut and paste a paragraph
Refine
Searc

Hadoop ecosystem.

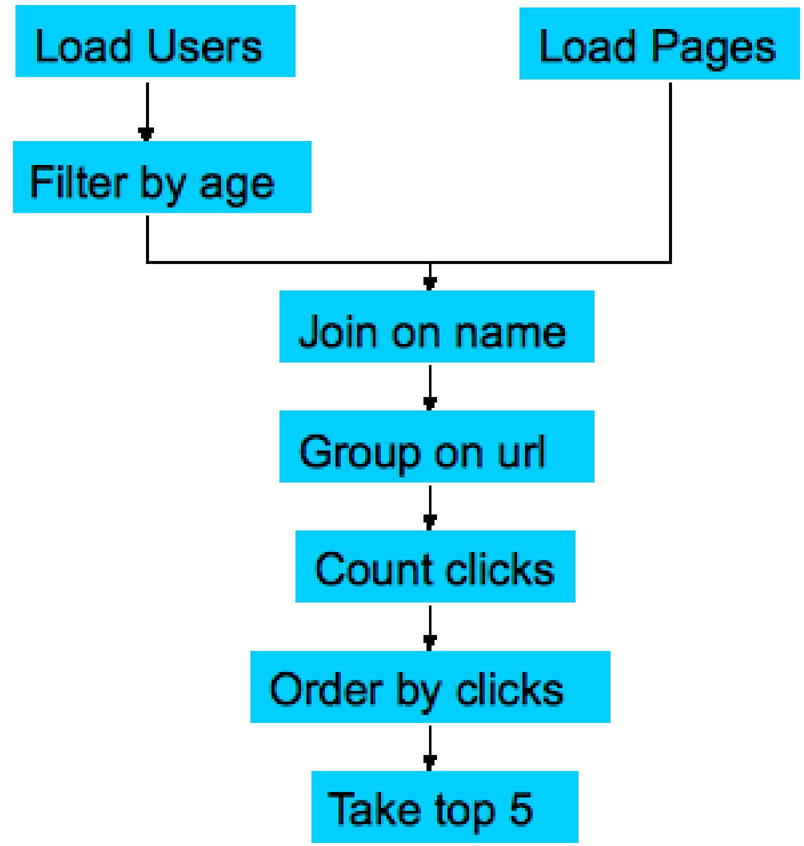
Higher level languages.

Cascading

HIVE


Y!

Suppose you have user data in one file, website data in another, and you need to find the top 5 most visited pages by users aged 18 - 25.



Example from PIG presentation at Apache Con EU 2009

Example from PIG presentation at Apache Con EU 2009

```
Users = load 'users' as (name, age);
Fltrd = filter Users by
        age >= 18 and age <= 25;
Pages = load 'pages' as (user, url);
Jnd = join Fltrd by name, Pages by user;
Grpd = group Jnd by url;
Smmd = foreach Grpd generate group,
       COUNT(Jnd) as clicks;
Srtd = order Smmd by clicks desc;
Top5 = limit Srtd 5;
store Top5 into 'top5sites';
```

Example from PIG presentation at Apache Con EU 2009

(Distributed) storage.

HYPERTABLE

H·BASE

Cassandra

A highly scalable, eventually consistent, distributed, structured key-value store.

apache CouchDB relax

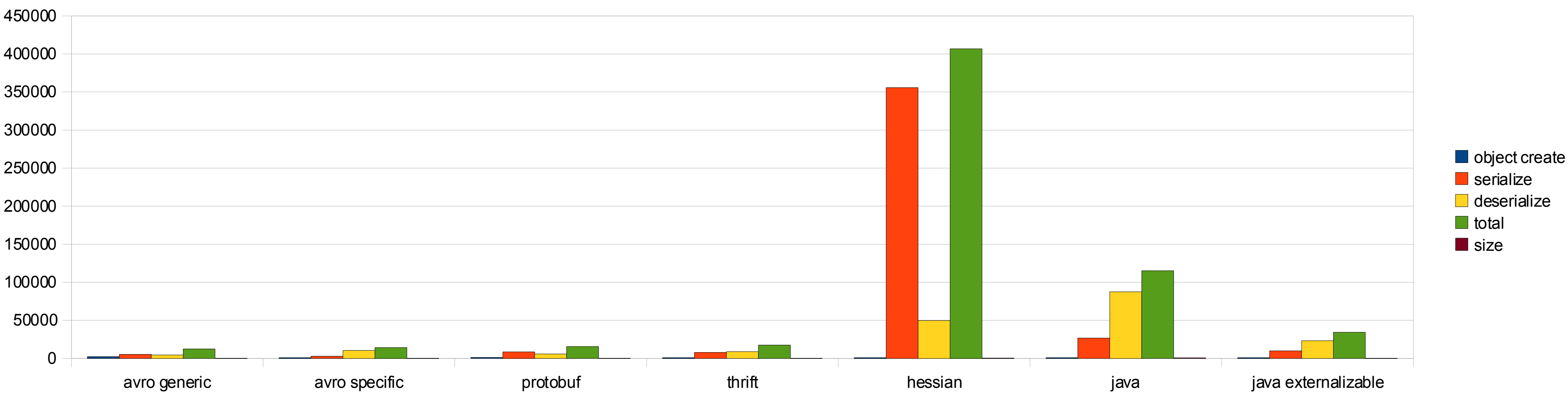
Libraries built on top.

Google code **protobuf**

Protocol Buffers - Google's data interchange format

AVRO

Thrift



MAHOUT

Apache Solr

AVRO

nutch

Thrift

Google code **protobuf**
Protocol Buffers - Google's data interchange format

katta

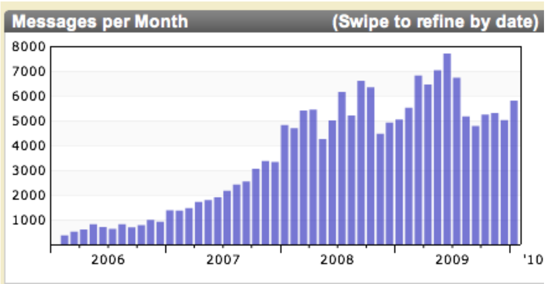

Jumpstart your project with proven code.
January 8, 2008 by dreizehn28
http://www.flickr.com/photos/1328/2176949559


Discuss ideas and problems online.
November 16, 2005 [phil h]
http://www.flickr.com/photos/hi-phi/64055296

http://www.flickr.com/photos/vonguard/4076391191
By: VonGuard

Become part of the community.

http://flickr.com/photos/iamdietrich/4075363845/
By: Iam Dietrich



http://www.flickr.com/photos/twleung/4110005817/
By: Ted Leung

*-user@hadoop.apache.org

*-dev@hadoop.apache.org

Image by: Patrick McEvoy

Interest in solving hard problems.

Being part of lively community.

Engineering best practices.

Bug reports, patches, features.

Documentation, code, examples.

Mar., 10th 2010: Hadoop* Get Together in Berlin

- Bob Schulze (eCircle/ Munich): Database and Table Design Tips with HBase
- Dragan Milosevic (zanox/ Berlin): Product Search and Reporting powered by Hadoop
- Chris Male (JTeam/ Amsterdam): Spatial Search

http://upcoming.yahoo.com/event/5280014/

* UIMA, Hbase, Lucene, Solr, katta, Mahout, CouchDB, pig, Hive, Cassandra, Cascading, JAQL, ... talks welcome as well.

Isabel Drost
Jan Lehnardt
newthinking store
Simon Willnauer


Berlin Buzzwords 2010
search - store - scale
N☆SQL

Berlin Buzzwords 2010

This is to announce the Berlin Buzzwords 2010 scalability conference. Berlin Buzzwords 2010 is scheduled for the start of June. Topics of interest include NoSQL databases, Hadoop, Lucene and others. Our goal is to bring developers and users together in central Europe for a conference featuring talks on scaling data analysis. The team organizing this event is deeply rooted in the Hadoop, Lucene, and CouchDB communities. Interested in helping? See the requests for helping hands. Also note that we are just getting off the ground. Please be patient as we get the various infrastructure pieces in place.

Berlin Buzzwords News

February 2009 - CfP to be published

The call for presentations will be published on this site in mid-February (including more detailed

June 7/8th: Berlin Buzzwords 2010

Store, Search, Scale

Solr
HBase
Lucene
Sphinx
Hadoop
Distributed computing
CouchDB
Business Intelligence
NoSQL
Cloud Computing
MongoDB
Scalability

*-user@hadoop.apache.org

*-dev@hadoop.apache.org

Image by: Patrick McEvoy

Interest in solving hard problems.

Being part of lively community.

Engineering best practices.

Bug reports, patches, features.

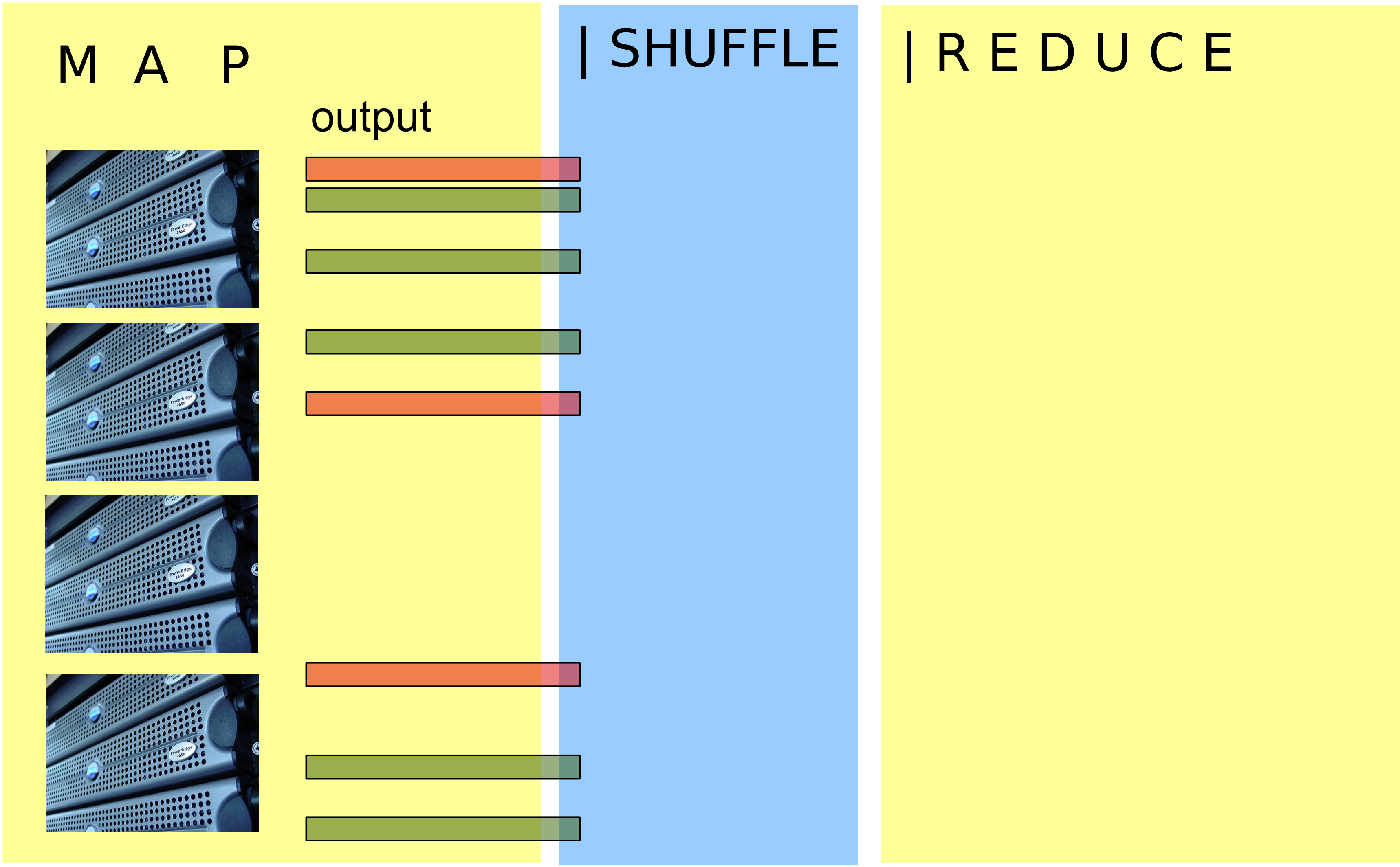
Documentation, code, examples.

M A P

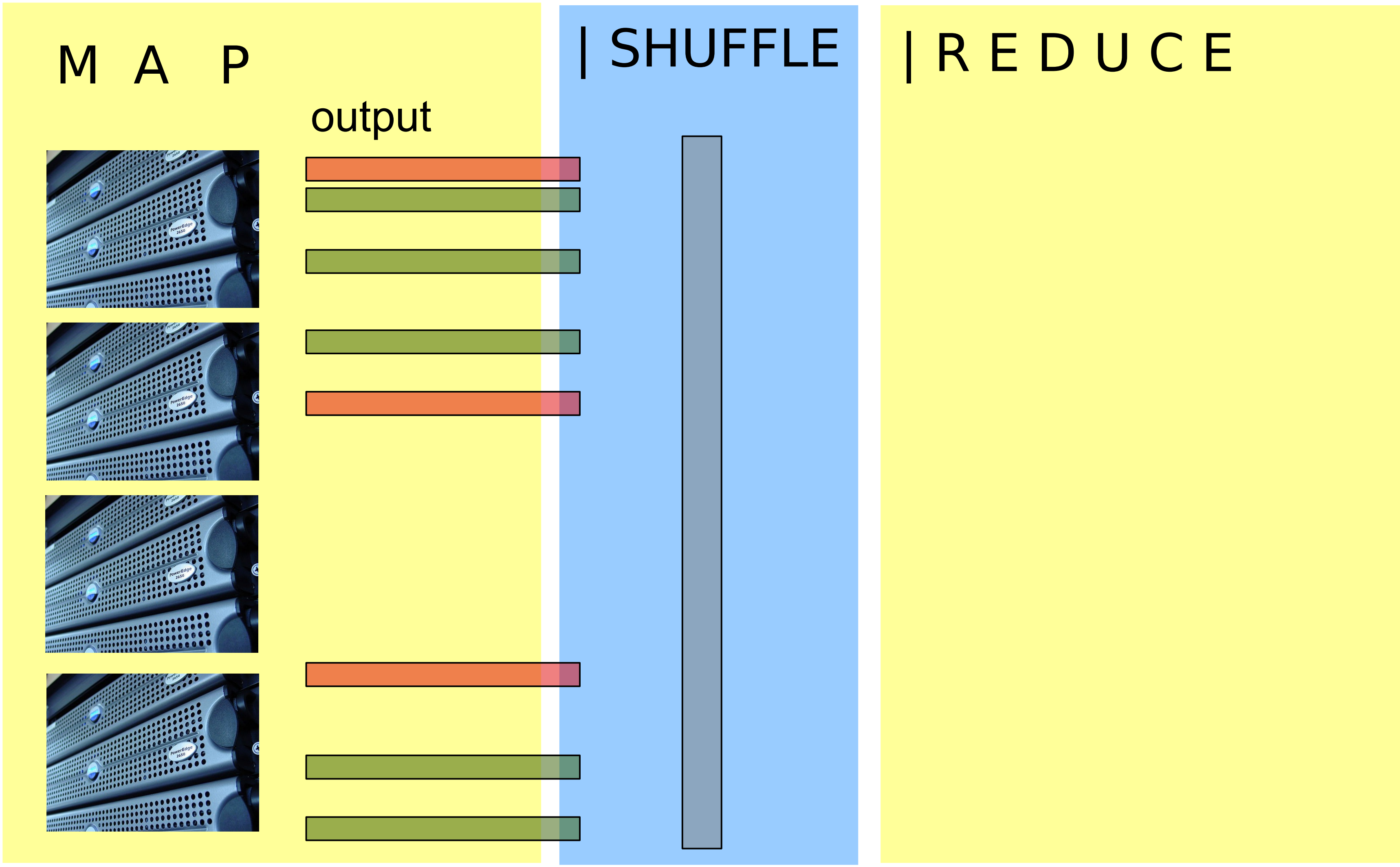
| SHUFFLE

| R E D U C E

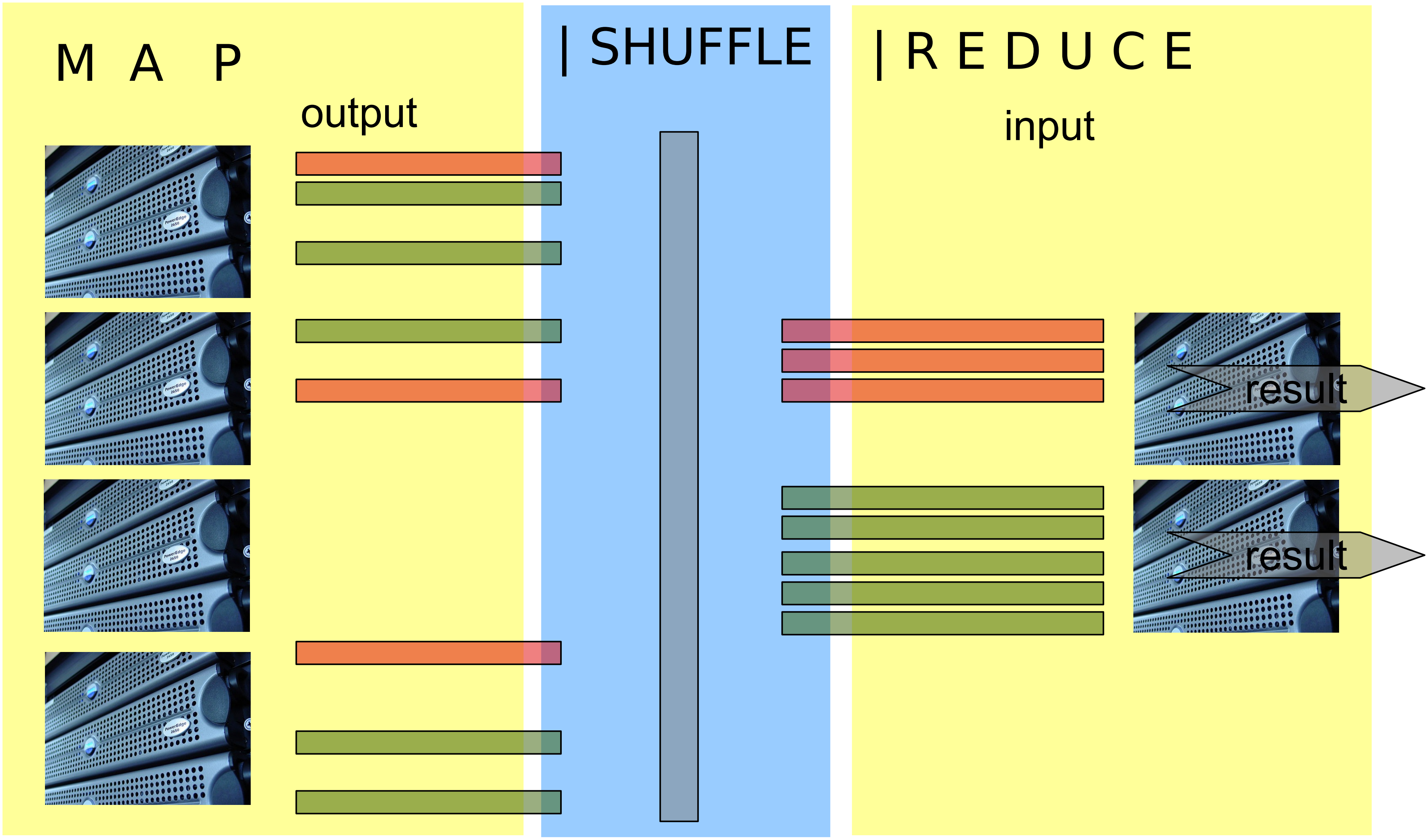
Local to data.



Local to data.
Outputs a lot less data.
Output can cheaply move.



Local to data.
Outputs a lot less data.
Output can cheaply move.



Local to data.
Outputs a lot less data.
Output can cheaply move.

Shuffle sorts input by key.
Reduces output significantly.