



# Cloud Computing at Yahoo!

Thomas Kwan  
Director, Research Operations  
Yahoo! Labs

# Overview

- Cloud Strategy
- Cloud Services
- Cloud Research Partnerships

# Yahoo! Cloud Strategy

1. Optimizing for Yahoo-scale
2. Fast product development & innovation
3. Robust data processing & serving environments
4. Reduced labor & costs for infrastructure
5. Internally focused, multi-year effort

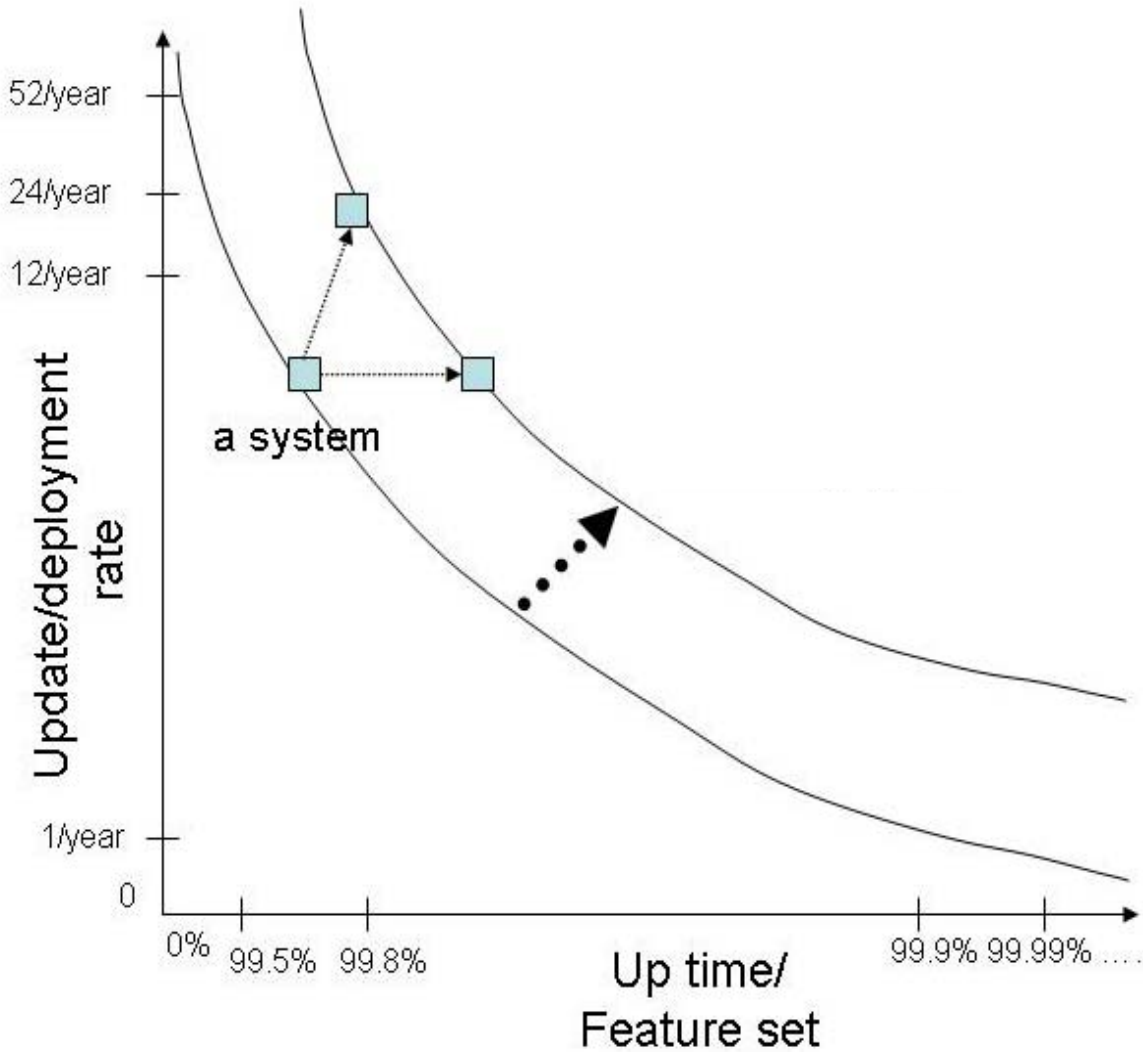


# 1. Optimizing for Yahoo-scale

- **Massive user base**
  - 500M+ unique users per month
  - Hundreds of peta bytes of storage
  - Hundreds of thousands of requests/second
- **Global scale**
  - Tens of globally distributed data centers
  - Serving each region at low latencies
- **Operational challenges**
  - Rapidly extracting value from data
  - Highly variable usage patterns



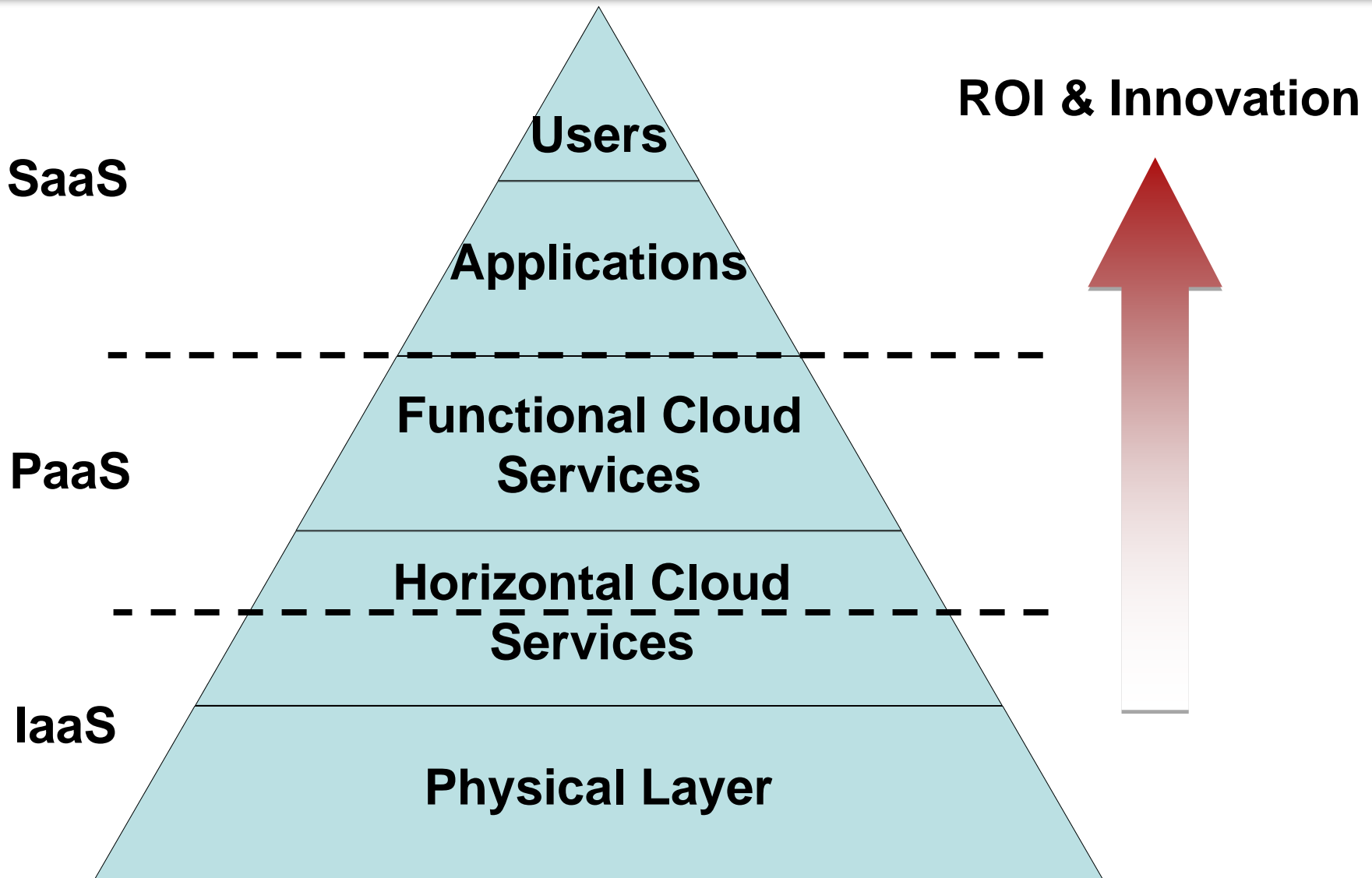
# 2. Fast Product Development...



Cloud increases agility and quality



# 2. ... Increase Innovation



# 3. Robust Processing & Serving

My Yahoo! | May 12, 2009

Sign In | New here? Sign Up

**YAHOO!** Web Images Video Local Shopping More

Web Search

**MY FAVORITES** + Add

- View Yahoo! Sites
- Yahoo! Mail
- Autos
- ebay
- Games
- Messenger
- Music
- MySpace
- omg!
- Personals
- Sports
- TV
- Weather

**RECOMMENDED**

- Deal Of The Day
- Buzz
- Shine

**Content Optimization**

**Search Index**

**TOP SEARCHES**

- Carrie Prejean
- Hubble Mission
- Detox Diets
- Milvina Dean
- Alzheimer's
- Leonard Nimoy
- Farrah Fawcett
- U.S. Stamps
- TV Recaps

**Machine Learning** (e.g. Spam filters)

**Attachment Storage**

**Ads Optimization**

**Image/Video Storage & Delivery**

**Miss California holds back tears**  
Carrie Prejean speaks out as she retains her title despite a recent photo scandal. Her emotional moment as she chokes up, Trump defends her

**NEWS WORLD LOCAL FINANCE**

- Pakistan dropping commandos into Taliban stronghold
- U.S. soldier charged with murder for shooting 5 comrades in...
- Details emerge on spying trial after Iran frees U.S....
- GM shares tumble to lowest level since Great Depression
- NTSB releases cockpit recordings from deadly Buffalo crash
- Bad boy doesn't always get the girl, tribal study shows
- Struggling Americans hosting parties to help pay rent
- LAPD takes in 1,700 weapons in gun buyback program - L.A. Times
- Lottery results - L.A. Times
- White House officials say no decision has been... - L.A. Times

updated 10:32 am PDT More: [News](#) [Popular](#) [Buzz](#)

**Dow:** 8,385.40 -0.39% **Nasdaq:** 1,700.55 -1.77%

Sponsored by: **Scottrade** Enter stock symbol **Get Quotes**

**700** What's Your Credit Score? A bad credit score is 600 or below. See yours in 2 easy steps for \$0. By Experian®

**Toyota Sales Event - Ad Feedback**

**SPOTLIGHT** « Prev Next »

**Most emailed news photos**

- A baby black jaguar is carried by its mother
- Lightning strikes lightning bolts races across space
- An osprey shows off its talents

**Image/Video Storage & Delivery**



# Technology Strategy: Design for...

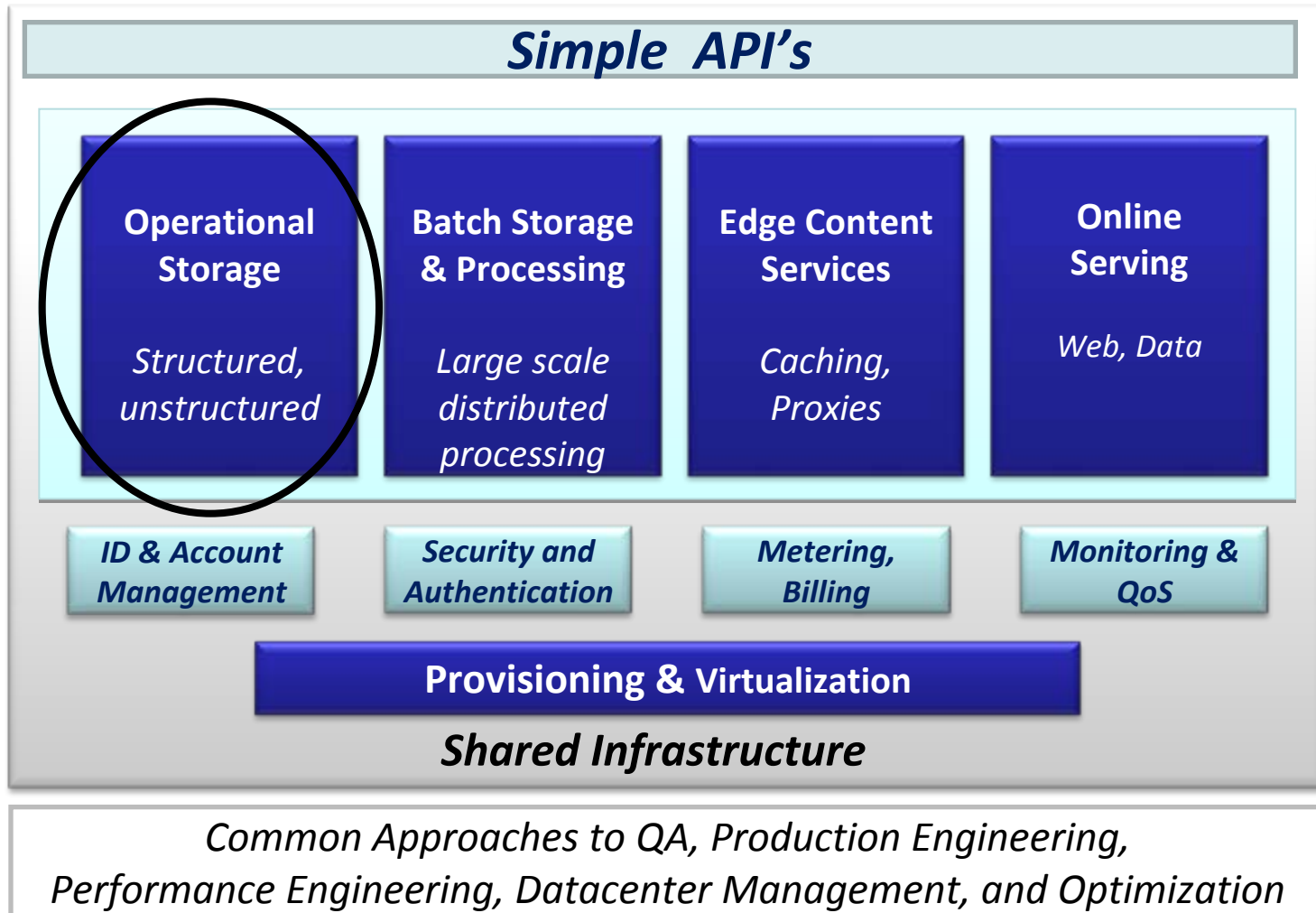
- **Open Source**
  - Hadoop, Pig, Traffic Server, etc.
- **High performance**
  - High throughput
  - Low latency
- **Multi-data center**
  - N-way replication
  - Consistency/availability tradeoffs
- **Flat, flexible infrastructure**



# Outline

- Cloud Strategy
- Cloud Services
- Cloud Research Partnerships

# Yahoo! Cloud Services



# Database in the Cloud

- **Reduce application development time**
- **Amortize operations/DBA costs**
- **Share best practices across applications**
- **Scale on-demand**



**Hosted, centrally-managed DB service**

# Structured Storage: PNUTS

- **Trade away “standard” DBMS features**
  - Complicated queries
  - Strong transactions
- **Must have**
  - Scalability
  - Flexibility
  - Availability

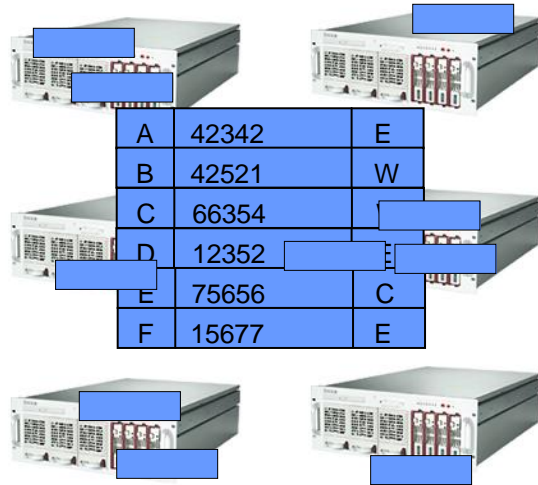


# PNUTS: Scalability and Flexibility

```
CREATE TABLE Parts (
  ID VARCHAR,
  StockNumber INT,
  Status VARCHAR
  ...
)
```

Structured, flexible schema

A	42342	E
B	42521	W
C	66354	W
D	12352	E
E	75656	C
F	15677	E



Parallel database  
(hashed or ordered tables)

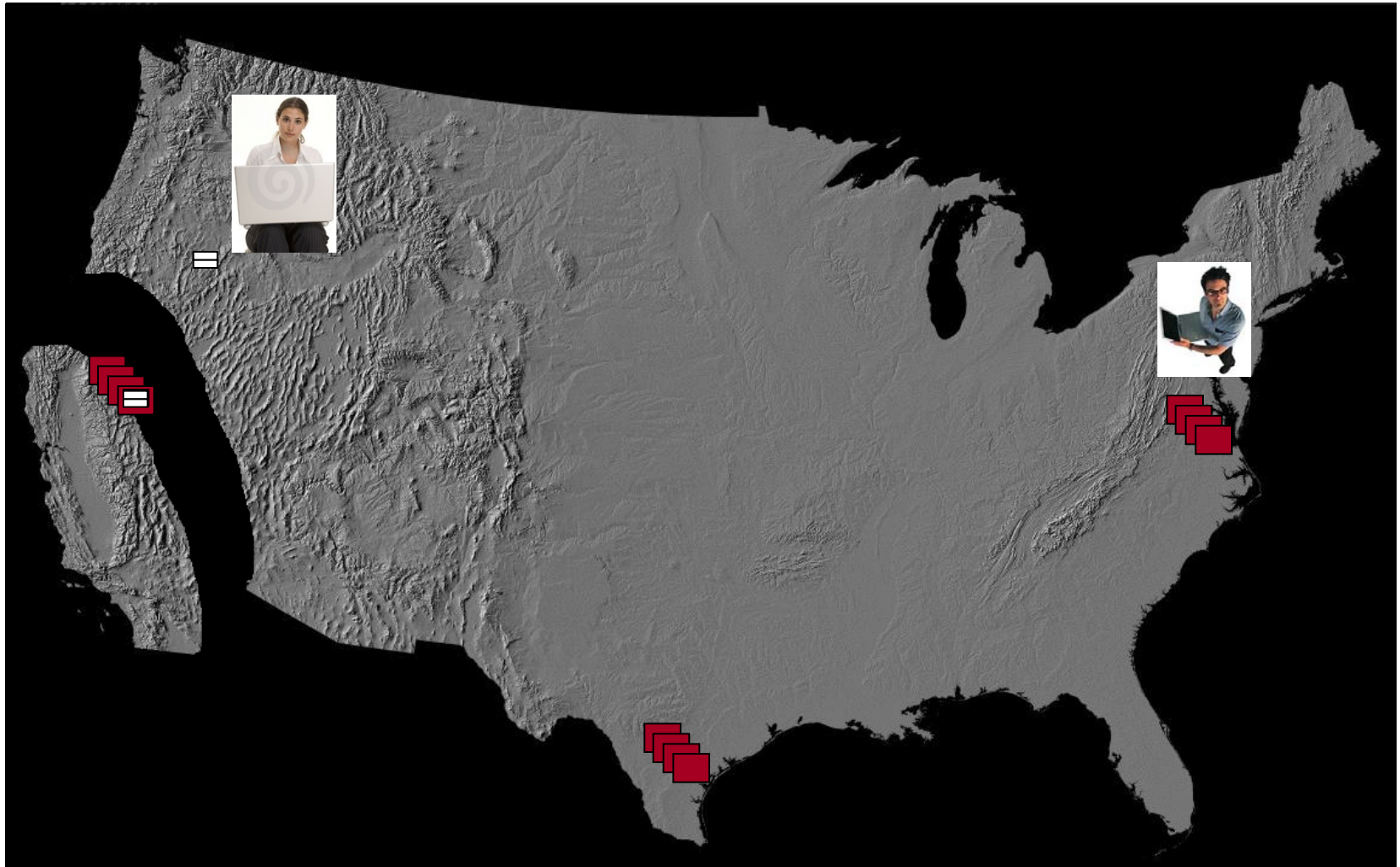


Hosted, managed infrastructure

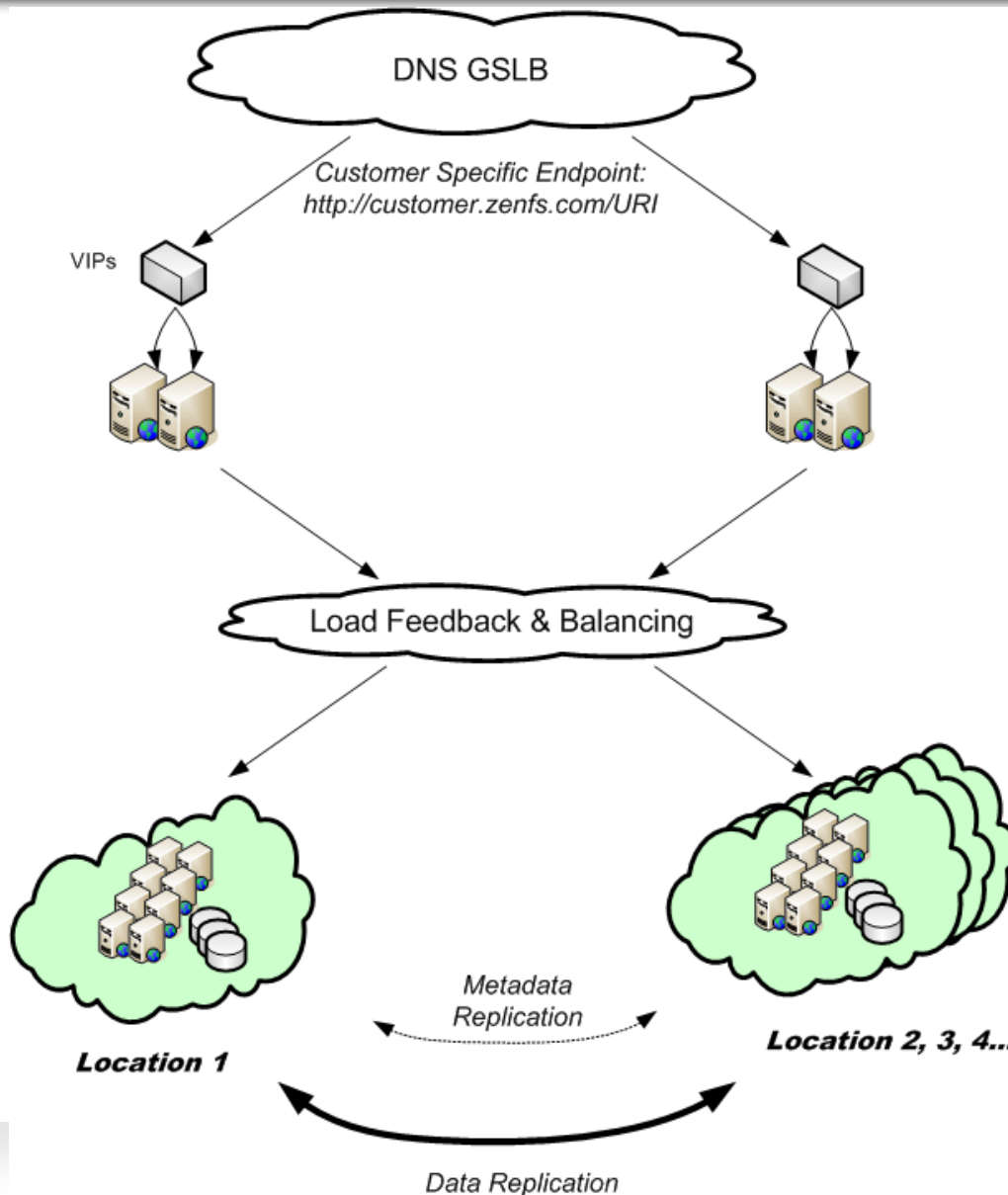
Geographic replication



# Availability & Low-Latency Access: Asynchronous Replication



# Unstructured Storage: MObsStor



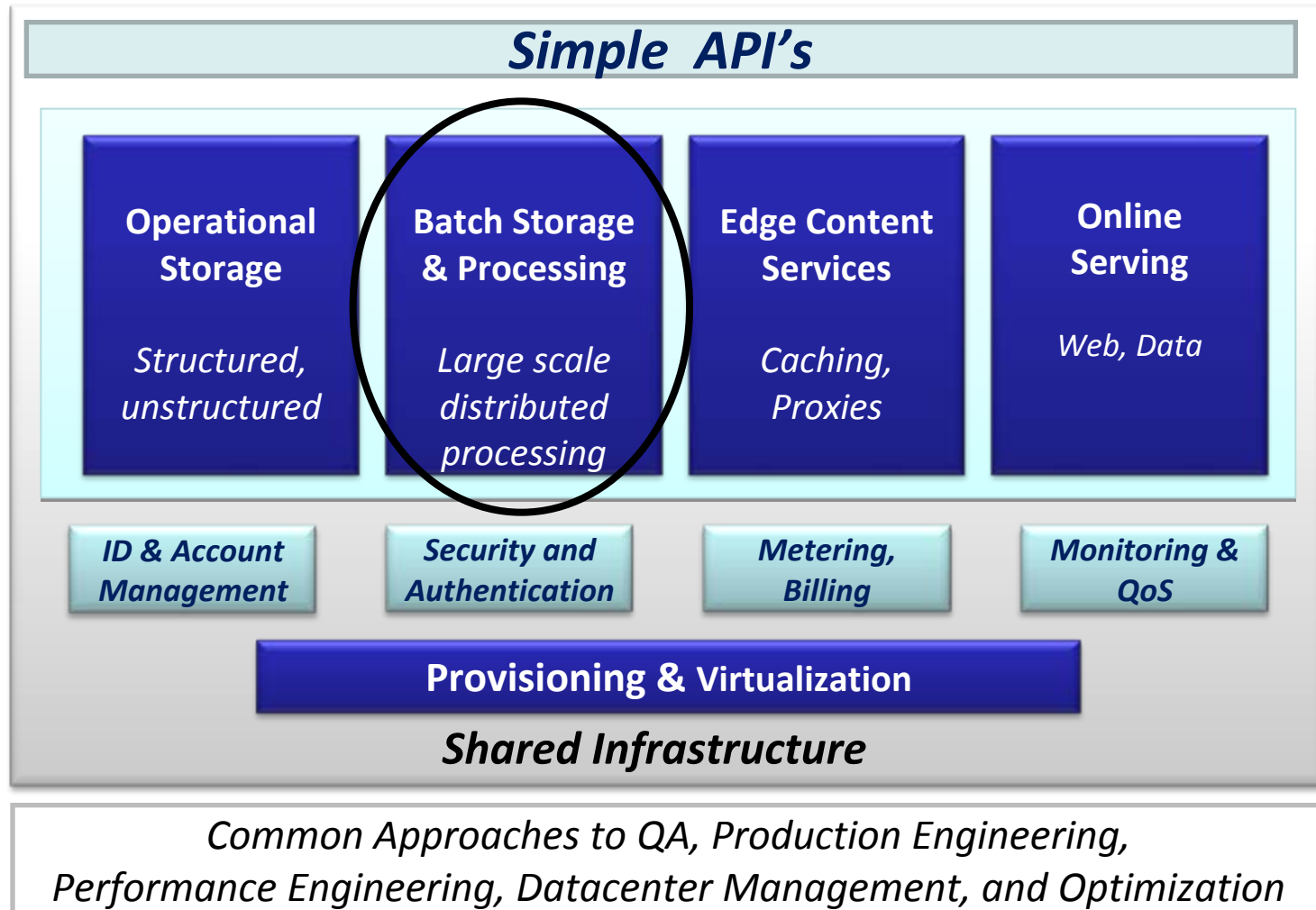
## Global Servers

- Caching
- Protocol termination
- Authentication
- Content routing

## Local Servers

- Auto expiration
- De-duplication
- Object placement
- Re-replication

# Yahoo! Cloud Services





# Batch Processing: Hadoop

- **Apache Hadoop**
  - Open source project started in 2005
- **Milestones at Yahoo!**
  - Became primary contributor in 2006
  - Scaled from 20 to 4,000 node clusters
  - Began running production jobs in Q1 2008
- **Characteristics**
  - Portable (written in Java)
  - Uses commodity hardware



# Hadoop Distributed File System

- **Single petabyte file system**
  - Managed by a *namenode*
  - Support files read/write but append-only
  - Optimized for streaming reads of large files
- **Files are broken into large blocks**
  - Transparent to the client
  - Data is checksummed
  - Blocks replicated to several *datanodes*
- **Client talks to *namenode* & *datanodes***



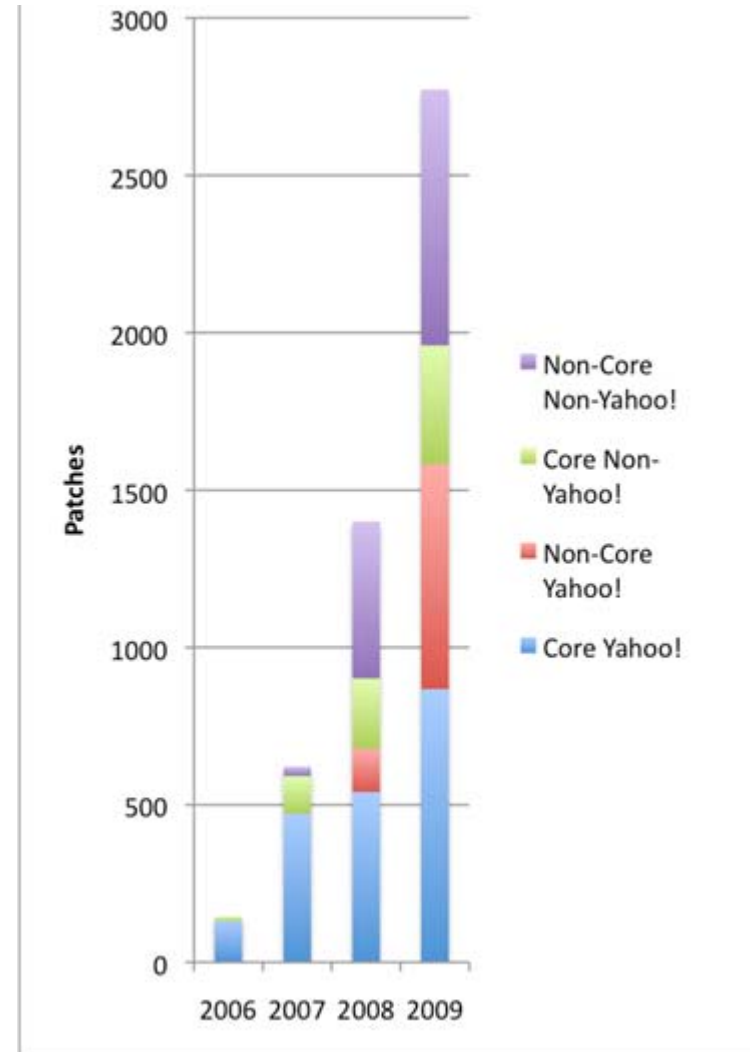
# Map/Reduce Support

- Programming model for distributed computing
- Similar to Unix pipeline:
  - `cat input | grep | sort | uniq -c | cat > output`
  - **Input | Map | Shuffle & Sort | Reduce | Output**
- Efficiency from streaming through data, reducing seeks, and pipelining
- Good fit for a lot of applications
  - Log processing
  - Web index building
  - Data mining and machine learning



# Yahoo! Contributions to Hadoop

- Each contribution is a patch
- Core/non-core sub-projects
- Core: HDFS and Map/Reduce
- Core contributors
  - 185 people (30% Yahoos)
  - 72% of patches from Yahoo!



# Hadoop Greatly Improves Productivity

- Easy to learn
- Key computation solved in days (not months)
- Production and research use same framework
- No need to find new hardware for experiments
- Projects move from research to production in days



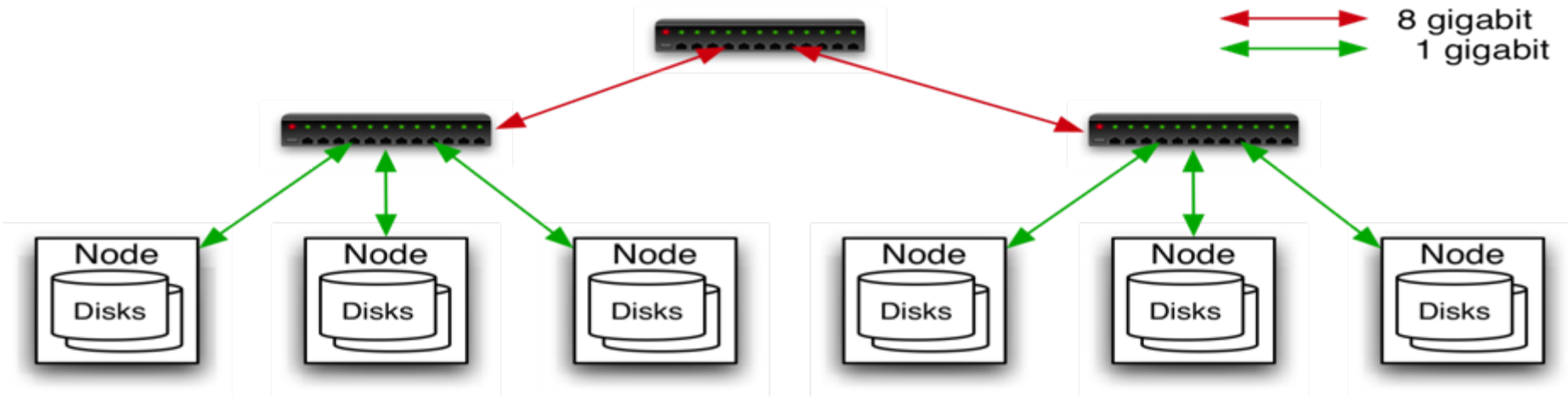
# Hadoop Application: Search Assist™

- Build Search Assist DB using 3 years of log-data



	Before Hadoop	After Hadoop
Time	26 days	20 minutes
Language	C++	Python
Development Time	2-3 weeks	2-3 days

# Hadoop Clusters at Yahoo!



- Commodity hardware
  - Linux PCs with 4 local disks
- Two-level architecture
  - 40 nodes/rack, 1Gb/s intra-rack, 8 Gb/s uplink
- Total 25,000+ nodes (~200,000 cores)

# Hadoop Statistics at Yahoo!

	2008	2009
Webmap	<ul style="list-style-type: none"><li>~70 hours runtime</li><li>~300 TB shuffling</li><li>~200 TB output</li><li>1480 nodes</li></ul>	<ul style="list-style-type: none"><li>~73 hours runtime</li><li>~490 TB shuffling</li><li>~280 TB output</li><li>2500 nodes</li></ul>
Sort benchmarks (Jim Gray contest)	<ul style="list-style-type: none"><li>1 Terabyte sorted</li><li>• 209 seconds</li><li>• 900 nodes</li></ul>	<ul style="list-style-type: none"><li>1 Terabyte sorted</li><li>• 62 seconds; 1500 nodes</li><li>1 Petabyte sorted</li><li>• 16.25 hours; 3700 nodes</li></ul>
Largest cluster	<ul style="list-style-type: none"><li>2000 nodes</li><li>• 6PB raw disk</li><li>• 16TB of RAM</li><li>• 16K CPUs</li></ul>	<ul style="list-style-type: none"><li>4000 nodes</li><li>• 16PB raw disk</li><li>• 64TB of RAM</li><li>• 32K CPUs</li><li>• (40% faster CPUs too)</li></ul>

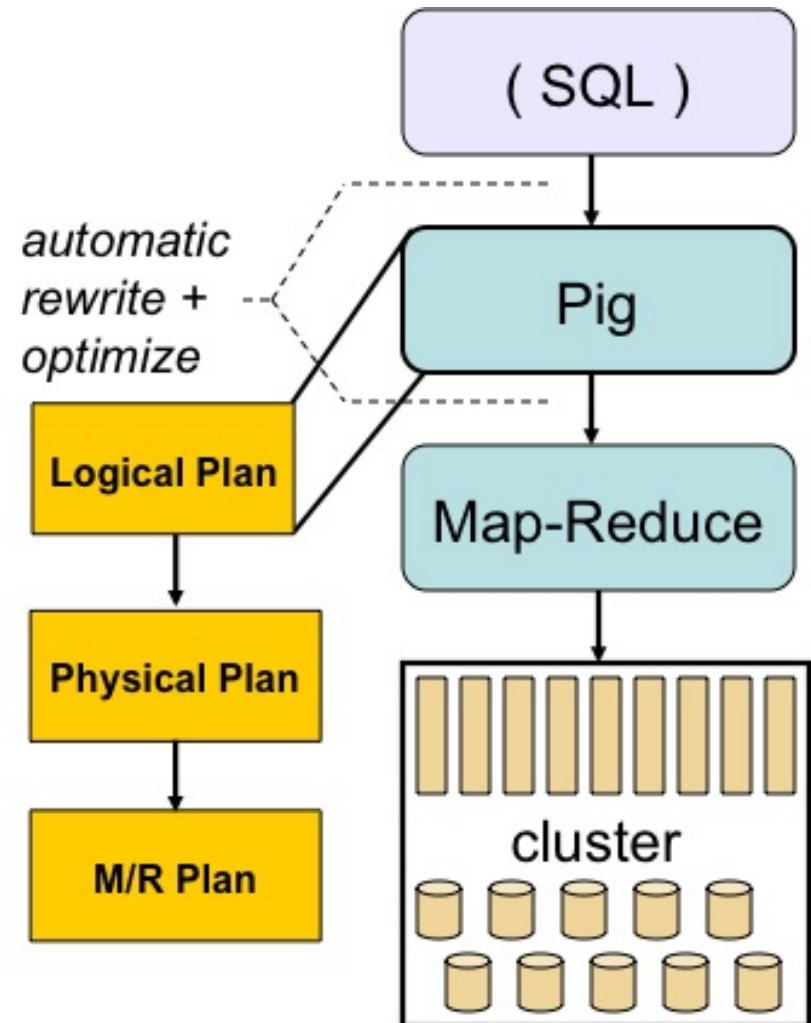




# Processing Large Data Sets: Pig

Pig:

- High-level procedural language
- Parser, optimizer and distributed query execution
- Insulate against Hadoop complexity (e.g., version upgrades)



# Pig Increases Programmer Productivity

In Native Hadoop:

In Pig:

```

import java.io.IOException;
import java.util.ArrayList;
import java.util.List;
import java.util.HashMap;
import java.util.HashSet;
import java.util.Iterator;
import java.util.Set;
import java.util.Arrays;
import java.util.Date;
import java.util.Calendar;
import java.util.GregorianCalendar;
import java.util.Locale;
import java.util.TimeZone;
import java.util.Date;
import java.util.Calendar;
import java.util.GregorianCalendar;
import java.util.Locale;
import java.util.TimeZone;

import org.apache.hadoop.conf.Configuration;
import org.apache.hadoop.fs.Path;
import org.apache.hadoop.io.IntWritable;
import org.apache.hadoop.io.LongWritable;
import org.apache.hadoop.io.Text;
import org.apache.hadoop.mapred.Mapper;
import org.apache.hadoop.mapred.Partitioner;
import org.apache.hadoop.mapred.Reducer;
import org.apache.hadoop.mapred.MapperCombiner;
import org.apache.hadoop.mapred.ReducerCombiner;
import org.apache.hadoop.mapred.MapperCombinerReducer;
import org.apache.hadoop.mapred.ReducerCombinerMapper;
import org.apache.hadoop.mapred.MapperCombinerReducerCombiner;
import org.apache.hadoop.mapred.ReducerCombinerMapperCombiner;
import org.apache.hadoop.mapred.MapperCombinerReducerCombinerMapper;
import org.apache.hadoop.mapred.ReducerCombinerMapperCombinerCombiner;

public class WordCountMapper extends Mapper<Text, Text, Text, Text> {
    public void setup(Conf conf) throws IOException {
        // ...
    }

    public void map(LongWritable key, Text value, Reporter reporter) throws IOException {
        // ...
    }
}

public class WordCountReducer extends Reducer<Text, Text, Text, Text> {
    public void setup(Conf conf) throws IOException {
        // ...
    }

    public void reduce(Text key, Iterator<Text> values, Reporter reporter) throws IOException {
        // ...
    }
}

public class WordCountMapperCombiner extends MapperCombiner {
    // ...
}

public class WordCountReducerCombiner extends ReducerCombiner {
    // ...
}

public class WordCountMapperCombinerReducer extends MapperCombinerReducer {
    // ...
}

public class WordCountReducerCombinerMapper extends ReducerCombinerMapper {
    // ...
}

public class WordCountMapperCombinerReducerCombiner extends MapperCombinerReducerCombiner {
    // ...
}

public class WordCountReducerCombinerMapperCombiner extends ReducerCombinerMapperCombiner {
    // ...
}

public class WordCountMapperCombinerReducerCombinerMapper extends MapperCombinerReducerCombinerMapper {
    // ...
}

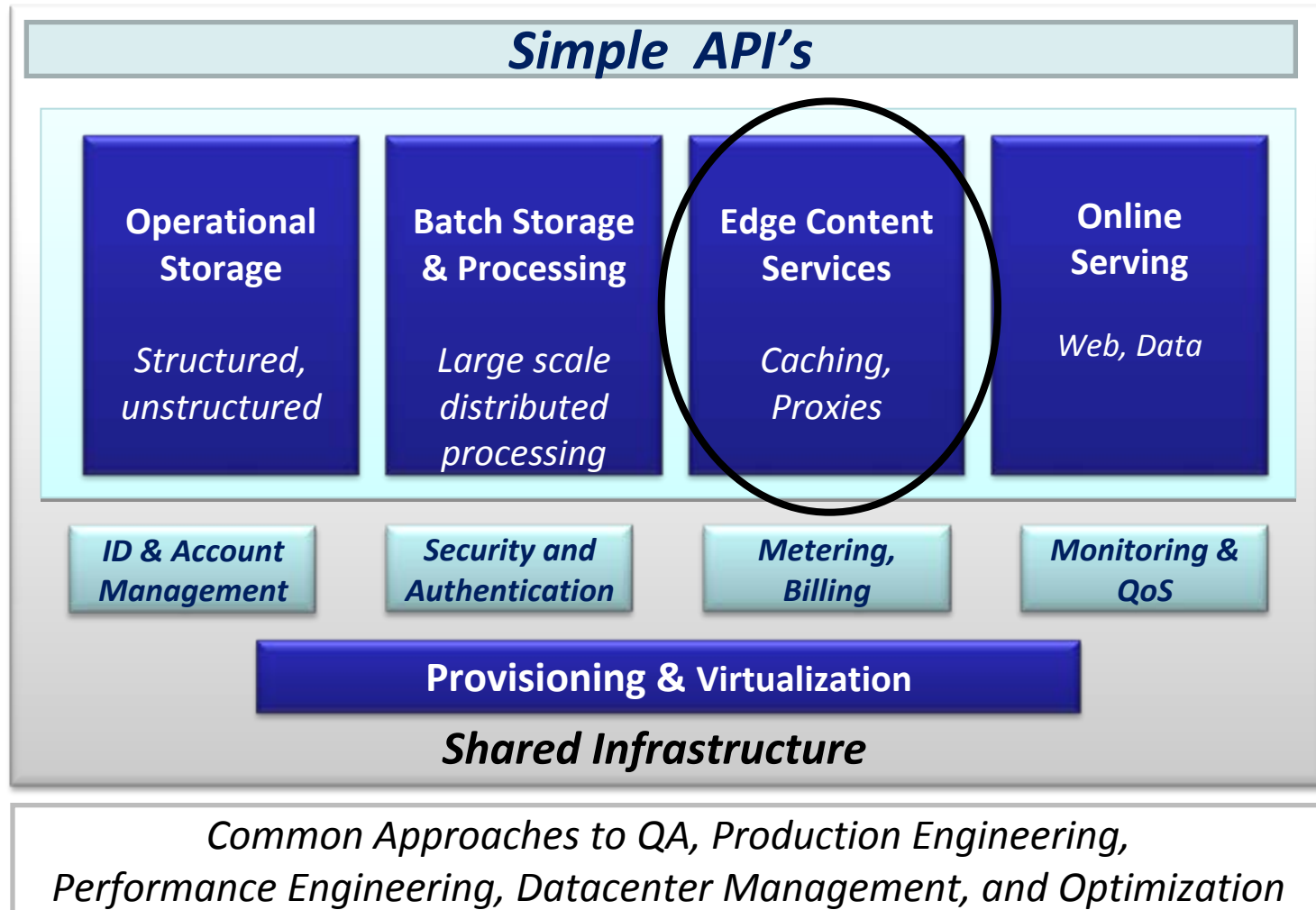
public class WordCountReducerCombinerMapperCombinerCombiner extends ReducerCombinerMapperCombinerCombiner {
    // ...
}
    
```

Users = load 'users' as (name, age);  
 Filtered = filter Users by age >= 18 and age <= 25;  
 Pages = load 'pages' as (user, url);  
 Joined = join Filtered by name, Pages by user;  
 Grouped = group Joined by url;  
 Summed = foreach Grouped generate group, COUNT(Joined) as clicks;  
 Sorted = order Summed by clicks desc;  
 Top5 = limit Sorted 5;  
 store Top5 into 'top5sites';

30%+ of Hadoop jobs at Yahoo! are in Pig



# Yahoo! Cloud Services



# Yahoo! Traffic Server

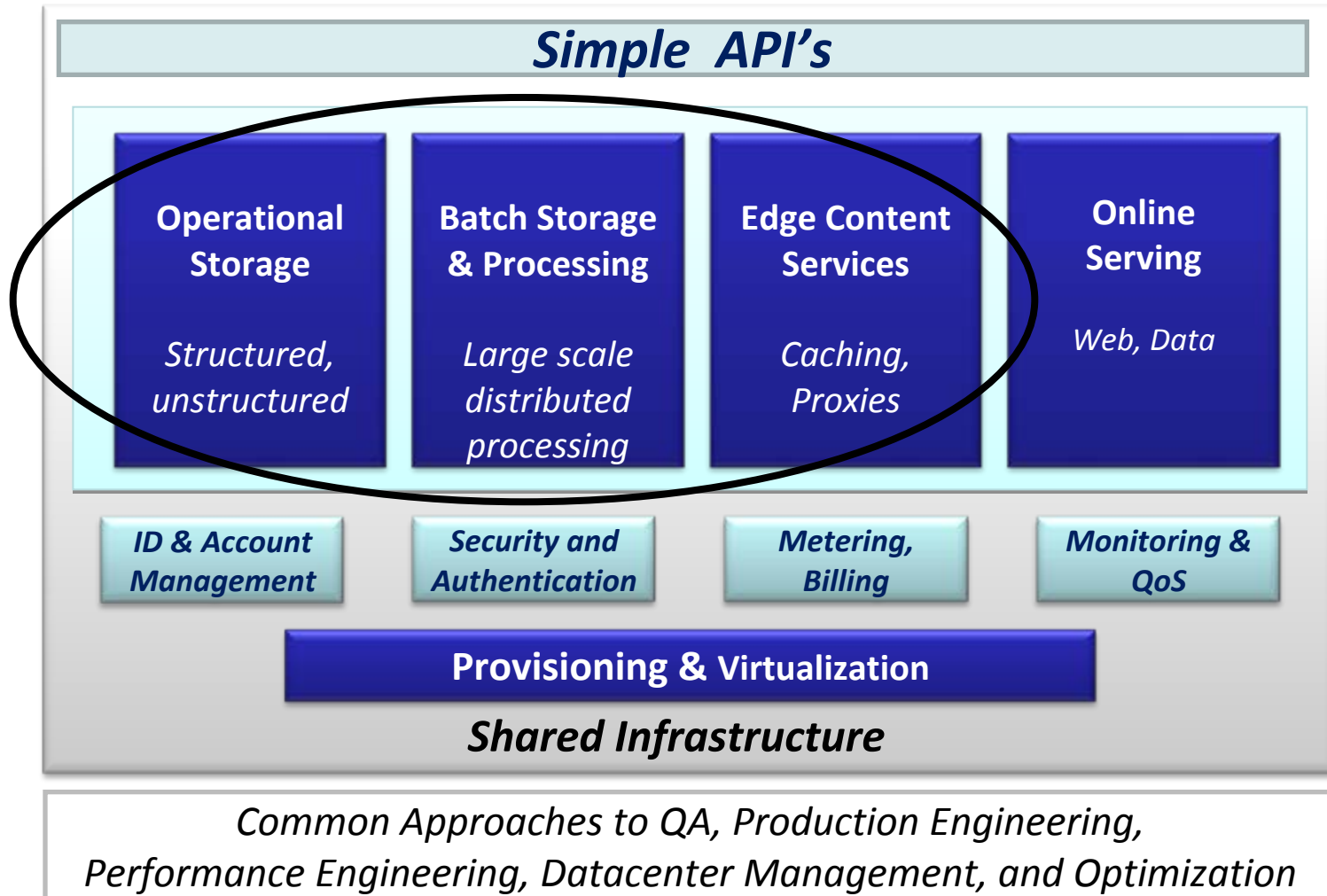
- **Open Source**
  - Released to Apache in November 2009
- **Caching proxy**
  - Can process 35,000 requests/second per server
  - Efficient use of multi-core CPUs
  - Extensible framework
- **Yahoo! usage**
  - Serve 400+ TB of data per day
  - Serve 30 billion Web objects per day

# Open Source Website

- For more information:
  - <http://apache.org>
  - <http://hadoop.apache.org/>
  - <http://developer.yahoo.com/hadoop/>



# Yahoo! Cloud Services



# Outline

- Cloud Strategy
- Cloud Services
- Cloud Research Partnerships

# Cloud Partnership with CMU: Nov. 2007

SCIENTIFIC  
AMERICAN

## Yahoo, Carnegie Mellon Switch On Supercomputer

By [WebProNews Staff](#) - Mon, 11/12/2007 - 12:08



The M45 supercomputer provided by Yahoo opened its ports to its partners at Carnegie Mellon University, where the initiative should help boost research that benefits the broader Internet community.

November 13, 2007 | [10 comments](#)

## Yahoo! Unleashes Teraflops of Processing Power for Research

ellon

### News

#### Yahoo! Reaches for the Stars with M45 Supercomputing Project

Faculty and students at the nation's best universities are hungry for an Internet-scale computing environment, but it's almost impossible to find this kind of computing power on a university campus.

Not any more. Yahoo! is bringing large-scale supercomputing to the academic research community through its newly launched M45 project. Named after a well-known open star cluster, M45 is a 4,000-processor supercomputer that's one of the fifty most powerful systems in the world. The goal of the project: help academic researchers tackle some of the most complicated computing tasks known to humanity.





# Open Cirrus Launched in July 2008

## The New York Times

July 29, 2008, 3:19 PM

### The Virtuous Competition in Cloud Computing Research

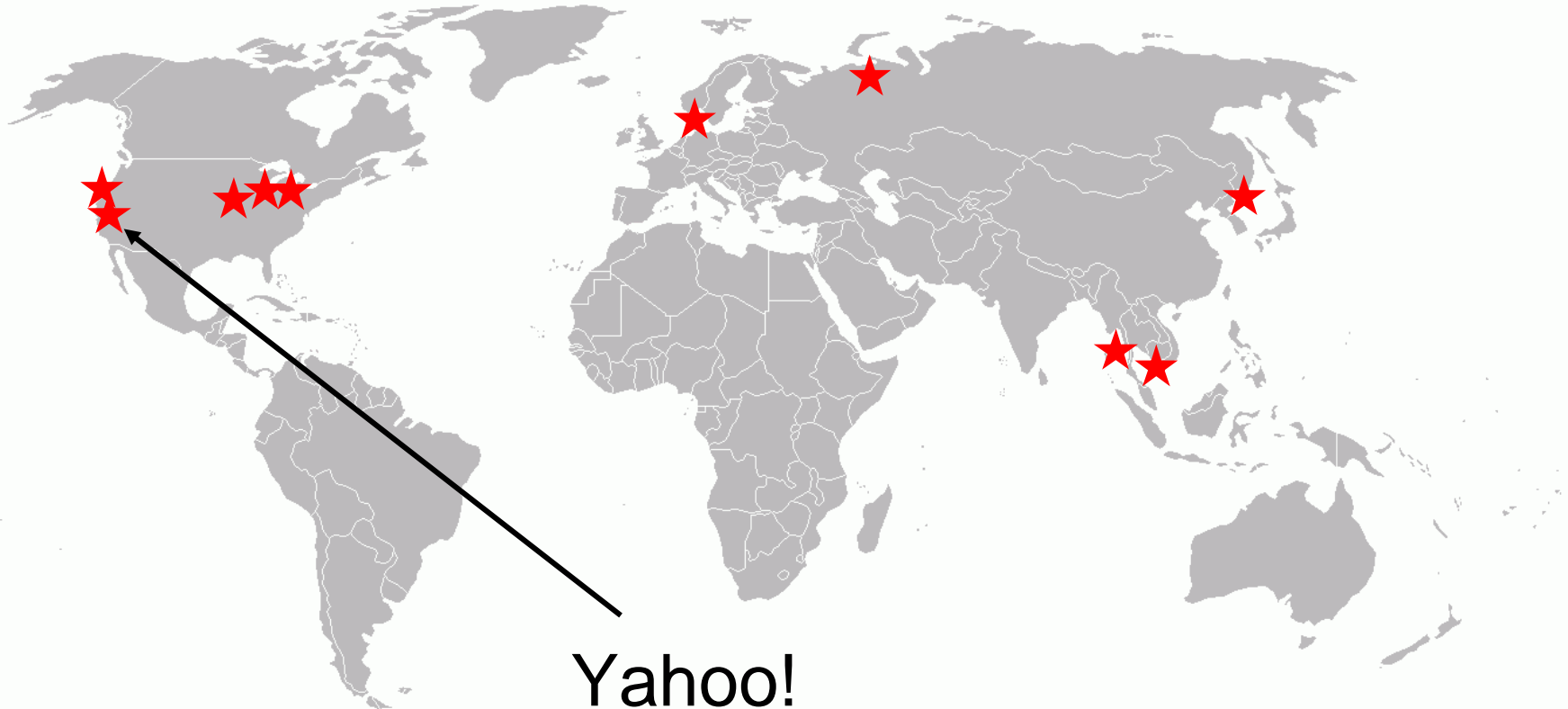
By STEVE LOHR

One more sign that we've entered the cloud computing era: the big corporate players are competing with one another to rev up academic research initiatives (partly with an eye toward wooing future computer scientists to work for them, of course).

Yahoo, Hewlett-Packard and Intel announced a research venture on Tuesday that spans the United States, Germany and Singapore. The goal is to advance Internet-scale computing — the proverbial “cloud,” in which more computing

# Open Cirrus: 10 Sites Worldwide

- U.S.: HP, Intel, Yahoo!, Illinois, Carnegie Mellon
- International: RAS, KIT, IDA, MIMOS, ETRI



# Yahoo!'s M45 Cluster

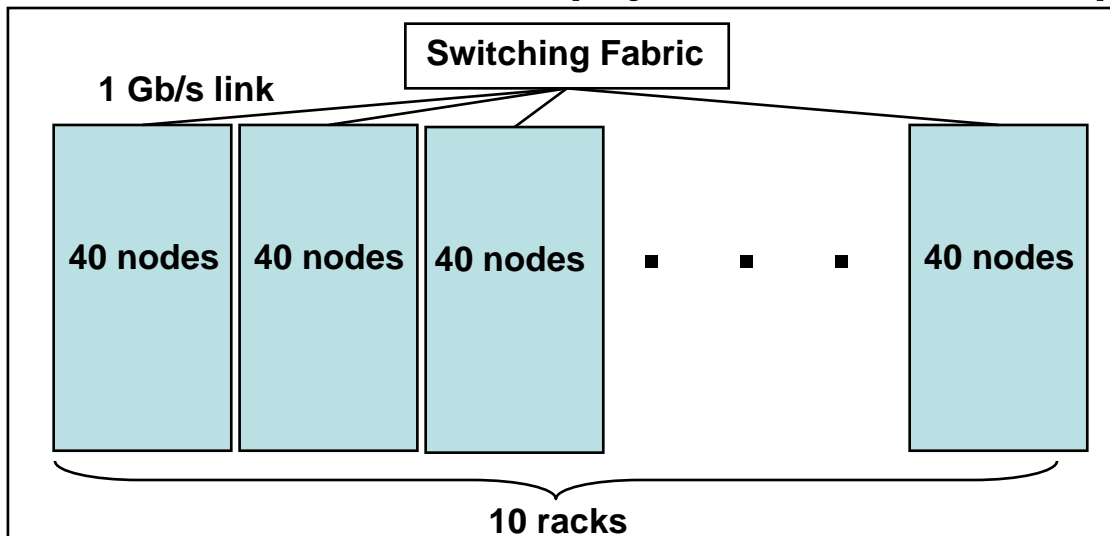
500 nodes



Each node



400 node Hadoop job execution partition



3,200 cores  
2.4 TB RAM  
1.2 PB disk

# M45 Research Partnerships

Carnegie Mellon



Berkeley  
University of California



Cornell University



UMASS  
AMHERST



# M45 Projects (Carnegie Mellon)

PI	Area	Research Project
Garth Gibson	Systems	Storage for DISC
Priya Narasimhan	Systems	Diagnosis/visualization
Tom Mitchell	ML	Read the Web
Noah Smith	NLP	Grammar induction
Jamie Callan	NLP	Educational learning
Stephan Vogel	NLP	Machine translation
Christos Faloutsos	Algorithms	Graph mining
Robert Kraut	HCI	Wikipedia user analysis



# Selected M45 Research Results

- Garth Gibson (CMU)
  - Framework for converting triplicated HDFS blocks into RAID 5/6 encodings in the background. Lowered capacity overhead by 33%
  - M45 logs key to analysis
- Priya Narasimhan (CMU)
  - Developed log-analysis techniques and visualization tools to troubleshoot a variety of performance problems in Hadoop clusters
  - M45 logs key to analysis
- Tom Mitchell (CMU)
  - Knowledge extraction from hundreds of millions of Web pages
  - Information extraction at quality and scale not reported before
- Noah Smith (CMU)
  - NLP models that require large datasets or computational resources
  - Inference algorithms tested in hours or days instead of months



# Selected M45 Research Results

- Christos Faloutsos (CMU)
  - Developed peta-scale graph mining system
  - Enabled analysis of very large graphs with more than billions of nodes and edges
- Robert Kraut and Niki Kittur (CMU)
  - Analyzed Wikipedia “wisdom of crowds”
  - Found identifying with group can shift contributor to align with group goal
- Stephan Vogel (CMU)
  - Built open-source Hadoop-based machine translation system
  - Improved machine translation quality enormously



# M45 Projects (Berkeley, Cornell, UMass)

PI	Area	Research Project
Joe Hellerstein	Systems	Hadoop online
Michael Jordan	ML	Statistical ML
Dan Starr	App	Astrophysics/Hadoop
Kimmen Sjolander	App	Biomedical computing
Bart Selman	AI	Computational Decision Making
David Smith	IR	Mining Million Books





# M45 Research Results

Key contributions:

- Over 40 technical publications to date
- Enabled training of new generation of scientists
- Built academic community around Hadoop



# Summary

- Cloud Strategy
- Cloud Services
- Cloud Research Partnerships

# Thank You

