Application Development for the Cloud:
A Paradigm Shift

Ramesh Rangachar

Intelsat

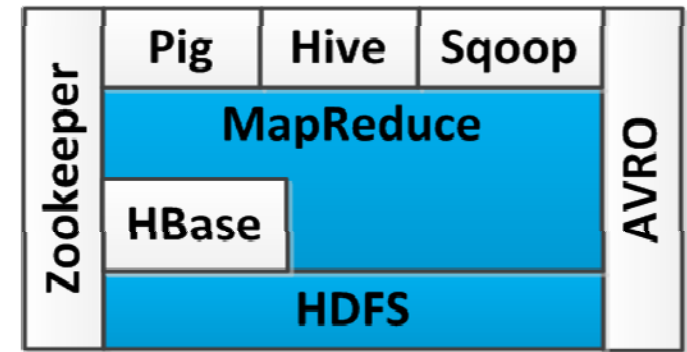# Motivation for the Paradigm Shift

- **The performance of the traditional applications is becoming inadequate**
  - Data is growing faster than Moore's Law [IDC study]
  - The new buzzword is Big Data
  - Application performance and scalability are limited by the performance and scalability of storage (disk I/O) and network
- **Applications must capitalize on the benefits provided by the cloud environment**
  - No significant improvements in "scaling up" (scale vertically or upgrade a single node)
  - "Scaling out" (scale horizontally or add more nodes) is a better approach
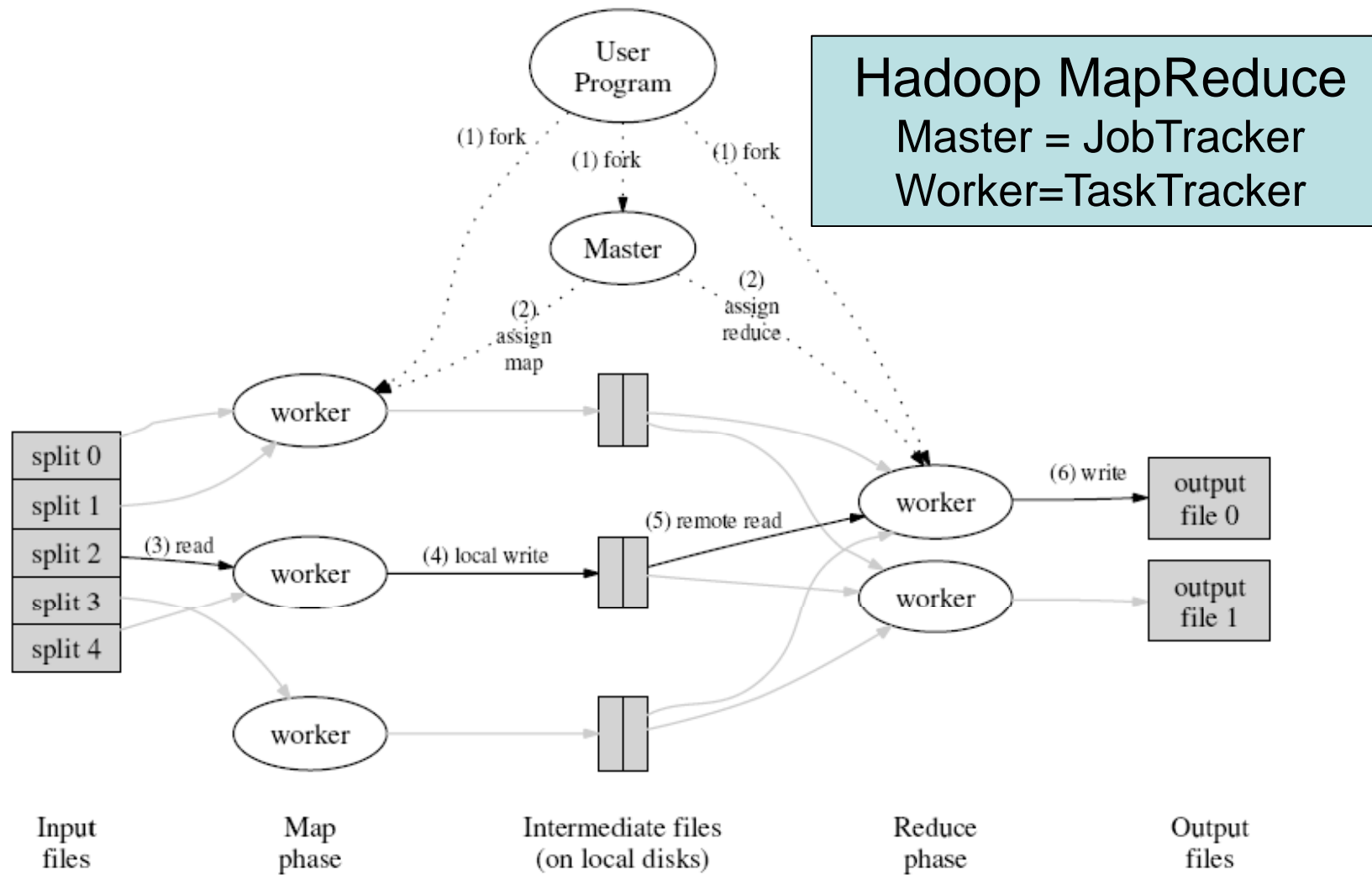
How are Others Solving Big Data Problem?

- **Google invented a software framework called MapReduce to support distributed computing on large data sets on clusters of computers**
  - **Google also invented Google Bigtable, a distributed storage system for managing petabytes of data**
- **Hadoop: An open source Apache project inspired by the work at Google. A software framework for reliable, scalable, distributed computing**
  - **There are a number of other Hadoop related projects at Apache**
  - **Major contributors include Yahoo, Facebook, and Twitter**
- **S4 Distributed Stream Computing Platform: An Apache incubator program. Supports development of applications for processing continuous unbounded streams of data**
- **Many more …**
- **We'll review some technologies in the Hadoop ecosystem**

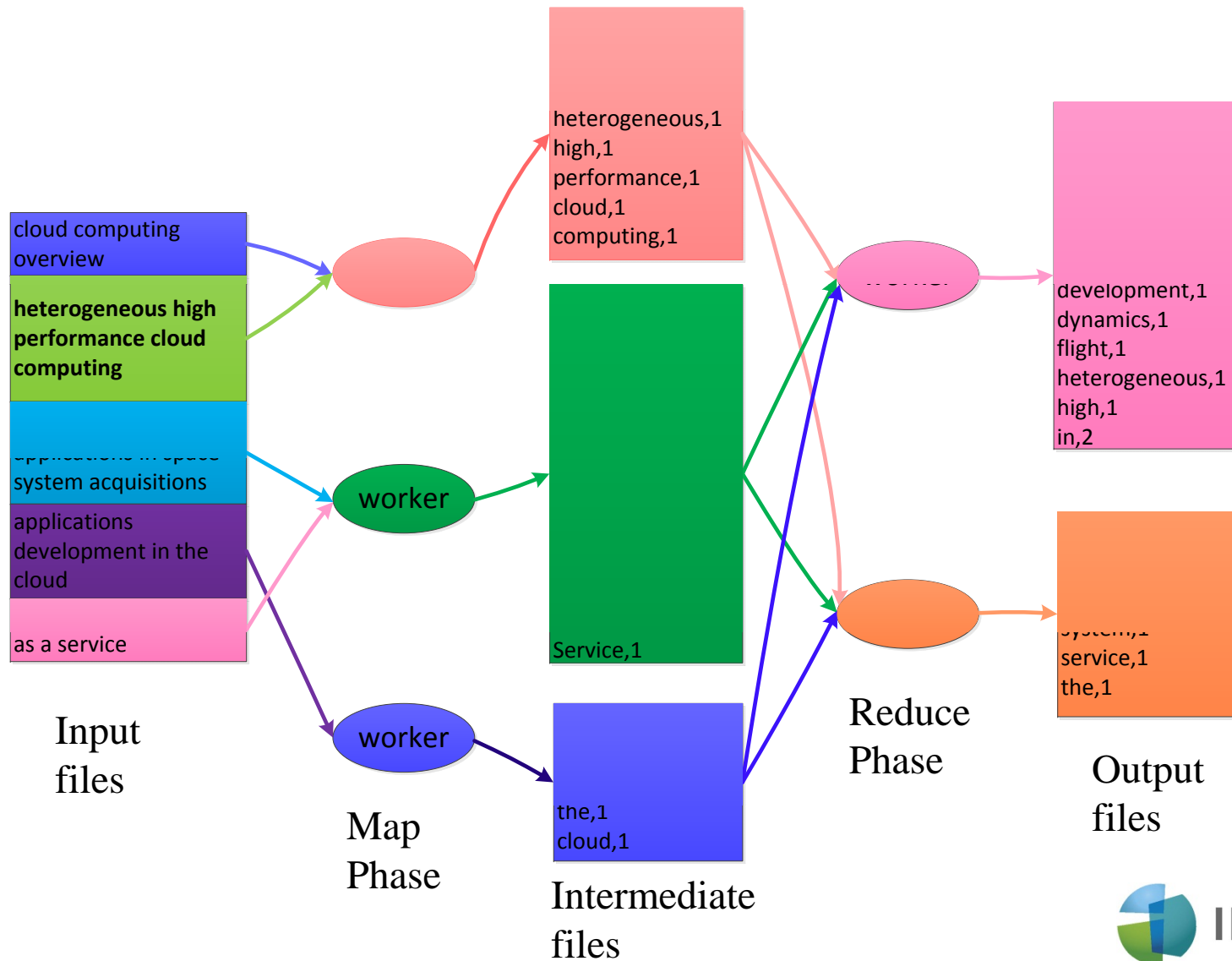INTELSAT.

What is Apache Hadoop?

- **The Apache™ Hadoop™ project develops open-source software for reliable, scalable, distributed computing**

  - **A framework for the distributed processing of large data sets across clusters of computers using a simple programming model**

  - **Designed to scale from single server to thousands of servers. Each server offers local computation and storage**

- **The Hadoop core consists of:**

  - **Hadoop Distributed File System (HDFS™): A distributed file system that provides high-throughput access to application data**

  - **Hadoop MapReduce: A software framework for distributed processing of large data sets on clusters**

MapReduce Data Flow

Hadoop MapReduce
Master = JobTracker
Worker=TaskTracker

Mapreduce Example: Word Count

Input files

cloud computing overview

heterogeneous high performance cloud computing

applications in space system acquisitions

applications development in the cloud

as a service

Map Phase

worker

worker

worker

Intermediate files

heterogeneous,1
high,1
performance,1
cloud,1
computing,1

Service,1

the,1
cloud,1

Reduce Phase

worker

development,1
dynamics,1
flight,1
heterogeneous,1
high,1
in,2

system,1
service,1
the,1

Output files

INTELSAT.

HDFS

- **Very large distributed file system on commodity hardware**

- **Designed to scale to petabytes of storage**

- **Single Namespace for entire cluster**

    - **Optimized for batch processing**

    - **Write-once-read-many access model**

    - **Read, write, delete, append-only**

- **NameNode handles replication, deletion, creation**

- **DataNode handles data retrieval**

HDFS

- **Files are broken up into blocks**

  - **Block size is 64 MB (can be modified)**

  - **Blocks replicated on at least 3 nodes**

  - **If a node fails, blocks get replicated automatically to other nodes**

| NameNode | ⟷ | BackupNode |

DataNode   DataNode   DataNode   DataNode   DataNode
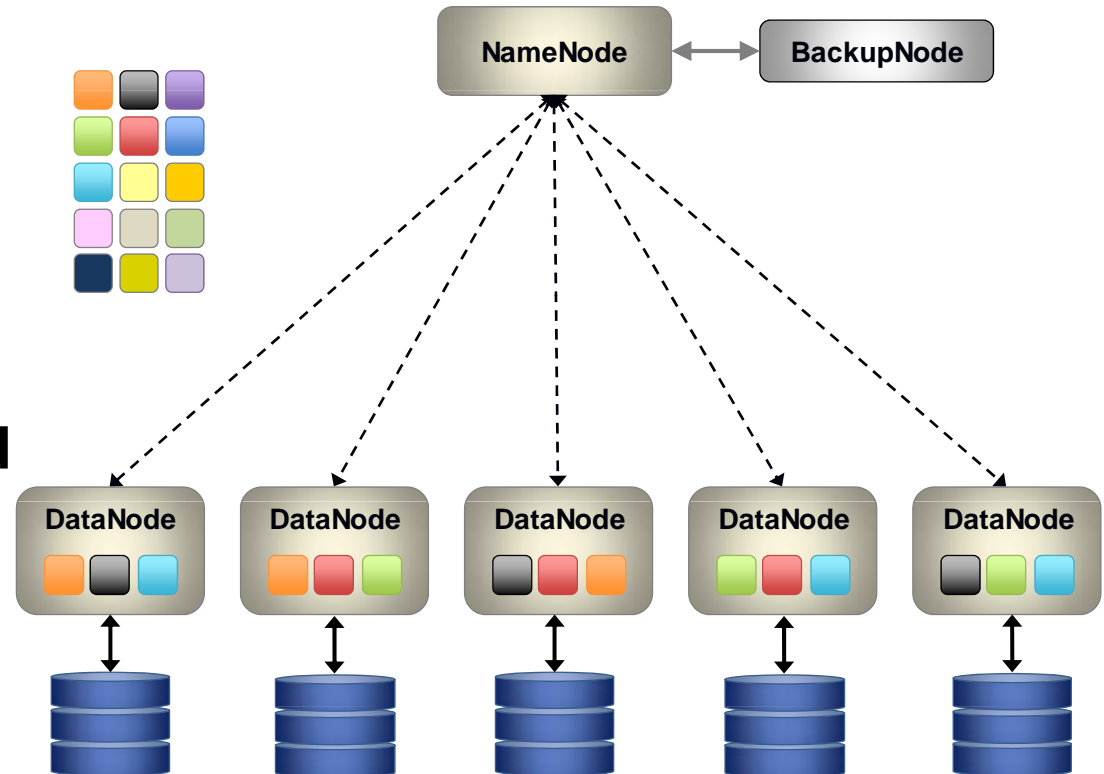
Hadoop Cluster at Yahoo!

- **Largest cluster:**
  - **4000 nodes, 16PB raw disk, 64TB RAM**

Who Uses Hadoop?

- **Adobe**
- **AOL**
- **Disney**
- **Ebay**
- **Facebook**
- **Google**
- **Hulu**
- **IBM**

- **LinkedIn**
- **New York Times**
- **Powerset/Microsoft**
- **Rackspace**
- **StumbleUpon**
- **Twitter**
- **Yahoo!**

Full list: http://wiki.apache.org/hadoop/PoweredBy

Examples

- **New York Times**
  - **Converted public domain articles from 1851-1922 to PDF**
  - **Stored 4 TB of scanned images as input on Amazon S3**
  - **Ran 100 Amazon EC2 instances for around 24 hours**
  - **Produced 11 Million articles, 1.5 TB of PDF files as output**
  - **http://open.blogs.nytimes.com/2007/11/01/self-service-prorated-super-computing-fun/**
- **Visa**
  - **Processing 73 billion transactions, amounting to 36 TB of data**
  - **Traditional methods: 1 month**
  - **Hadoop: 13 minutes**
  - **How Hadoop Tames Enterprises' Big Data. Sreedhar Kajeepeta, InformationWeek Reports, Feb 2012**

Sorting Benchmarks

- **Hadoop Sorts a Petabyte in 16.25 Hours and a Terabyte in 62 Sec**

  – **http://developer.yahoo.com/blogs/hadoop/posts/2009/05/hadoop_sorts_a_petabyte_in_162/onds**
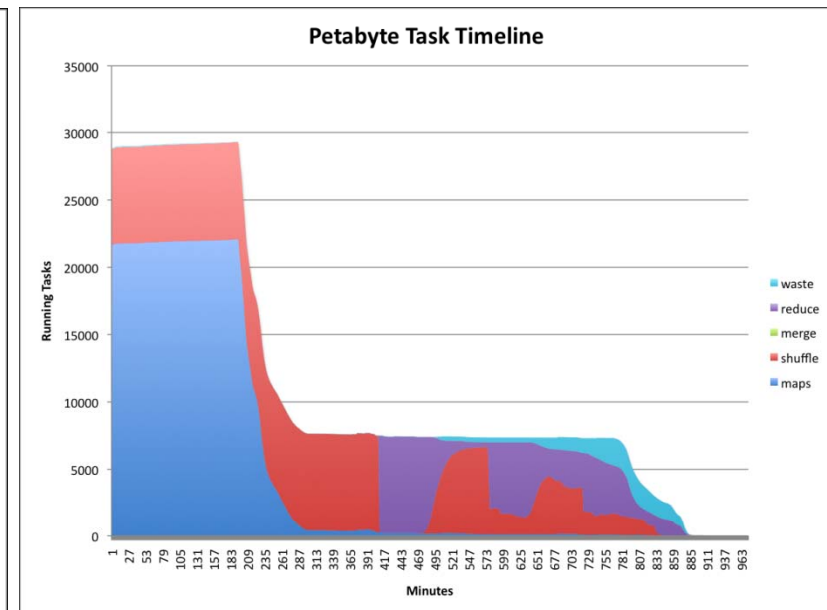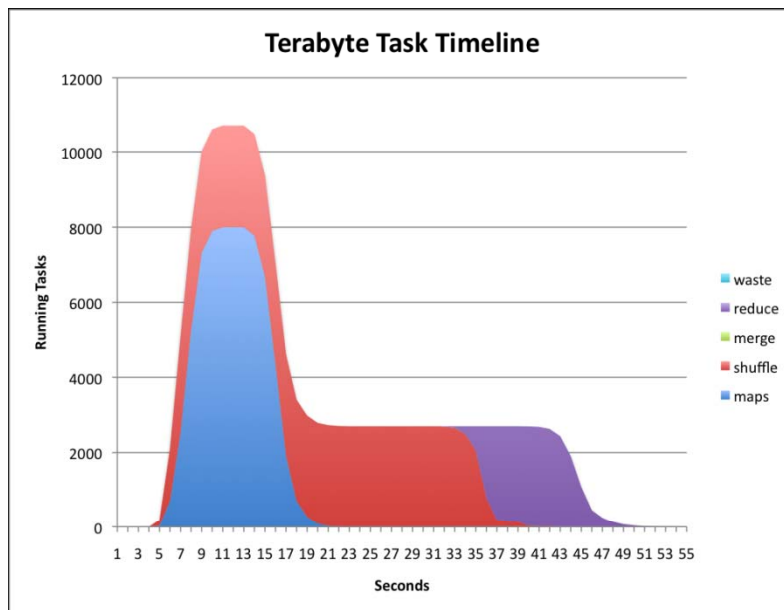
The Hadoop Ecosystem
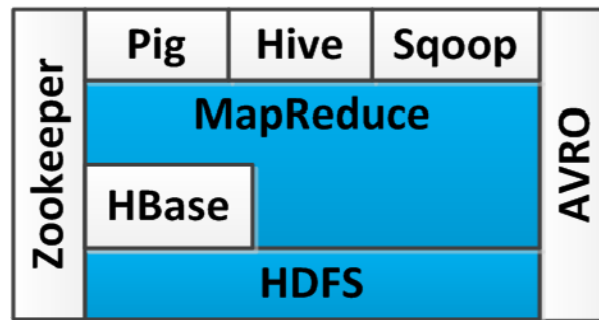
- **HBase™: Initiated by Powerset. A scalable, distributed database that supports structured data storage for large tables**

- **Hive™: Initiated by Facebook. A data warehouse infrastructure that provides data summarization and ad hoc querying**

- **Pig™: Initiated by Yahoo. A high-level data-flow language and execution framework for parallel computation**

- **ZooKeeper™: Initiated by Yahoo. A high-performance coordination service for distributed applications**

- **Avro™: A data serialization system**

HBase

- **Open source project modeled after Google's BigTable**

- **Distributed, large-scale data store**

- **Efficient at random reads/writes**

- **Use Hbase when you:**
  - **need to store large amounts of data**
  - **need high write throughput**
  - **need efficient random access within large data sets**
  - **need to scale gracefully with data**
  - **need to save time-series ordered data**
  - **don't need full RDMS capabilities (cross row/cross table transactions, joins, etc.)**

HBase Use Cases

- **Mozilla's Socorro - crash reporting system**
  - **2.5 Million crash reports/day**
- **Facebook**
  - **Real-time Messaging System: HBase to Store 135+ Billion Messages a Month**
  - **Realtime Analytics System: HBase to Process 20 Billion Events Per Day**
- **Twitter, Yahoo!, TrendMicro, StumbleUpon, …**
- **See http://wiki.apache.org/hadoop/Hbase/PoweredBy**

INTELSAT.

OpenTSDB
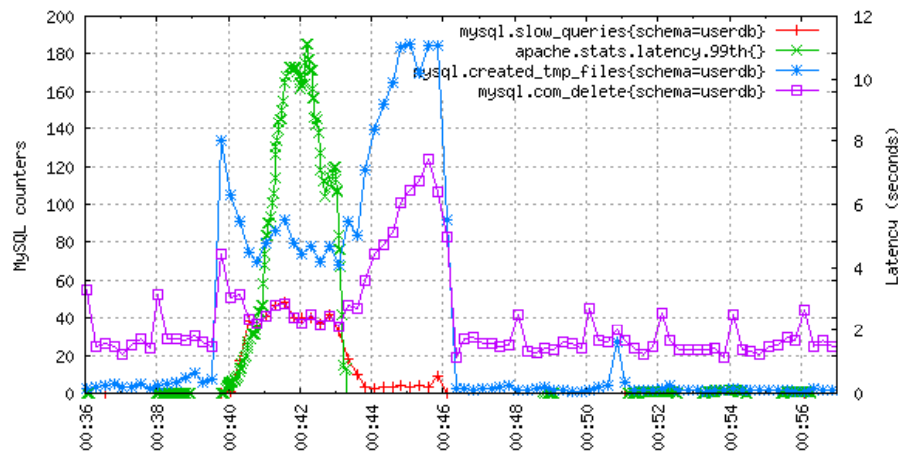
- 15464 points retrieved, 932 points plotted in 100ms
- Generating custom graphs and correlating events is easy

- **A distributed, scalable, time series database developed by StumbleUpon**
  - **Open-source monitoring system built on Hbase**
  - **It is also a data plotting system.**
  - **Used to monitor computers in their datacenter**
  - **Collects 600 million data points per day, over 70 billion data points per year**
- **Store them forever, retain granular data**

Typical Use Cases in Ground Systems Domain

- **Hadoop can be used to scale out applications that fit the "write-once-read-many" model**

- **Typical use cases:**

  - **Telemetry analysis and visualization**

  - **Real time displays**

  - **Storage and retrieval of spacecraft and ground system events**

  - **Payload data processing**

Summary

- **In order to capitalize on the benefits of the cloud environment, applications must be designed to "scale out" rather than "scale up"**

- **Robust open source tools to support computing on a cluster of computers are available and in use today**

- **The Hadoop Ecosystem is used in many enterprises**

- **The performance and scalability of ground systems can be improved by using the growing ecosystem of Hadoop tools and technologies**

- **Applications that fit the "write-once-read-many" model and process large amount of data benefit from this approach**

INTELSAT

References

- **Jeffrey Dean and Sanjay Ghemawat. "MapReduce: Simplified Data Processing on Large Clusters." Communications of the ACM, Volume 51, Issue 1, pp. 107-113, 2008**
- **Apache Hadoop: http://hadoop.apache.org/**
- **Apache Hadoop: an introduction. Todd Lipcon, May 27, 2010**
- **Hadoop Update, Open Cirrus Summit 2009.Eric Baldeschwieler, VP Hadoop Development, Yahoo!**
- **Cloudera Training: http://www.cloudera.com/hadoop-training**
- **Realtime Apache Hadoop at Facebook. Dhruba Borthakur, Oct 5, 2011 at UC Berkeley Graduate Course**
- **Apache Hbase: hbase.apache.org**
- **http://www.slideshare.net/kevinweil/nosql-attwitter-nosql-eu-2010**
- **http://www.slideshare.net/ydn/3-hadooppigattwitterhadoopsummit2010**
- **OpenTSDB, The Distributed, Scalable, Time Series Database. Benoît "tsuna" Sigoure, StumbleUpon**