




## **La convergence du stockage et d'Hadoop pour l'entreprise**

*Xavier Guerin VP Southern Europe & Benelux, MapR Technologies*

- 
- *“The use of big data will become a key basis of competition and growth for individual firms. In most industries, established competitors and new entrants alike will leverage data-driven strategies to innovate, compete, and capture value from deep and up-to-real-time information.”*  
– **McKinsey & Company**
  - *“The size, complexity of formats and speed of delivery exceeds the capabilities of traditional data management technologies”*  
– **Gartner**
  - *“The bringing together of a vast amount of data from public and private sources, combined with the intuition of business and thought leaders and the speed and affordability of today's computers, is what Big Data is all about.”*  
– **IDC**

# The Hadoop Advantage

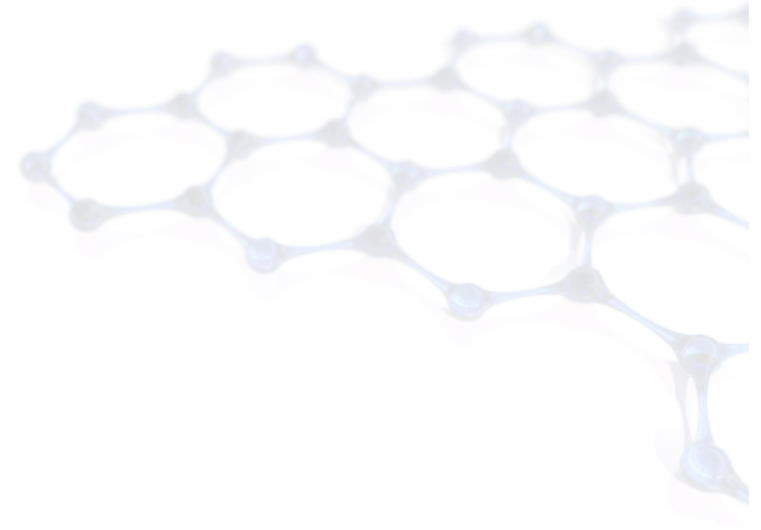
- Fueling an industry revolution by providing infinite capability to store and process big data
- Expanding analytics across data types
- Compelling economics
  - 20 to 100X more cost effective than traditional enterprise storage or data warehouses



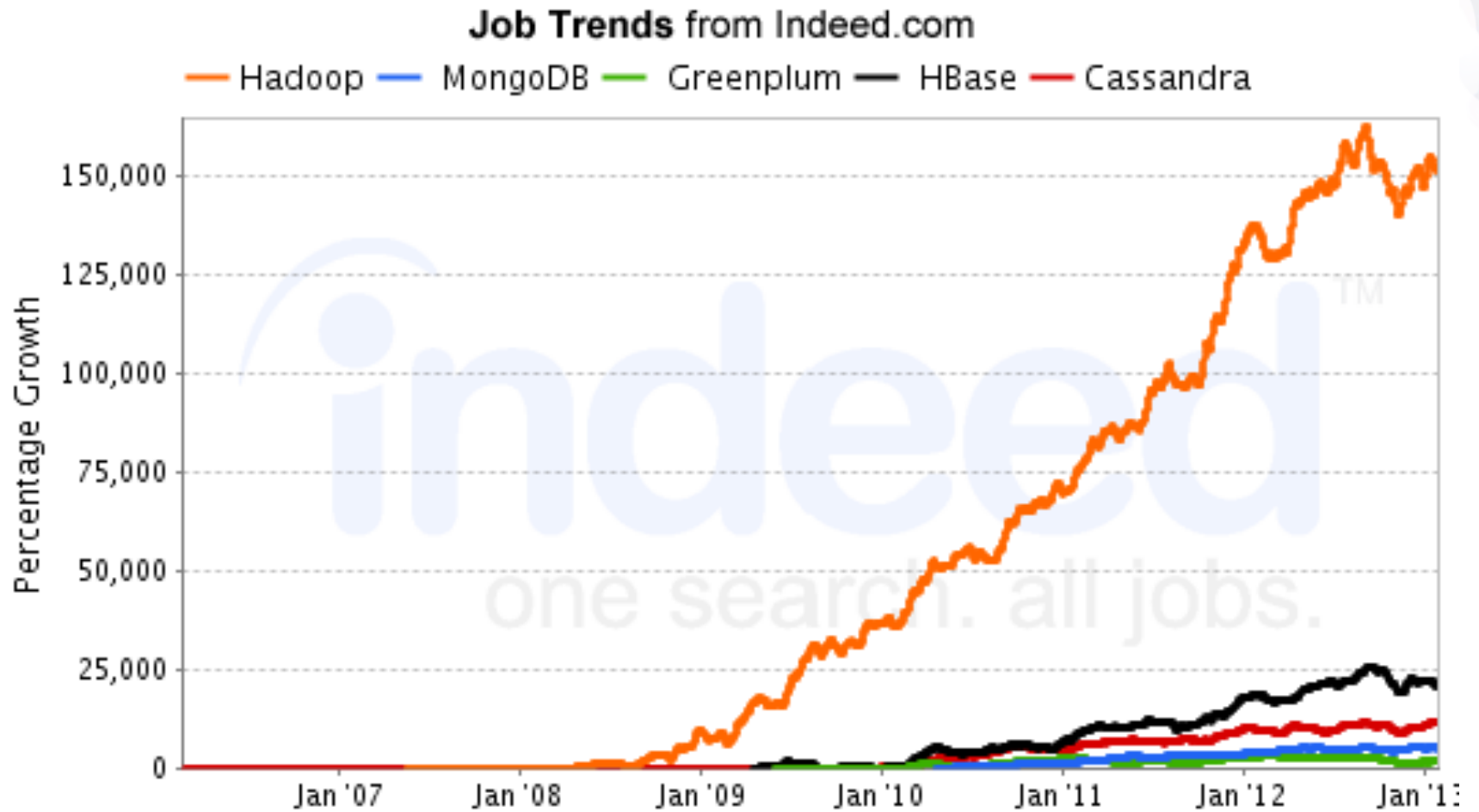
Pioneered at **Google**<sup>™</sup>

# Biggest challenges

- 80% Finding talent
- 72% Training and education
- 76% Identifying the correct tools
- 32% Siloed data, non cooperation



# Hadoop Growth



# Professional Services and Training

- Consulting
  - Use Case Identification
  - Architecture/Solution Design
  - Implementation and Migration
  - Leverage MapR Partners
- MapR Academy
  - Hands-on Classes
  - Self-service/Classroom/Onsite
  - Administrator/Developer Training
    - Hive/Pig/MapReduce
    - HBase



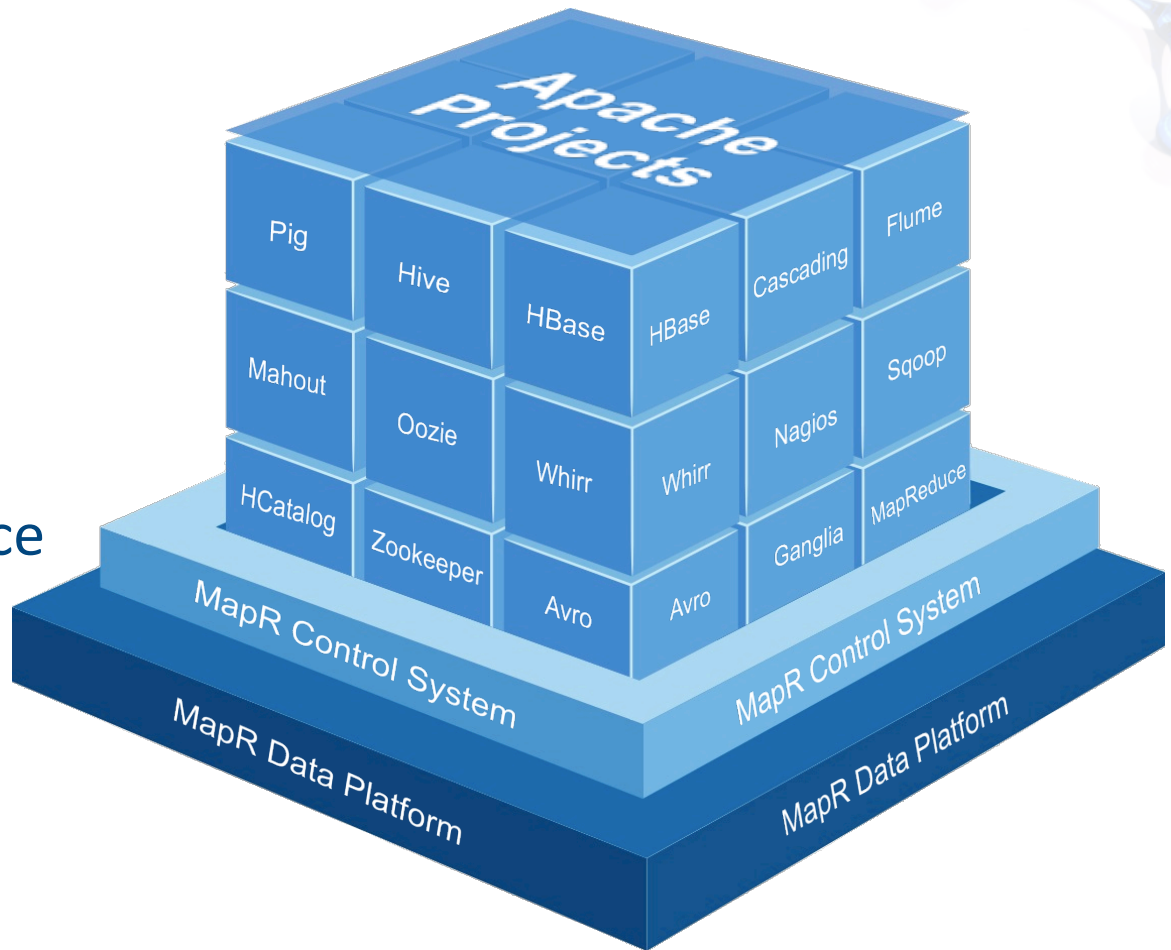
# MapR Technologies Overview

- Enterprise-grade platform for Hadoop
  - Founders from Google, Microsoft, IBM, EMC
- Deployed at 1000s of companies
- Massive amounts of data analyzed on MapR
  - Hundreds of billions of events daily
  - 90% of the world's Internet population monthly
  - \$1 trillion in retail purchases annually



# MapR Distribution for Apache Hadoop

- Complete Hadoop distribution
- Comprehensive management suite
- Industry-standard interfaces
- Combines open source packages with enterprise-grade dependability
- Higher performance





# One Platform for Big Data

**Batch**  
Log File Analysis  
Data Warehouse Offload  
Fraud Detection  
Clickstream Analytics

**Interactive**  
Forensic Analysis  
Analytic Modeling  
BI User Focus

**Real-Time**  
Sensor Analysis  
"Titterscraping"  
Telematics  
Process Optimization

Batch

Interactive

Real-time



...

99.999%  
HA

Data  
Protection

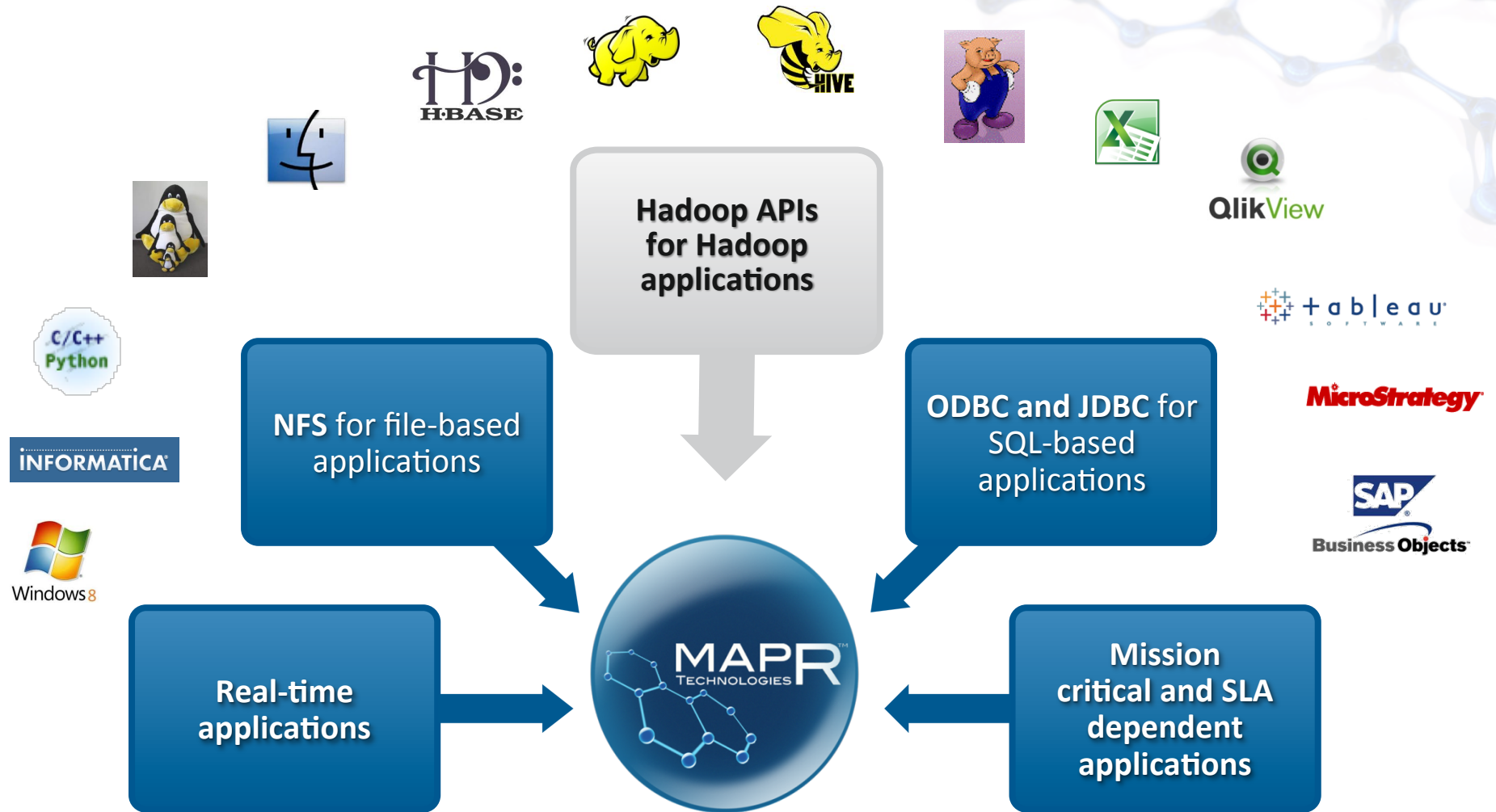
Disaster  
Recovery

Scalability  
&  
Performance

Enterprise  
Integration

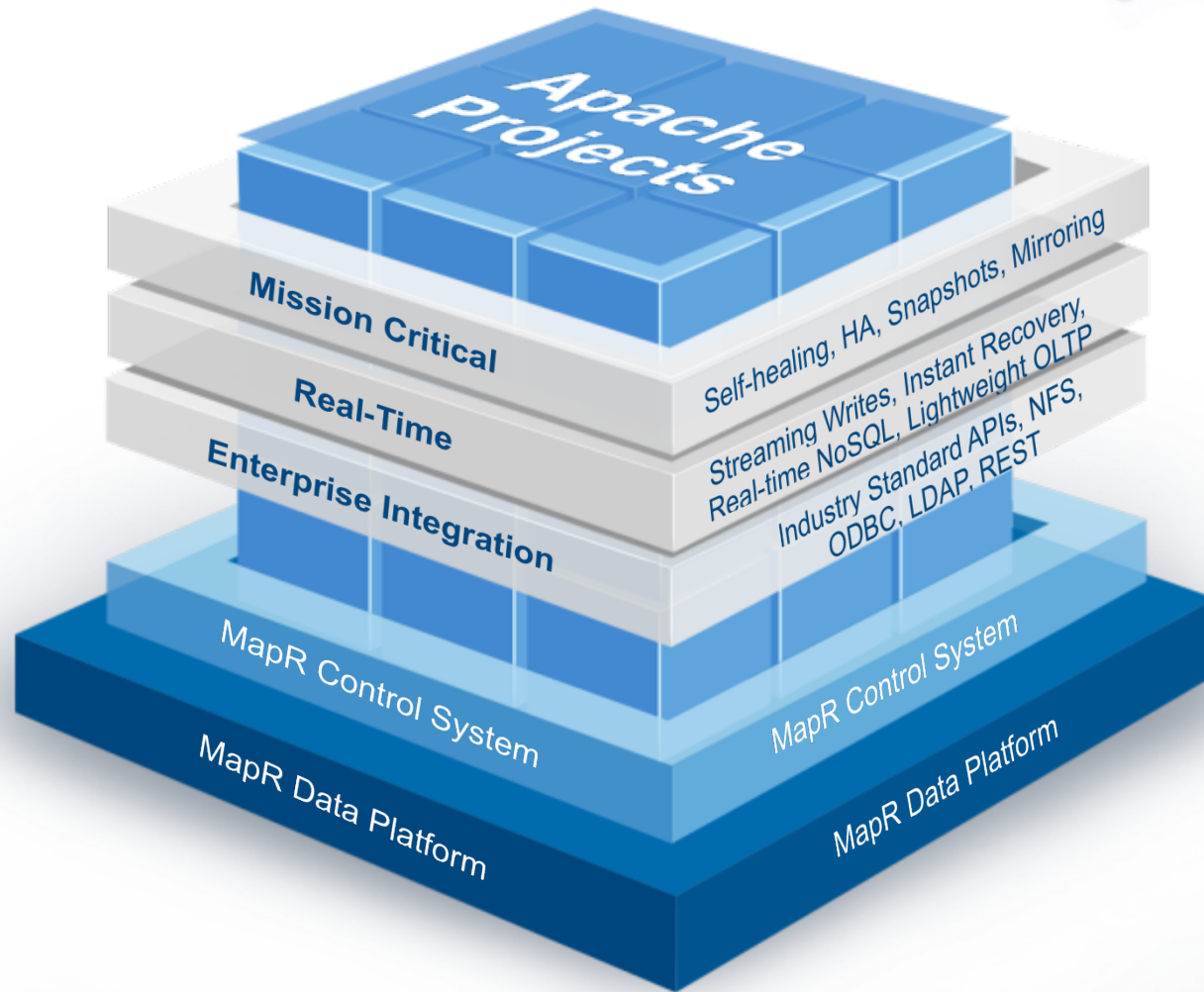
Multi-  
tenancy

# Easy Integration with the Enterprise



Blue = MapR Innovations

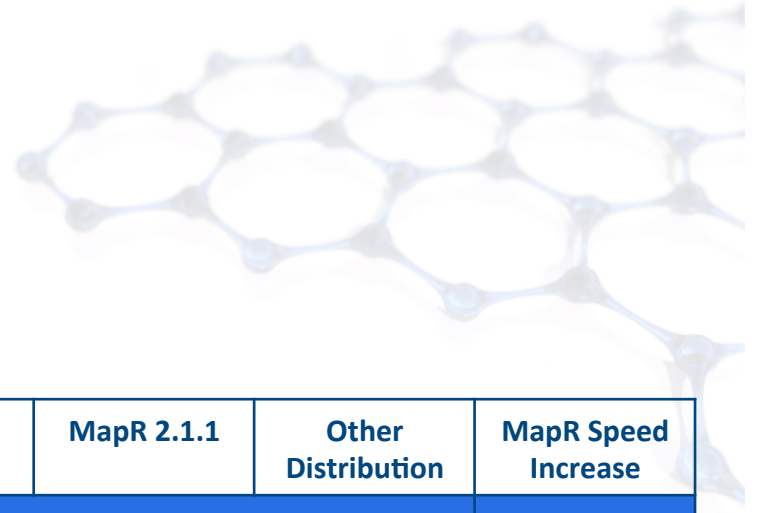
# MapR: The Enterprise-Grade Distribution



# The Convergent Platform

	MAPR	OTHERS
Open Data Platform with POSIX and NFS	✓	X
99.999% High Availability	✓	X
Disaster Recovery	✓	X
Data Protection	✓	X
Unlimited Scalability	✓	X
2x Performance	✓	X
Multi-tenancy	✓	X
Enterprise-grade NoSQL	✓	X
No Vendor Lock-In	✓	✓

# World Record Performance



**New MinuteSort  
World Record  
1.5 TB in 1 minute  
2103 nodes**

**Previous Record: 1.4 TB**



Benchmark	MapR 2.1.1	Other Distribution	MapR Speed Increase
<b>Terasort (1x replication, compression disabled)</b>			
Total	13m 35s	26m 6s	<b>1.9x</b>
Map	7m 58s	21m 8s	<b>2.7x</b>
Reduce	13m 32s	23m 37s	<b>1.7x</b>
<b>DFSIO throughput/node</b>			
Read	1003 MB/s	656 MB/s	<b>1.5x</b>
Write	924 MB/s	654 MB/s	<b>1.4x</b>
<b>YCSB (50% read, 50% update)</b>			
Throughput	36,584.4 op/s	12,500.5 op/s	<b>2.9x</b>
Runtime	3.80 hr	11.11 hr	<b>2.9x</b>
<b>YCSB (95% read, 5% update)</b>			
Throughput	24,704.3 op/s	10,776.4 op/s	<b>2.3x</b>
Runtime	0.56 hr	1.29 hr	<b>2.3x</b>

# MapR Supports Broad Set of Customers

## Global Credit Card Issuer

- Recommendation Engine
- Fraud detection and Prevention

## Leading Retailer

- Customer Behavior Analysis
- Brand Monitoring



- Customer targeting
- Viewer Behavioral analytics



- Intrusion detection & prevention
- Forensic analysis



- Recommendation Engine
- Family tree connections



- Global threat analytics
- Virus analysis



- Patient care monitoring



- Log analysis
- HBase



- Clickstream Analysis
- Quality profiling/field failure analysis



- Fraud Detection
- Channel analytics



- Advertising exchange analysis and optimization



- Monitoring and measuring online behavior



- Customer Revenue Analytics
- ETL Offload



- Customer targeting
- Social media analysis



- Enterprise Grade Platform
- COOP features

# Key Strategic Partners



# Industry Leaders Choose MapR in the Cloud



Amazon EMR is the largest Hadoop provider in revenue and # of clusters

Google chose MapR to provide Hadoop on Google Compute Engine



# MapR Editions

## MAPR M3 FREE

- Control System
- NFS Access
- Performance
- Unlimited Nodes
- Free

## MAPR M5 EDITION

- Control System
- NFS Access
- Performance
- High Availability
- Snapshots & Mirroring
- 24 X 7 Support
- Annual Subscription

## MAPR M7 EDITION

- All the Features of M5
- Simplified Administration for HBase
- Increased Performance
- Consistent Low Latency
- Unified Snapshots, Mirroring

Also available through:



**Thank You**

