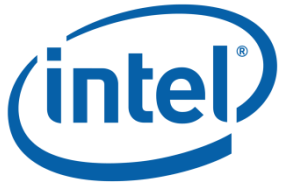


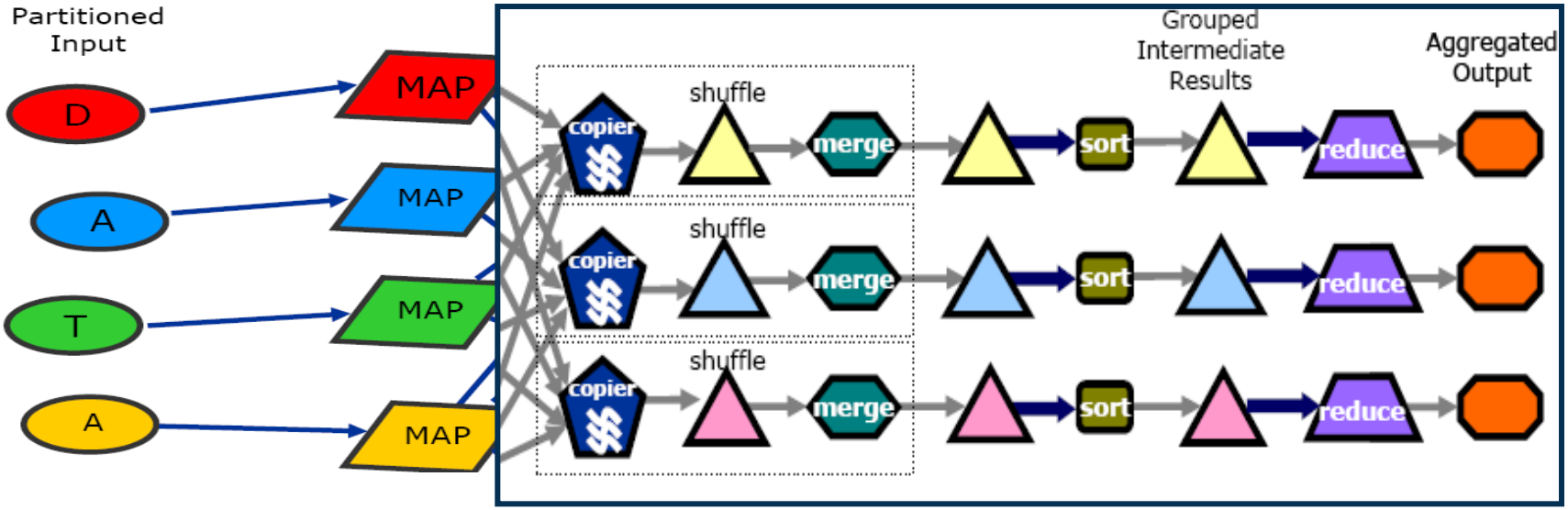
Benchmarking & Optimizing Hadoop

Agenda

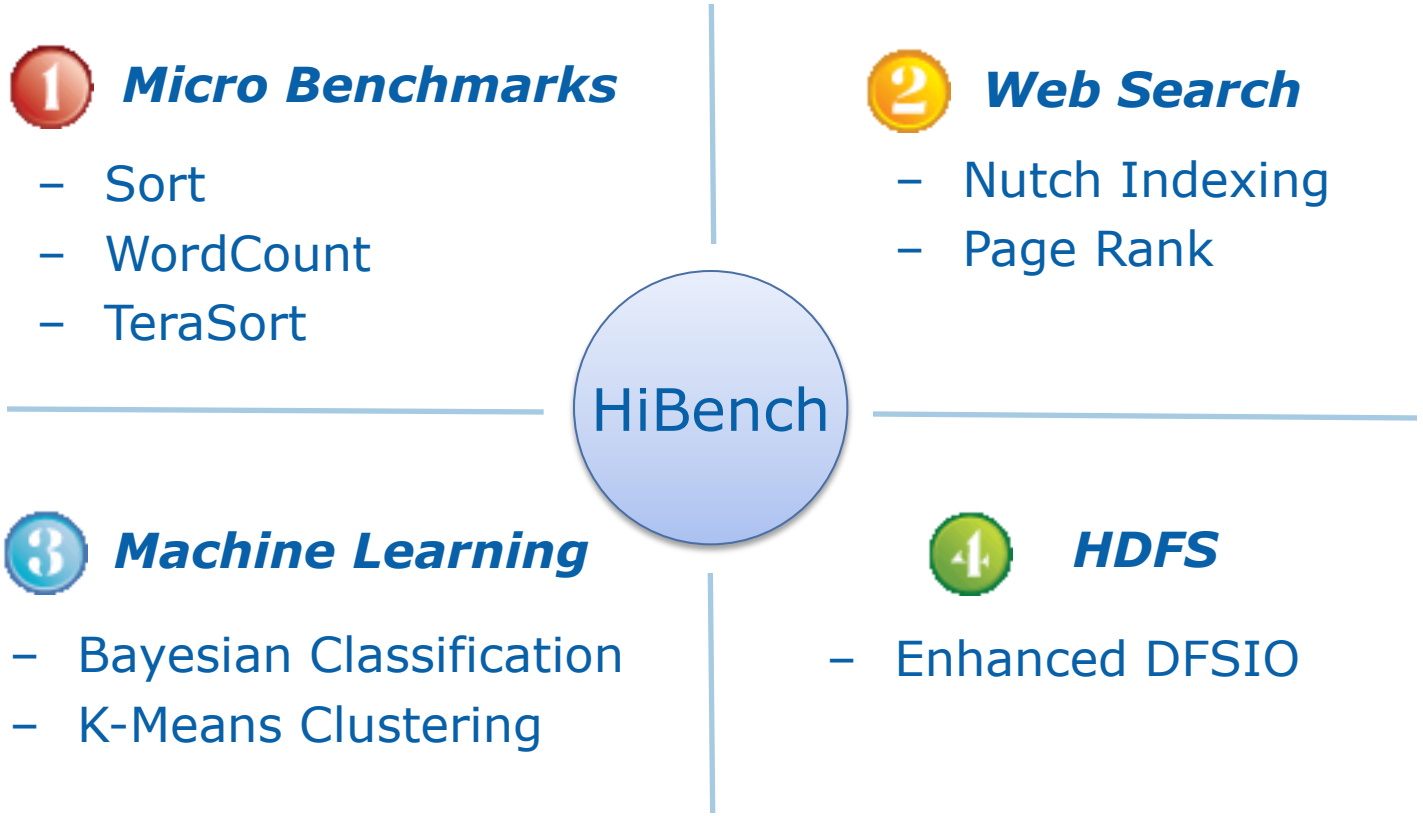
- MapReduce/Hadoop
- HiBench: The Benchmark Suite for Hadoop
- Using HiBench: Characterization & Evaluation
- Optimizing Hadoop Deployments

MapReduce/Hadoop

- MapReduce
  - Essentially a group-by-aggregation in parallel
  - Batch-style, throughput-oriented, data-parallel
- Hadoop
  - Most popular open source implementation of MapReduce



HiBench : A Benchmark Suite for Hadoop



A Comprehensive & Realistic Benchmark Suite

Using HiBench : Characterization & Evaluation

Characterization

- Understand the typical behavior of real-world Hadoop applications
- Understand the Hadoop framework and data flow model

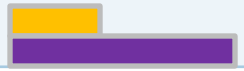
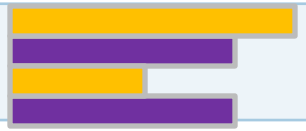
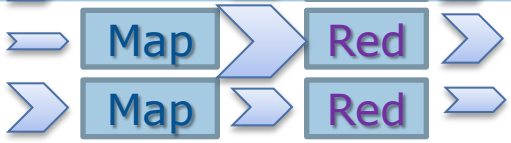
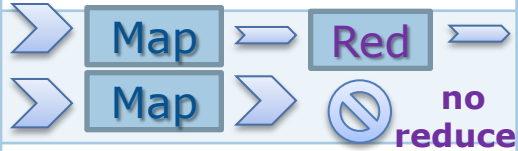
Evaluation on different server platforms

- Measure/Compare the performance of specific deployments
- Find out the bottleneck of certain deployment
- Find out the power efficiency of certain deployment choices

Evaluation on different Hadoop versions

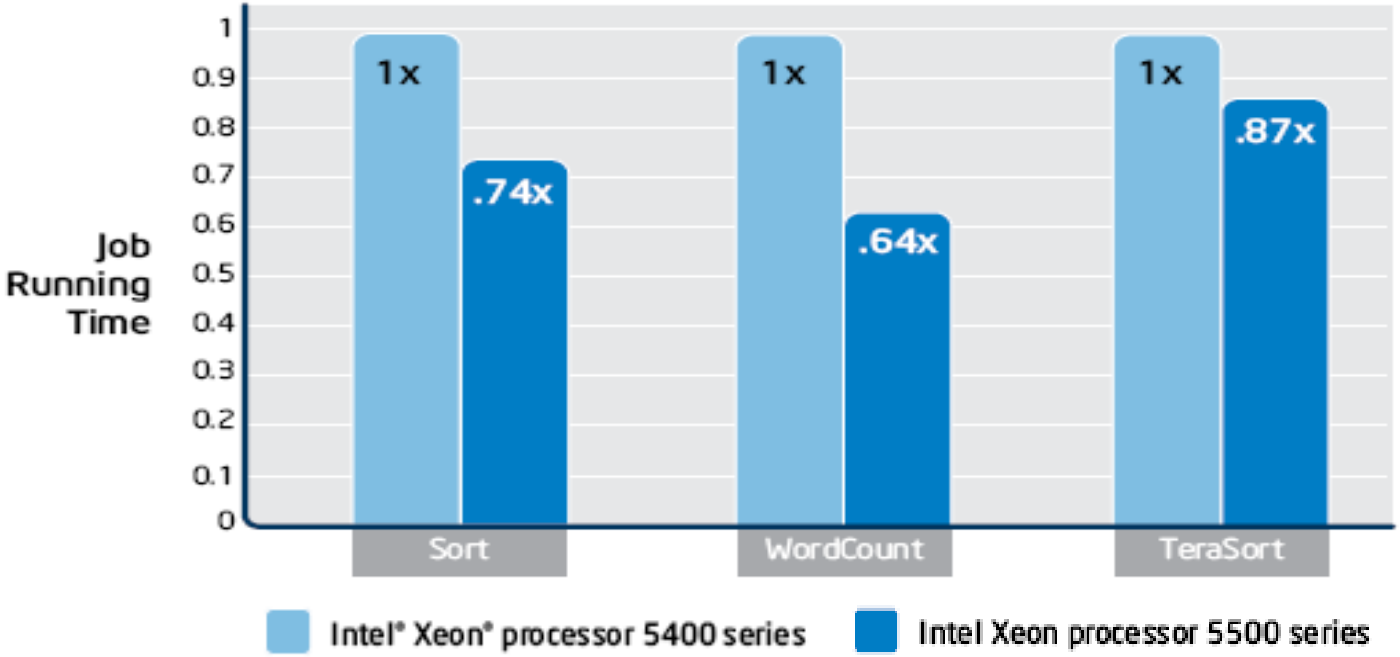
- Analyze the performance impact of new features and optimizations in newer versions

Characterization Results

Workload	System Resource Utilization	Data Access Patterns	Map/Reduce Stage Time Ratio
Sort	I/O bound		
WordCount	CPU bound		
TeraSort	Map stage : CPU-bound; Red stage : I/O-bound		
Nutch Indexing	I/O bound, high CPU utilization in map stage		
Page Rank (1st & 2nd job)	CPU-bound in all jobs		
Bayesian Classification (1st & 2nd job)	I/O bound, with high CPU utilization in map stage in the 1st job		
K-means Clustering	CPU bound in iteration; I/O bound in clustering		
Enhanced DFSIO	I/O-bound	trivial	trivial

Evaluation Results : Server Platforms
In Terms of Speed

HiBench Comparison
Between Two Generations of Intel® Xeon® Platforms
Speed Test (Lower Values are Better) *

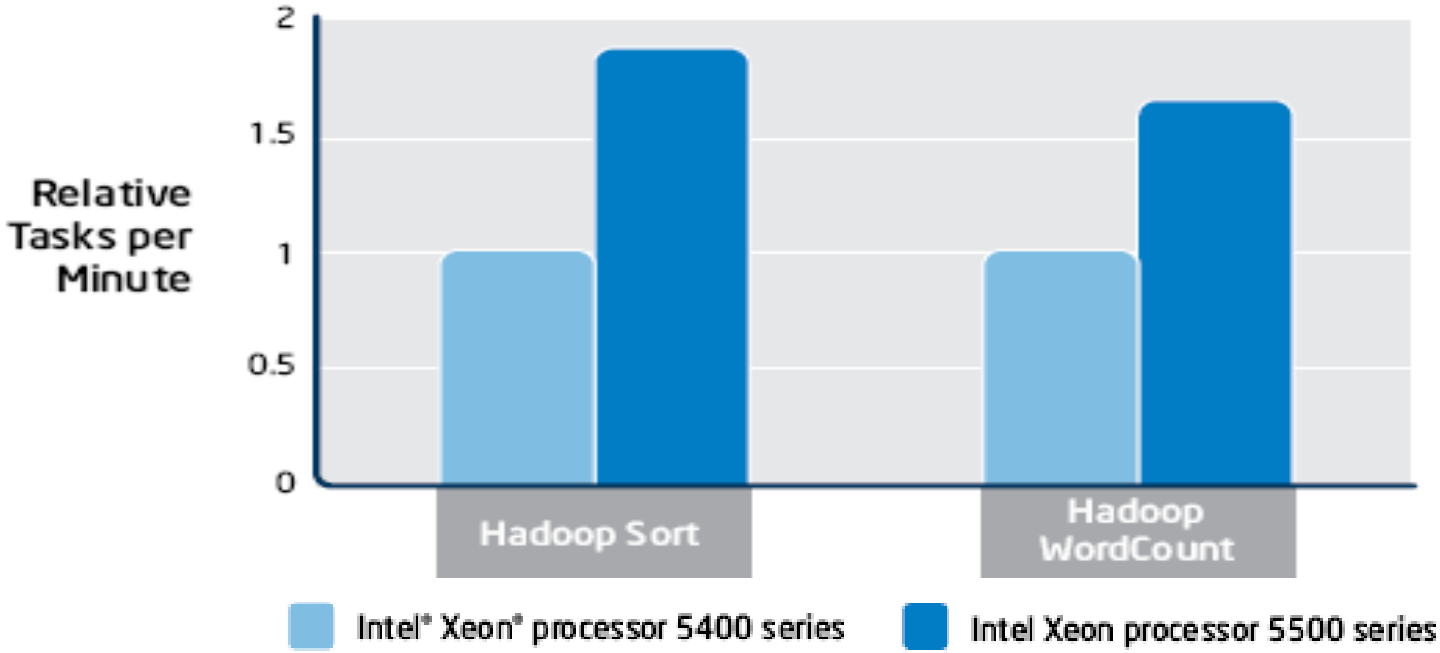


- The smaller the job running time is, the faster the job runs.
- The newer platform is up to 56% faster than the older platform.

[*] See more details in Intel Whitepaper: Optimization Hadoop Deployments, Available at : http://communities.intel.com/docs/DOC-4218

Evaluation Results : Server Platforms
In Terms of Throughput

HiBench Comparison
Between Two Generations of Intel® Xeon® Platforms
Throughput Test (Higher Values are Better)



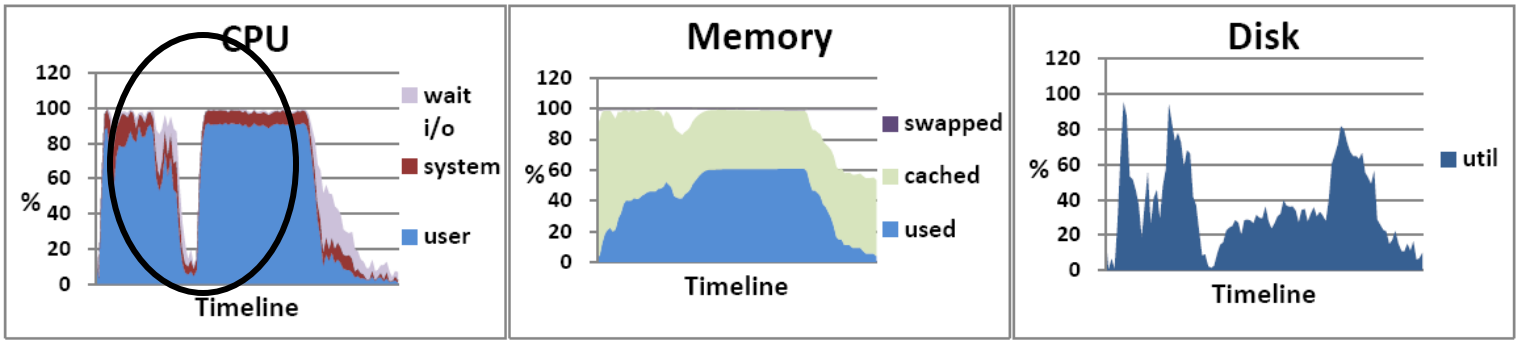
- Throughput = # of tasks completed / minute when cluster is at 100% utilization.
- The new platform provides up to 86% more throughput than the older platform.

[*] See more details in Intel Whitepaper: Optimization Hadoop Deployments, Available at : http://communities.intel.com/docs/DOC-4218

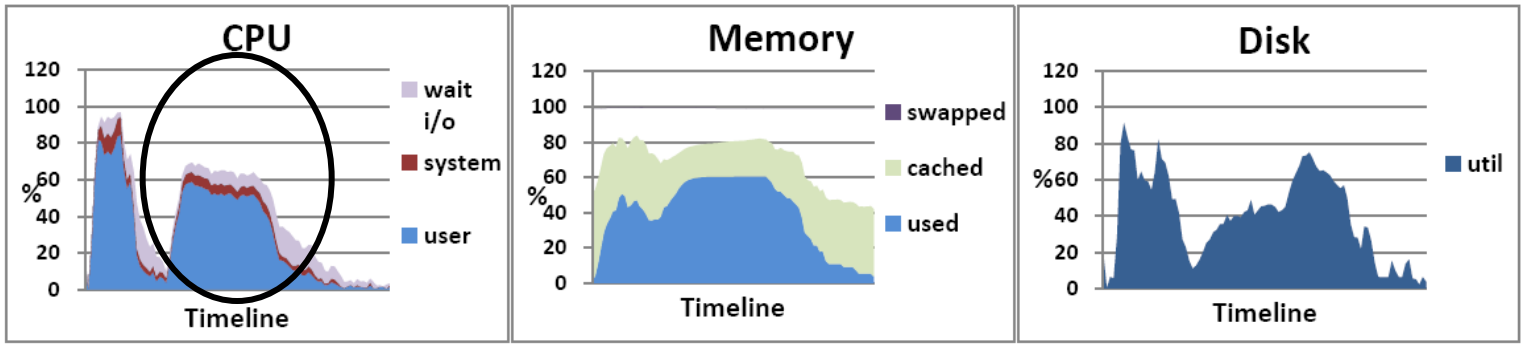
Evaluation Results : Server Platforms
In Terms of System Resource Utilization

HiBench Comparison
Between Two Generations of Intel® Xeon® Platforms
(CPU, Memory, Disk Utilization)

Intel Xeon 5400 series



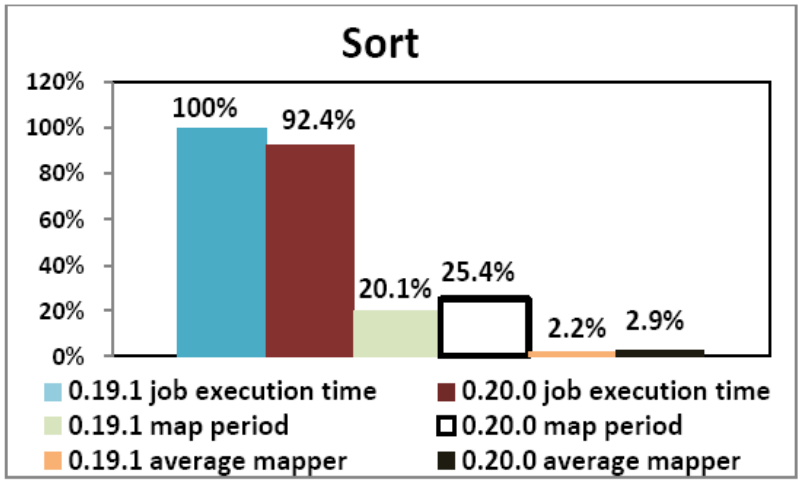
Intel Xeon 5500 series

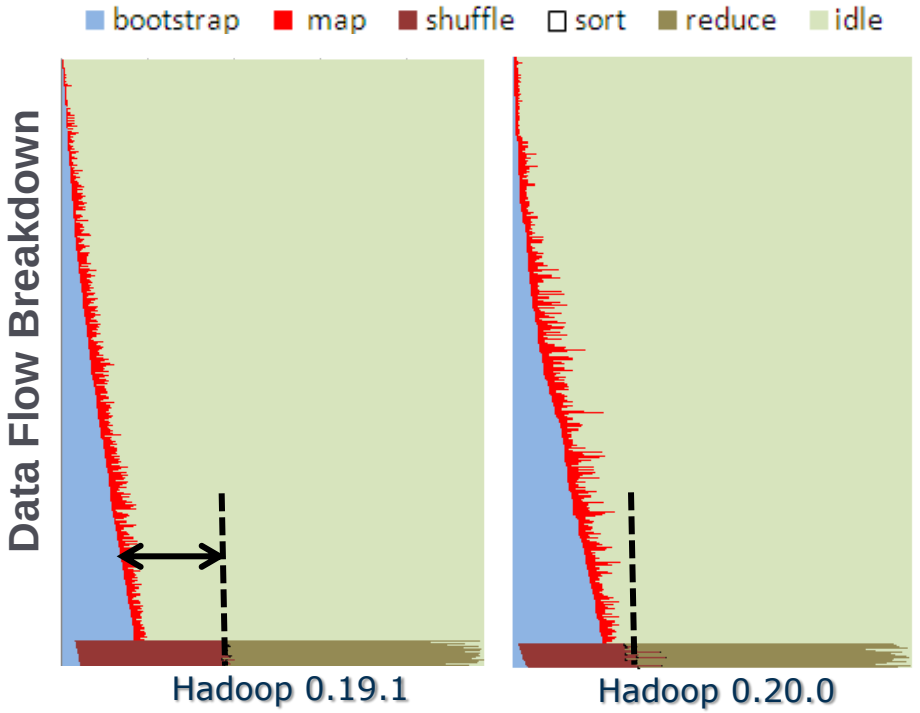


- Resource bottleneck may vary with different Hadoop deployment.
- In Reduce Stage, disk I/O is the bottleneck for Intel Xeon 5570 platform, while CPU is the bottleneck for 5400 series

Evaluation Results : Hadoop Versions

HiBench Comparison
Between Two Hadoop Versions (v0.19.1 and v0.20.0)
MapReduce Dataflow Model (Sort)

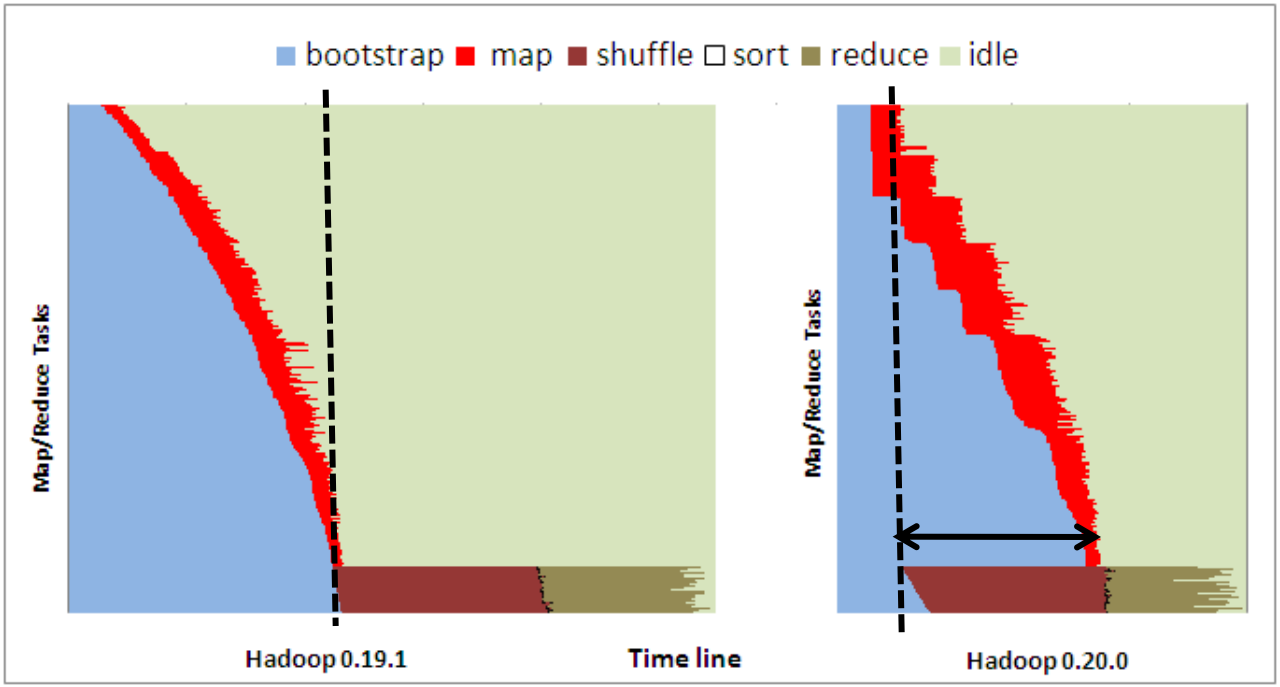




- Even the map stage is slower, the entire job is faster in 0.20.0 than 0.19.1.
- An improvement in the shuffle copying makes shuffle finish faster.

Evaluation Results : Hadoop Versions

HiBench Comparison
Between Two Hadoop Versions (v0.19.1 and v0.20.0)
MapReduce Dataflow Breakdown (Bayesian 2nd job)



- In v0.19.1, most of reduce tasks cannot start until all map tasks are done
- In v0.20.0. the improved task scheduler helps the reducers to start earlier.

Optimizing Hadoop* Deployments – Hardware

- Server Platform
  - Choose the optimal server platform for cost performance
  - Leverage the most current platform technologies
  - Use power optimized server board
- Memory
  - Supply sufficient memory for parallelism
  - ECC memory is recommended
- Hard Disk
  - RAID is not needed
  - NCQ makes the disk access faster
  - Solid State Disks are much faster and power efficient.

[*] See more details in Intel Whitepaper: Optimization Hadoop Deployments, Available at : http://communities.intel.com/docs/DOC-4218

Optimizing Hadoop* Deployments – Software

- Operation System
  - Use kernel version 2.6.30 or later
  - Optimize Linux configurations (e.g. noatime)
- Application Software
  - Use latest Java and proper JVM settings for servers
  - Use optimal Hadoop distributions
- Hadoop Configuration Tuning
  - Number of simultaneous map/reduce tasks
  - HDFS settings (block size, replication factor)
  - Tradeoff between system resources

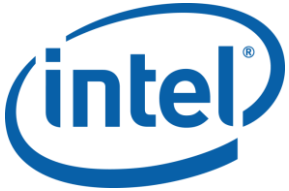
[*] See more details in Intel Whitepaper: Optimization Hadoop Deployments, Available at : http://communities.intel.com/docs/DOC-4218

Summary

- MapReduce/Hadoop becomes increasingly popular in Cloud.
- HiBench is a comprehensive and realistic benchmark suite for Hadoop; it can be used to evaluate Hadoop deployments and demonstrate Hadoop framework characteristics
- Based on our evaluation results, we provided suggestions for optimal Hadoop hardware and software configurations, which may help organizations to make deployment choices in the planning stage.

Q&A

Thanks



HiBench Workloads Details

Category	Workload	Why included
Micro Benchmarks	**Sort**	✓Representative of typical MapReduce jobs ✓Demonstrates intrinsic characteristics of MapReduce model
	WordCount	✓Representative of typical MapReduce jobs ✓Demonstrates another typical usage of MapReduce model
	TeraSort	✓A standard benchmark for large size data sorting (used by Google and Yahoo to demonstrate the power of their MapReduce clusters publicly)
Web Search	**Nutch Indexing**	✓Typical application area of MapReduce – text tokenization, indexing and search ✓Large scale indexing system is one of the most significant uses of MapReduce (e.g., in Google and Facebook)
	Page Rank	✓Typical application area of MapReduce – Web searching. ✓Page Rank is a popular web page ranking algorithm.
Machine Learning	**Mahout K-Means Clustering**	✓ Typical application area of MapReduce – large-scale data mining and machine learning (e.g., used in Google and Facebook) ✓ K-Means is a well-known clustering algorithm
	Mahout Bayesian Classification	✓ Typical application area of MapReduce – large-scale data mining and machine learning ✓ Bayesian is a well-known classification algorithm
HDFS	**Enhanced DFSIO**	✓Test the HDFS throughput of Hadoop cluster (e.g., original DFSIO has been used by Yahoo to evaluate their 4000-node Hadoop cluster publicly)

Cluster Configuration Information

Each cluster is configured with 1 master (running JobTracker and NameNode) and 4 slave nodes (running TaskTracker and DataNode), the server platform configuration of each slave node is listed below:

Intel® Xeon® X5460-based server

Processor: Dual-socket quad-core Intel® Xeon® X5460 3.16GHz

Processor Memory: 16GB (DDR2 FBDIM ECC 667MHz) RAM

Storage: 1 X 300GB 15K RPM SAS disk for system and log files, 4 X 1TB 7200RPM SATA for HDFS and intermediate results

Network: 1 Gigabit Ethernet NIC

BIOS: BIOS version S5000.86B.10.60.0091.100920081631EIST (Enhanced Intel SpeedStep Technology) disabled both hardware prefetcher and adjacent cache-line, prefetch disable

Intel® Xeon® X5570-based server

Processor: Dual-socket quad-core Intel® Xeon® X5570 2.93GHz

Processor Memory: 16GB (DDR3 ECC 1333MHz) RAM

Storage: 1 X 1TB 7200RPM SATA for system and log files, 4 X 1TB 7200RPM SATA for HDFS and intermediate results

Network: 1 Gigabit Ethernet NIC

BIOS: BIOS version 4.6.3 Both EIST (Enhanced Intel SpeedStep Technology) and Turbo mode disabled both hardware prefetcher and adjacent cache-line prefetch enabled, SMT (Simultaneous MultiThreading), enabled (Disabling hardware prefetcher and adjacent cache-line prefetch helps improve Hadoop performance on Xeon X5460 server according to our benchmarking.)