Comparing Application Performance on HPC-based Hadoop Platforms with Local Storage and Dedicated Storage

Zhuozhao Li[*], Haiying Shen[*], Jeffrey Denton[§] and Walter Ligon[§]

[*]Department of Computer Science, University of Virginia, USA

[§]Department of Electrical and Computer Engineering, Clemson University, USA

Outline

- Introduction of motivations
- System configuration
- Measurement study
- Performance Evaluation
- Conclusion and remark

Introduction

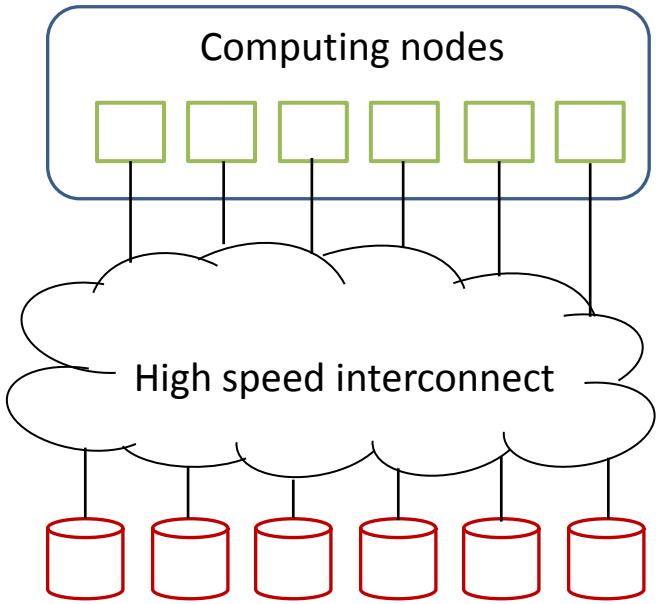
- Big Data Analytics
  - process several petabytes of data every day
  - run tens of thousands of jobs
  - Important to improve the performance
- MapReduce
  - distributed, parallel
  - data-intensive application
  - a cluster of computing nodes
- Hadoop
  - Facebook and Yahoo

Introduction

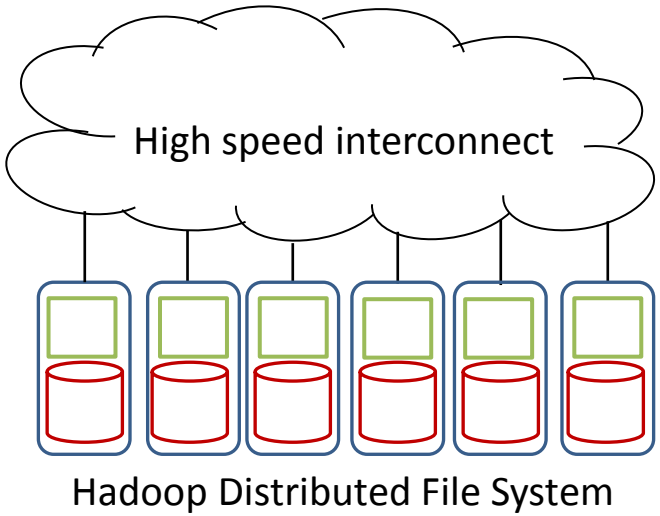
- High-performance computing (HPC) clusters are widely adopted to support CPU-intensive applications.
- HPC clusters also need to process data-intensive workloads.
- Many high-performance computing (HPC) sites extended their clusters to support Hadoop MapReduce.
- However, several settings are different between HPC and traditional data analytic clusters.

Introduction

- File systems?
  - HDFS and HPC remote file system



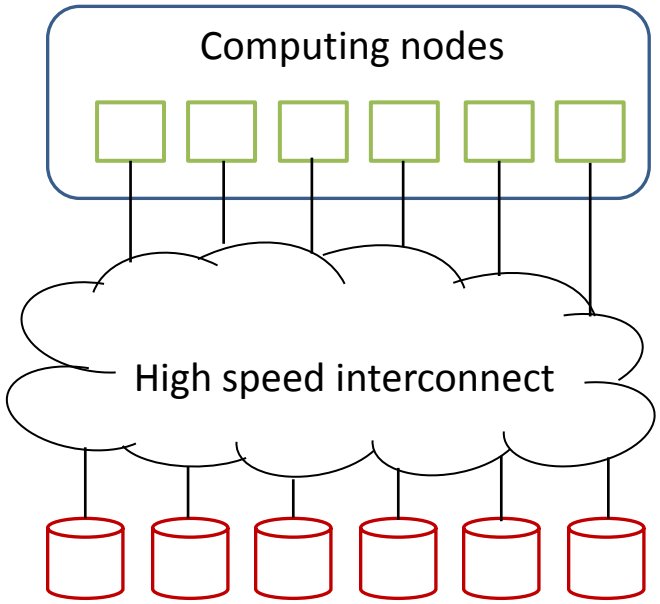
(a) A typical HPC cluster



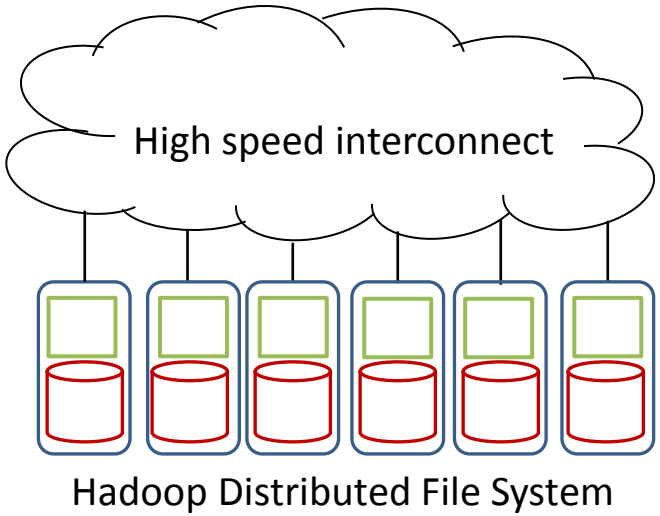
(b) A Hadoop cluster

Introduction

- Clemson Palmetto HPC cluster successfully configured Hadoop by replacing the local HDFS with the remote Orange File System (OFS).



(a) A typical HPC cluster



(b) A Hadoop cluster

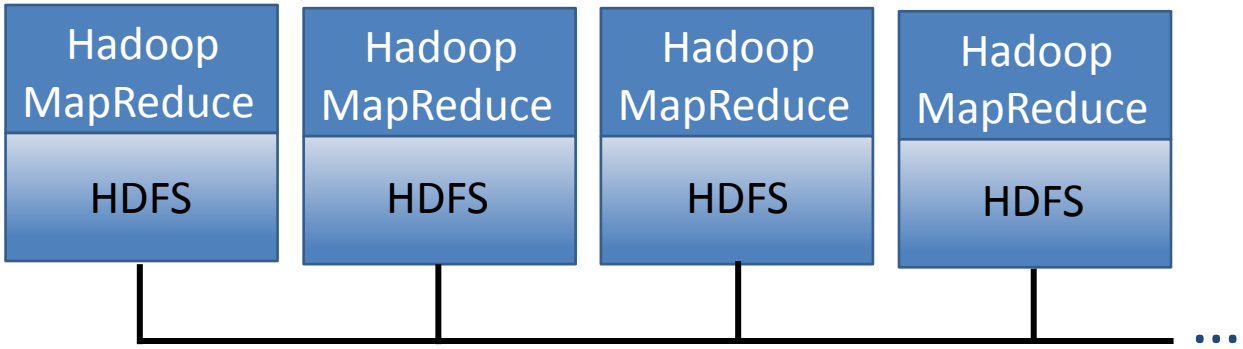
Goal

- Real MapReduce workload
  - A real world workload consists of many different types of applications with different job characteristics (data-intensive, CPU-intensive, I/O-intensive)[1].

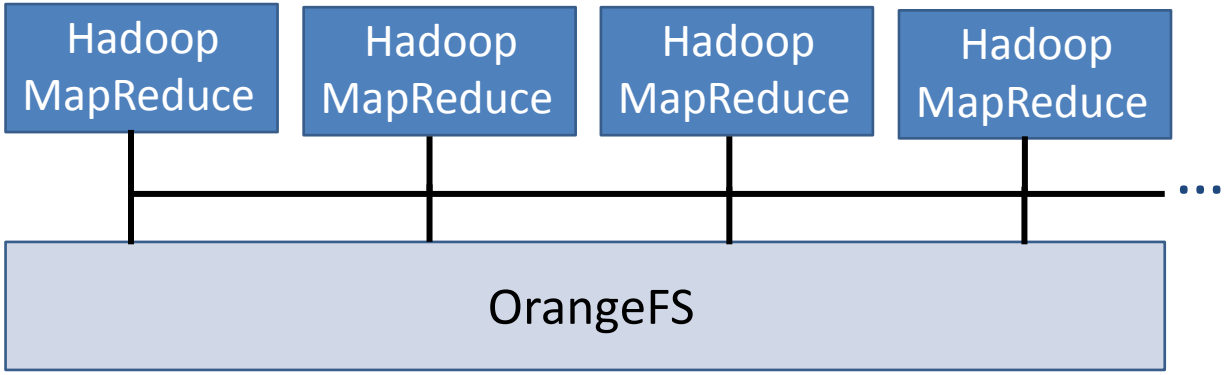
- To gain an insight of the two platforms, in this paper, we investigate the performance and resource utilization of different types of applications on the HPC-based Hadoop platforms with local storage and dedicated storage.
  - Hadoop with HDFS
  - Hadoop with OFS

[1] Y. Chen, A. Ganapathi, R. Griffith, and R. Katz. The Case for Evaluating MapReduce Performance Using Workload Suites. In Proc. of MASCOTS, 2011

Tranditional Hadoop with HDFS



Hadoop with OFS



Measurement Setting

- Clemson Palmetto HPC Cluster
- Hadoop Clusters
  - 40 machines
  - 8 cores
  - 16GB memory
  - 10Gbps Myrinet interconnect
- Hadoop 1.2.1, with the help of myHadoop
  - HDFS, local storage (HDD)
  - Remote file system (OrangeFS), a parallel file system
- Block sizes 128MB

OrangeFS

- OFS is an open source parallel file system, the next generation of Parallel Virtual File System (PVFS).
- The Palmetto HPC cluster at Clemson University has developed a Java Native Interface (JNI) shim to allow data to be passed between programs.
- The JNI shim allows Java code to execute functions present in the OrangeFS Direct Client Interface.
- We use 8 servers in total. Each OFS server has 5 HDDs to store data.
- Advantages over local file system
  - Easy to manage for a centralized storage, reliability and scale
  - More powerful than the local file system on HPC clusters

Measurement Application

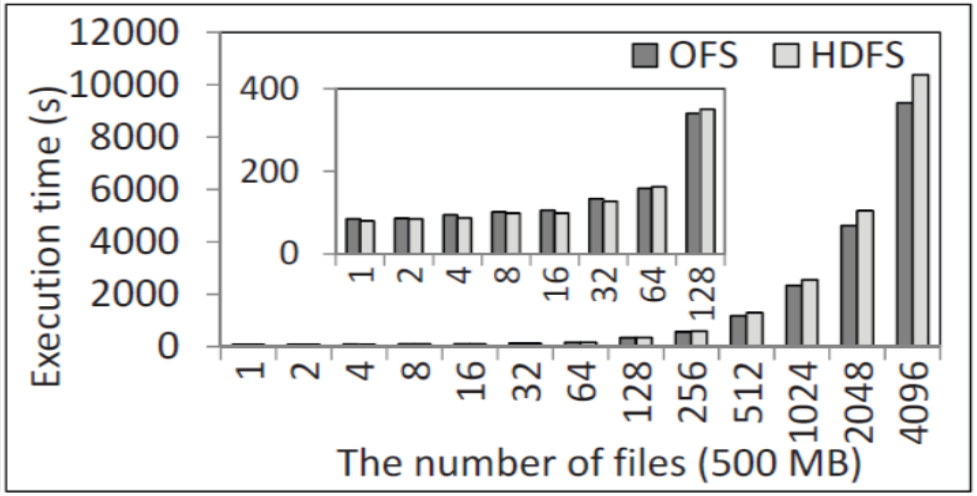
- Data-intensive application
  - A large amount of I/O read/write and a few amount of computation
  - WordCount, Grep
  - Input data generated from BigdataBench [1]
- I/O-intensive application
  - Purely consists of I/O read/write
  - Write and read test of TestDFSIO
- CPU-intensive application
  - A large amount of computation such as iterative computation
  - PiEstimator, PageRank

[1] L. Wang, J. Zhan, C. Luo, Y. Zhu, Q. Yang, Y. He, W. Gao, Z. Jia, Y. Shi, S. Zhang, et al. Bigdatabench: A big data benchmark suite from internet services. In Proc. of HPCA, 2014

Measurement Application

- Metrics
  - **Execution time**
  - Average map task execution time
  - Average reduce task execution time
  - CPU time
    - Use SYSSTAT utilities mpstat
  - Total transmitted data size
    - Developed a bash script to monitor the bandwidth consumption

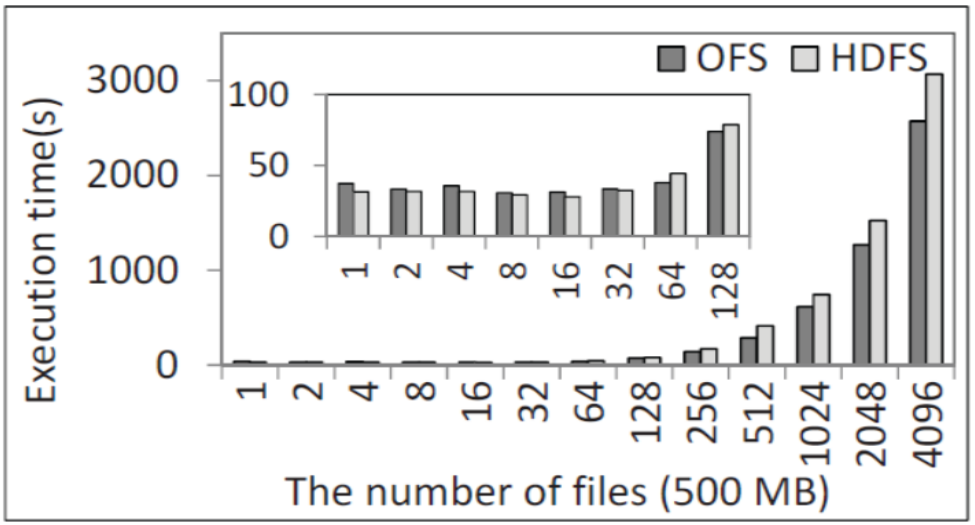
Measurement Analysis



Execution time of WordCount

Small input size:
OFS worse than HDFS

Large input size
OFS better than HDFS

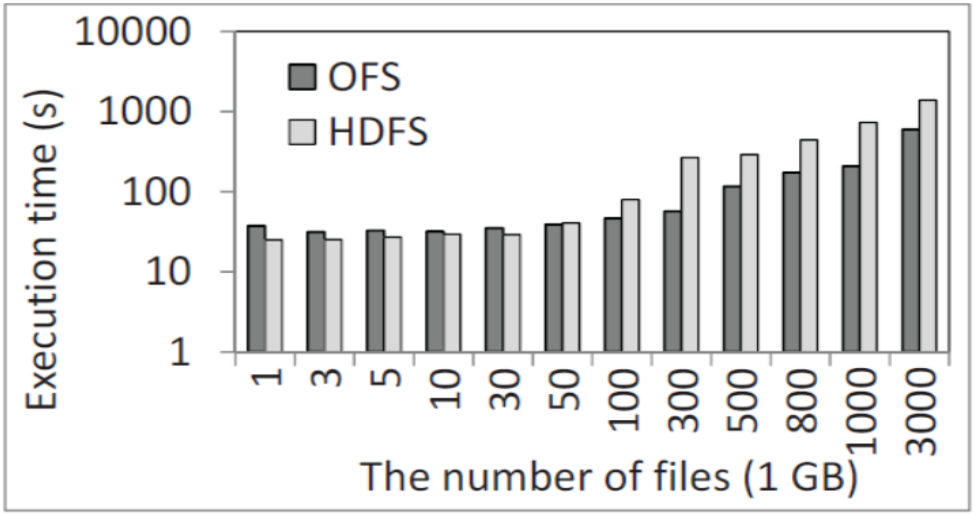


Execution time of Grep

Small input size:
OFS worse than HDFS

Large input size:
OFS better than HDFS

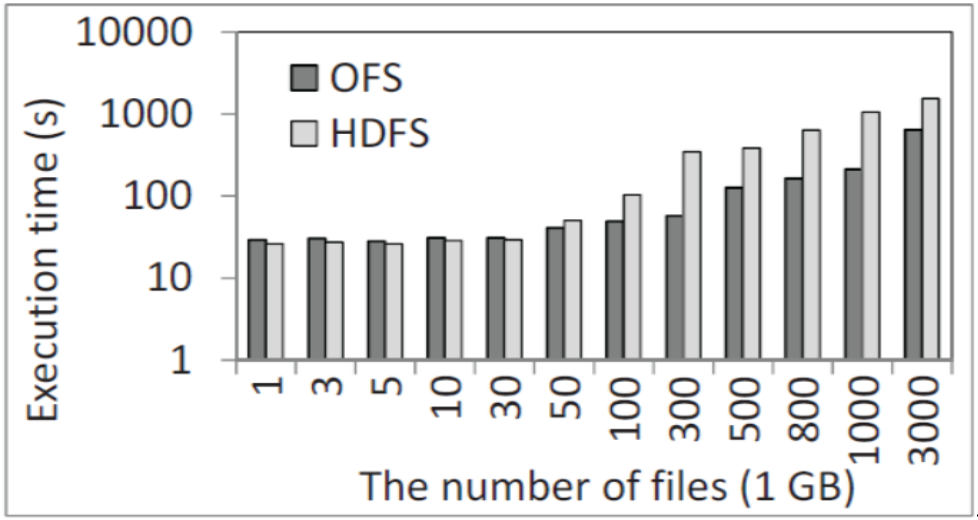
Measurement Analysis



Execution time of write test

Small input size:
OFS worse than HDFS

Large input size
OFS better than HDFS



Execution time of read test

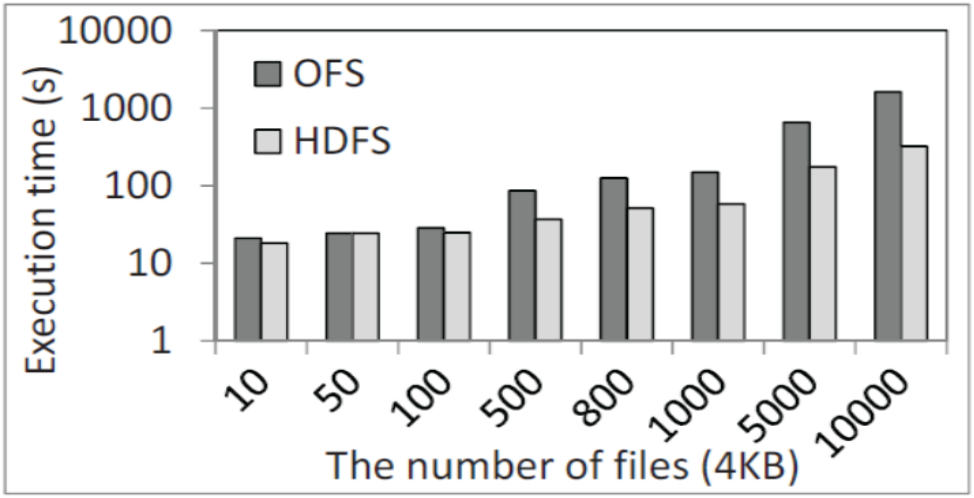
Small input size:
OFS worse than HDFS

Large input size:
OFS better than HDFS

Measurement Analysis

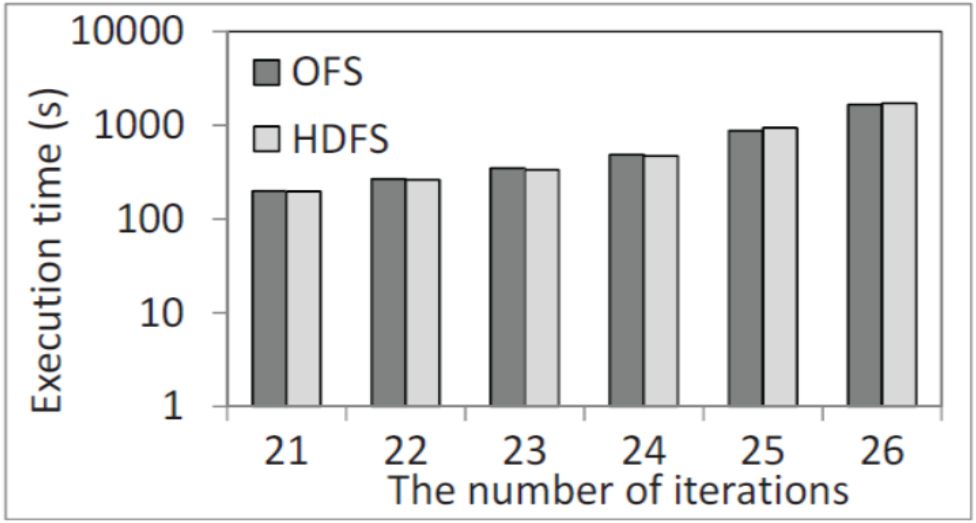
- If an application has a large input data size, OFS is a better platform
  - Better I/O performance
- If an application has a small input data size, HDFS is a better platform
  - Avoid network latency for small files
- The more computations an application has, the less influence from the I/O performance on the execution time and the less performance difference between the two platforms.
- Since the I/O-waiting CPU time occupies less percentage of total CPU time for data-intensive applications, the performance difference between OFS and HDFS for data-intensive applications is not as large as I/O-intensive applications
  - Concluded from CPU time metric, please refer to the paper for more details.

Measurement Analysis



Execution time of PiEstimator

OFS always worse than HDFS



Execution time of PageRank

Quite similar performance for OFS and HDFS

Measurement Analysis

- Although CPU-intensive applications can contain large size input files, a large amount of calculations dominate the CPU time for this kind of applications, which makes the I/O performance play a much less important role in determining the application performance.
- If CPU-intensive applications have a large number of small-size input files, HDFS is better platform that can avoid high user-level CPU time for communication setup with the remote storage in OFS.
- If CPU-intensive applications have large-size input files, both HDFS and OFS produce comparable performance.

Performance Evaluation

- Same cluster configurations
  - Hadoop with HDFS
  - Hadoop with OFS
  - 40 compute nodes
- Facebook-2009 synthesized trace
- Validate that the measurement results and show that Hadoop with OFS can provide better performance for some applications on HPC clusters

Discussion

- In the paper, we provide the measurement results in details.
  - analysis in details
  - provide reasons for the observations
- We expect that this gives a guidance to users on how to select the best platforms
  - selecting file systems
- Clouds, e.g., EC2
  - data is stored in a dedicated storage (e.g., Amazon S3)

Conclusion

- Conducted performance measurement study of data-intensive, I/O-intensive and CPU-intensive applications on HPC-based Hadoop platforms
  - Traditional Hadoop with HDFS
  - Hadoop with OFS
- Expect that our measurement results can help users to select the most appropriate platforms for different applications with different characteristics
- Future Work
  - Investigate Hadoop YARN on HPC clusters
  - Whether it is feasible to configure Hadoop with remote file system on Cloud Environment

UNIVERSITY of VIRGINIA

Thank you!

Questions & Comments?

Zhuozhao Li

zl5uq@Virginia.edu

Ph.D. Candidate

Pervasive Communication Laboratory

University of Virginia