Since the beginning of creation humans have valued data and time

Introducing iSCSI Protocol on Online Based MapReduce Mechanism

International Conference On Computational Science and Its Application, ICCSA-2014

Guimaraes, Portugal.

Presented By

Sung-Soon Park

Contents

- Introduction
- Motivation
- Proposed Schemes
- Experiments and Result Analysis
- Conclusion

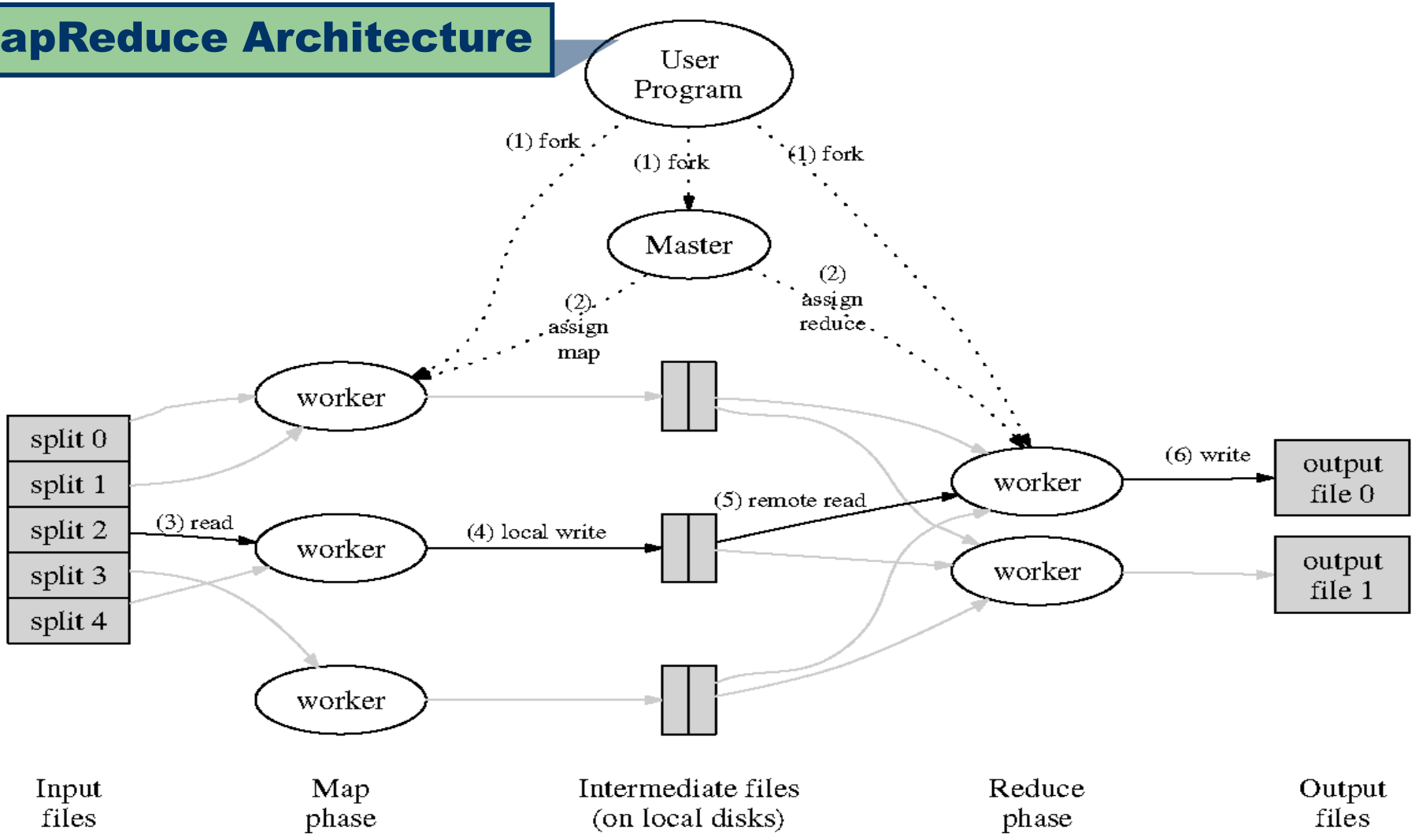
Introduction (1/4)

- Big data is a collection of huge amount of data sets. Data can be classified into three types:
  - **Structured data: processed by NoSQL, MongoDB, and TerraStore**
  - **Unstructured data**
  - **Semi-structured data**

- It will be better if ***MapReduce*** take little time to process the huge amount of data. Our goal is to optimize MapReduce.

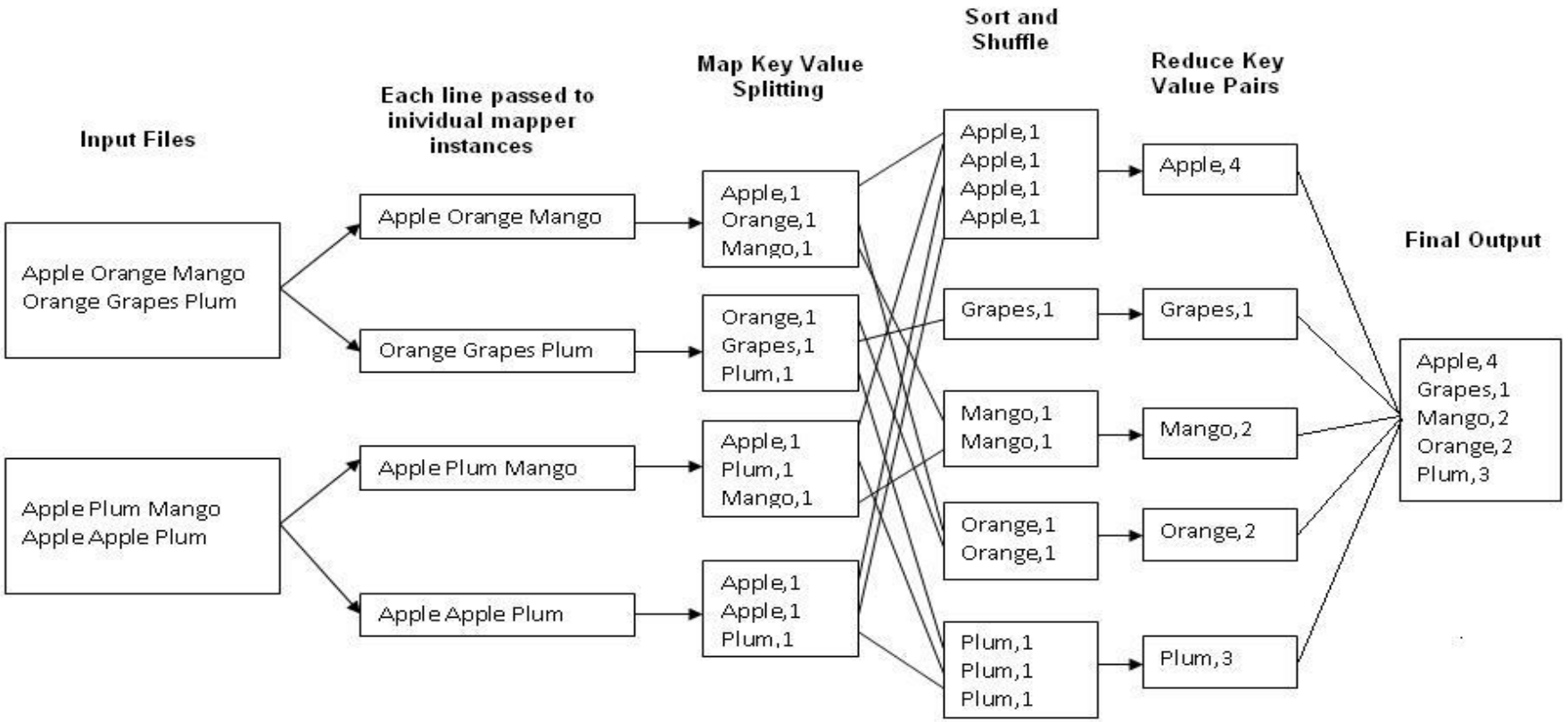
Introduction (2/4)

MapReduce Architecture



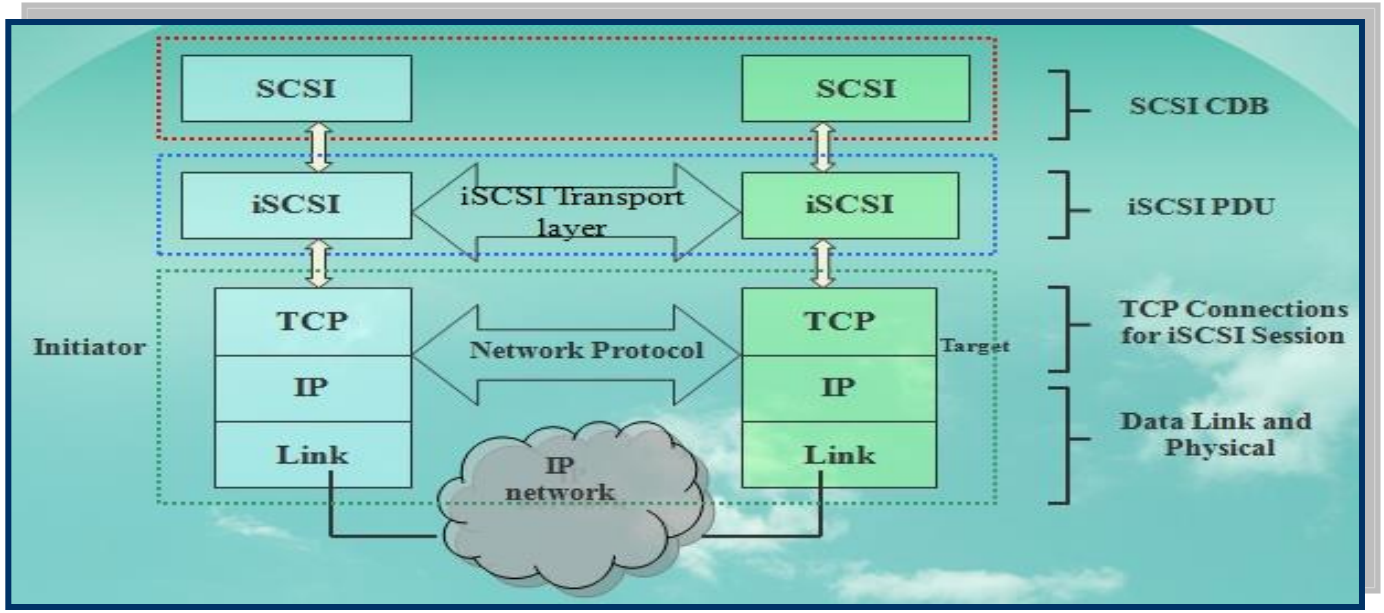
Introduction (3/4)

MapReduce Example



Introduction (4/4)

- Problems of MapReduce:
  - Reducer waits for full completion of Map task
  - Data overload and connection problem
  - Time of completion
- **Solution: Introducing iSCSI protocol**



Motivation

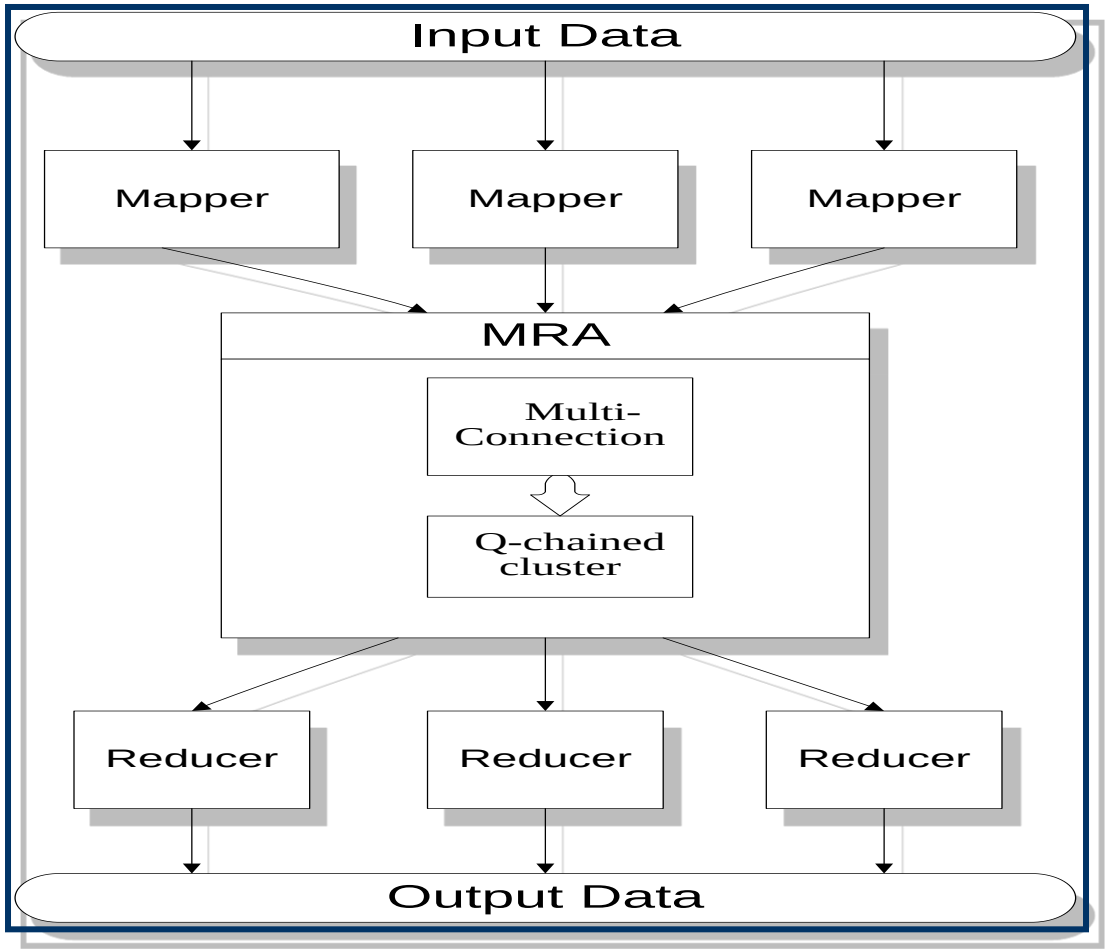
- In pipelining mechanism[**ref**], they used naïve implementation to send data directly from map to reduce tasks: ***MapReduce***
- Big Data transition Map to Reduce process by using TCP: o*ccur data overload and connection problem.*
- Then, Map task retransmit the whole data.
- It has no mechanism to handle data overload

- **MRA [MapReduce Agent]**
- Transmit data without retransmission
- Has a mechanism to handle data overload

[**ref: Tyson Condie, Neil Conway, Peter Alvaro, Joseph M. Hellerstein UC Berkeley: MapReduce Online. Khaled Elmeleegy, Russell Sears(Yahoo! Research)2012]**

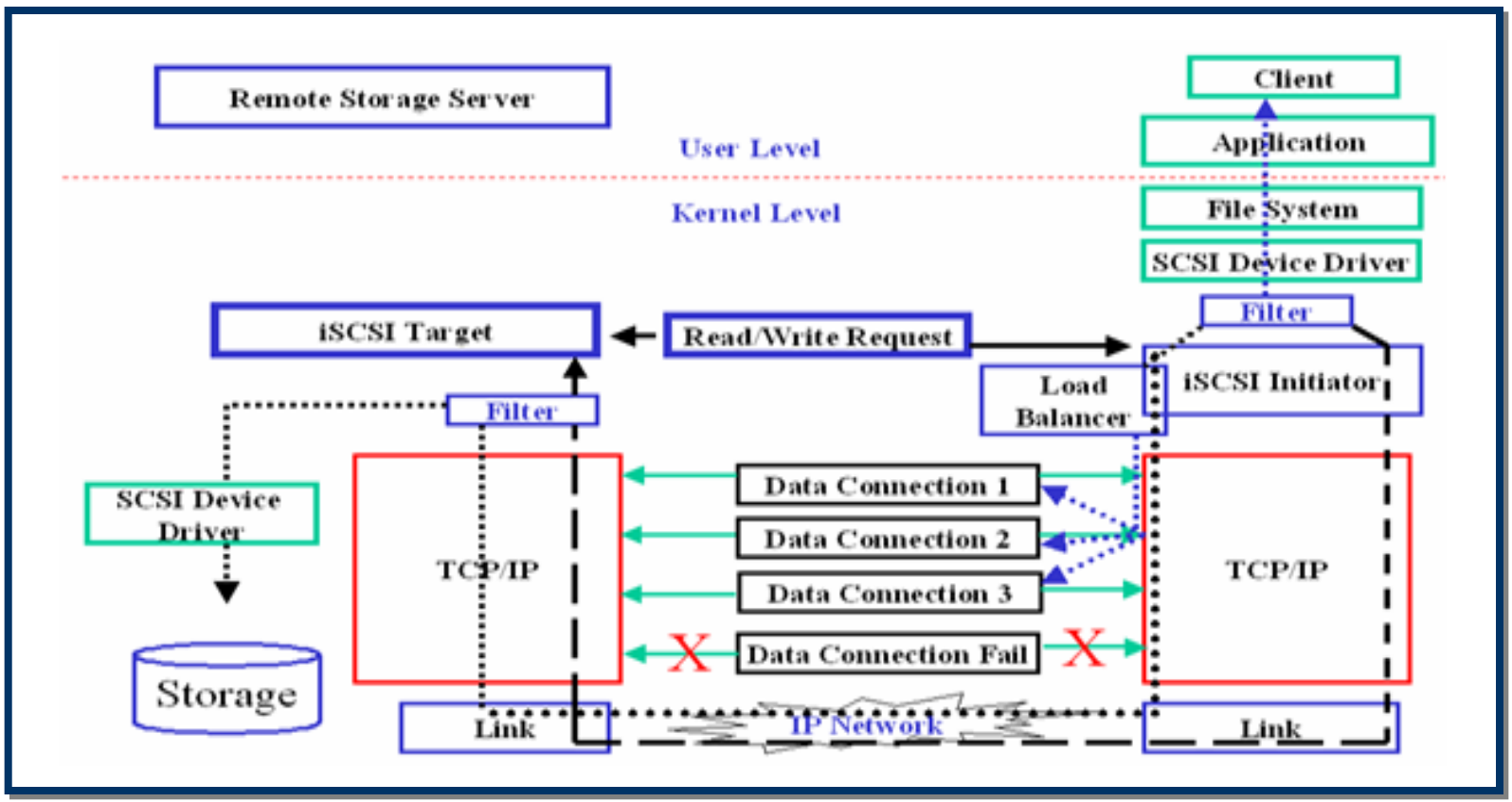
Proposed Model (1/3)

MapReduce Agent(MRA)



Proposed Model (2/3)

Overview of Multi-connection and ErrorRecovery Method of iSCSI



[ref: S.M.Allayear, Sung Soon Park: iSCSI Multi-Connection and Error Recovery Method for Remote Storage System in Mobile Appliance. The 2006 International Conference on Computational and It's Applications (ICCSA2006), Glasgow-Scotland. Springer-Verlag Berlin Heidelberg 2006, (SCI Indexed) LNCS 3981, pp.641-650.

Proposed Model (3/3)

Q-Chained Load Balancer:

Being able to fully balance, **the workload** among the data connections in the event of packet losses due to bad channel characteristics.

Data Connection	0	1	2	3	4	5
Primary data	Q_0	Q_1	Q_2	Q_3	Q_4	Q_5
Recovery data	-	-	-	-	-	-

< data transmission with multi TCP connections >

Data Connection	0	1	2	3	4	5
Primary data	Q_0	F	Q_2	Q_3	Q_4	Q_5
Recovery data	-	F		-	-	-

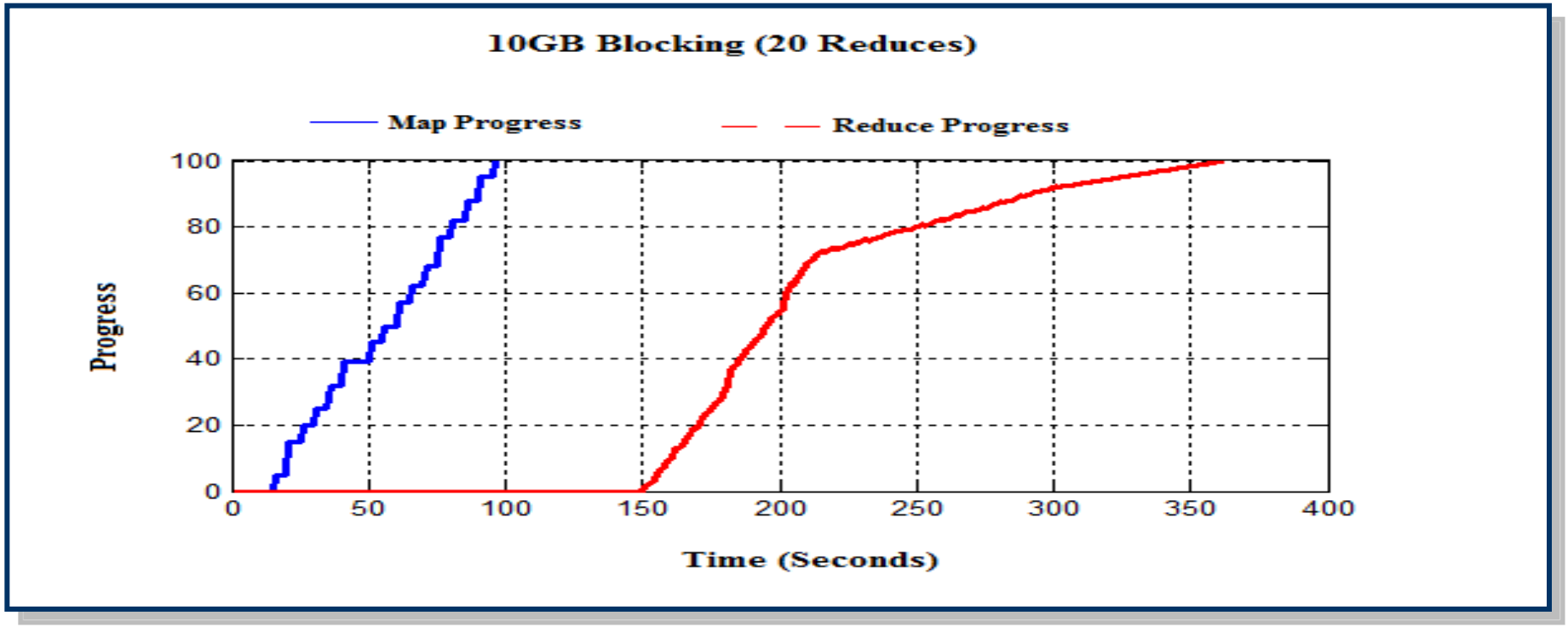
< typical takeover scheme without Load Balancing >

Data Connection	0	1	2	3	4	5
Primary data	Q_0	F	$\frac{1}{5}Q_2$	$\frac{2}{5}Q_3$	$\frac{3}{5}Q_4$	$\frac{4}{5}Q_5$
Recovery data	$\frac{1}{5}q_5$	F	q_1	$\frac{4}{5}q_2$	$\frac{3}{5}q_3$	$\frac{2}{5}q_4$

< Q-Chained Cluster Load Balancer >

Performance Evaluation (1/6)

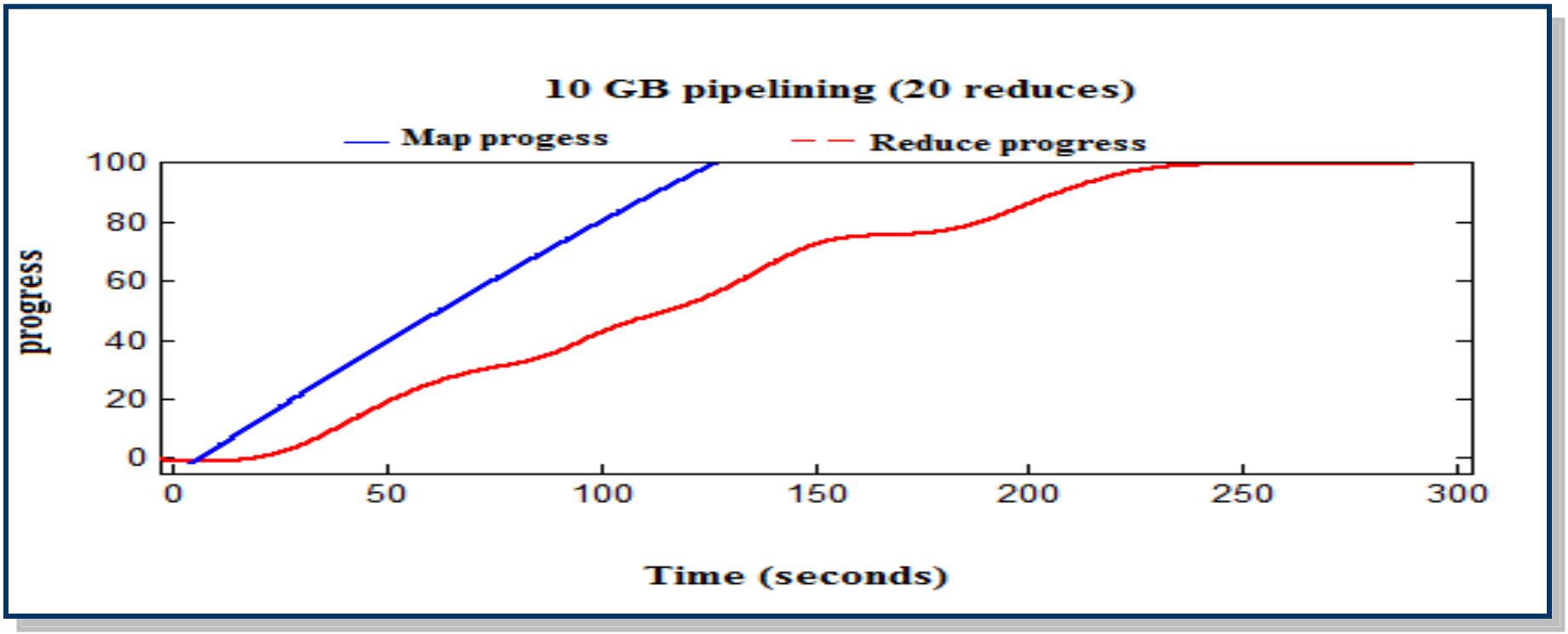
10 GB Blocking



- CDF of map and reduce task completion times for **a 10GB wordcount job** using **20 map tasks** and **20 reduce tasks** (**512MB block size**).
- The total job runtimes were **361 seconds** for blocking.

Performance Evaluation (2/6)

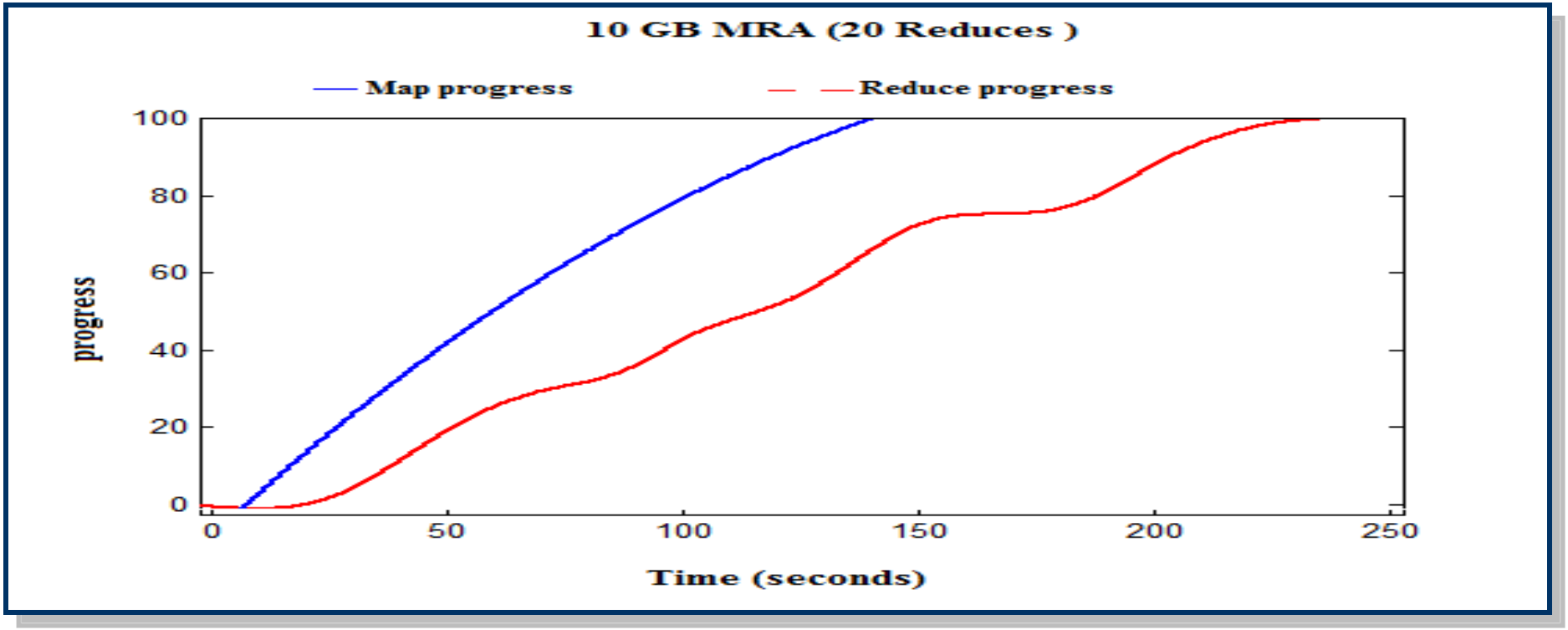
10 GB pipelining



- CDF of map and reduce task completion times for a 10GB wordcount job using 20 map tasks and 20 reduce tasks (512MB block size).
- The total job runtimes were **290 seconds** for pipelining.

Performance Evaluation (3/6)

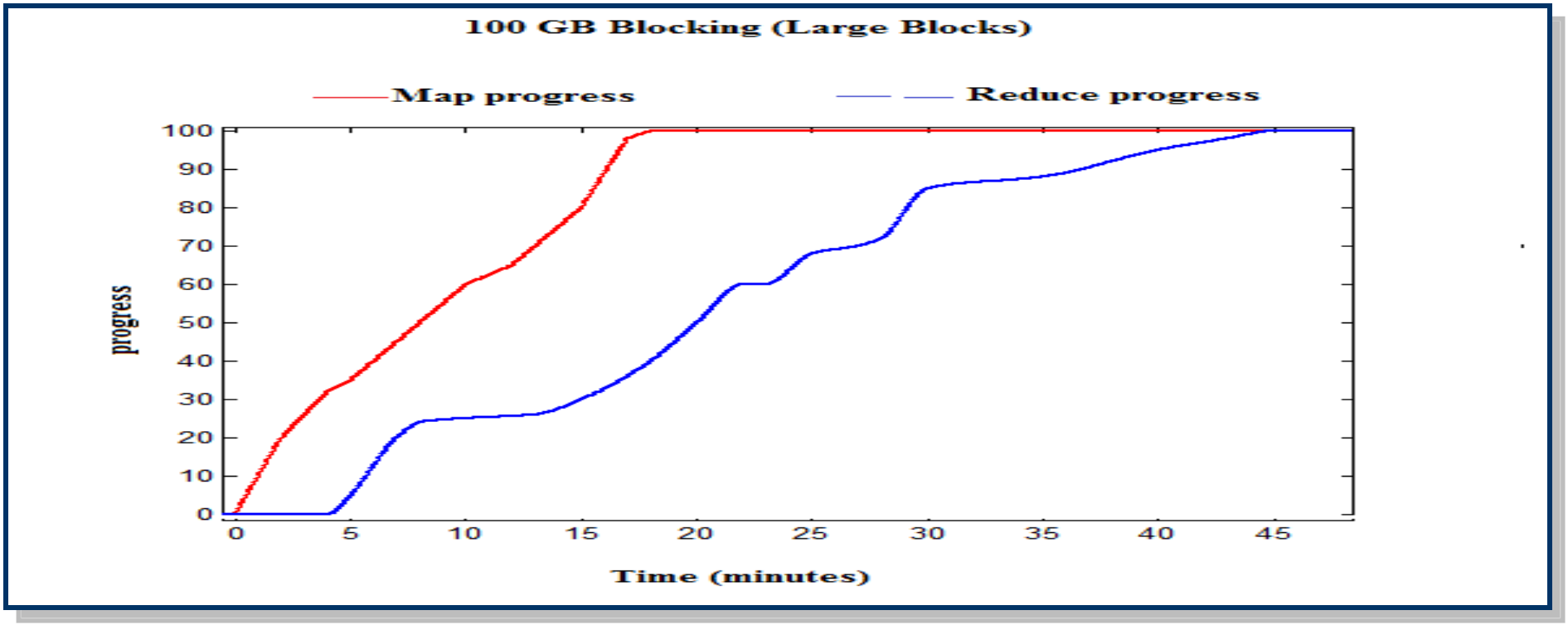
10 GB MRA



- CDF of map and reduce task completion times for a 10GB wordcount job using 20 map tasks and 20 reduce tasks (512MB block size).
- The total job runtimes were **240 seconds** for MRA.

Performance Evaluation (4/6)

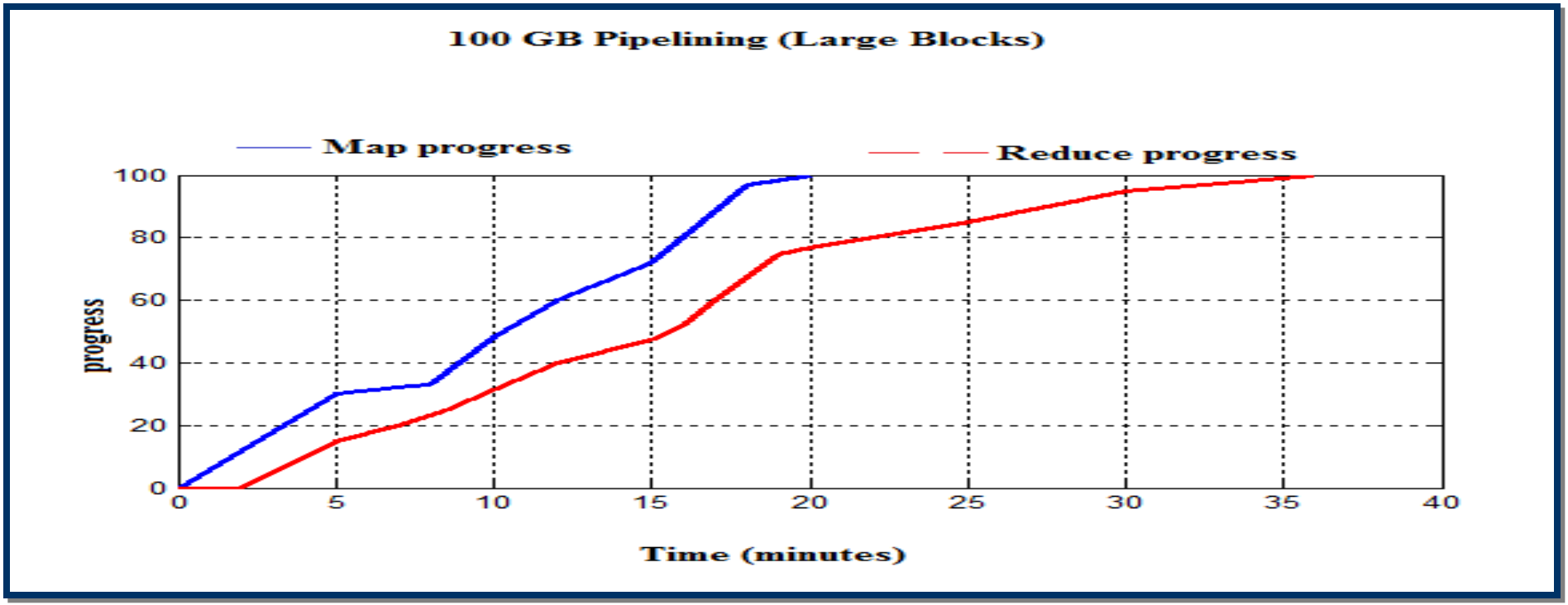
100 GB Blocking



- CDF of map and reduce task completion times for **a 100GB wordcount job** using **240 map tasks** and **60 reduce tasks** (**512MB block size**).
- The total job runtimes were **48 minutes** for blocking.

Performance Evaluation (5/6)

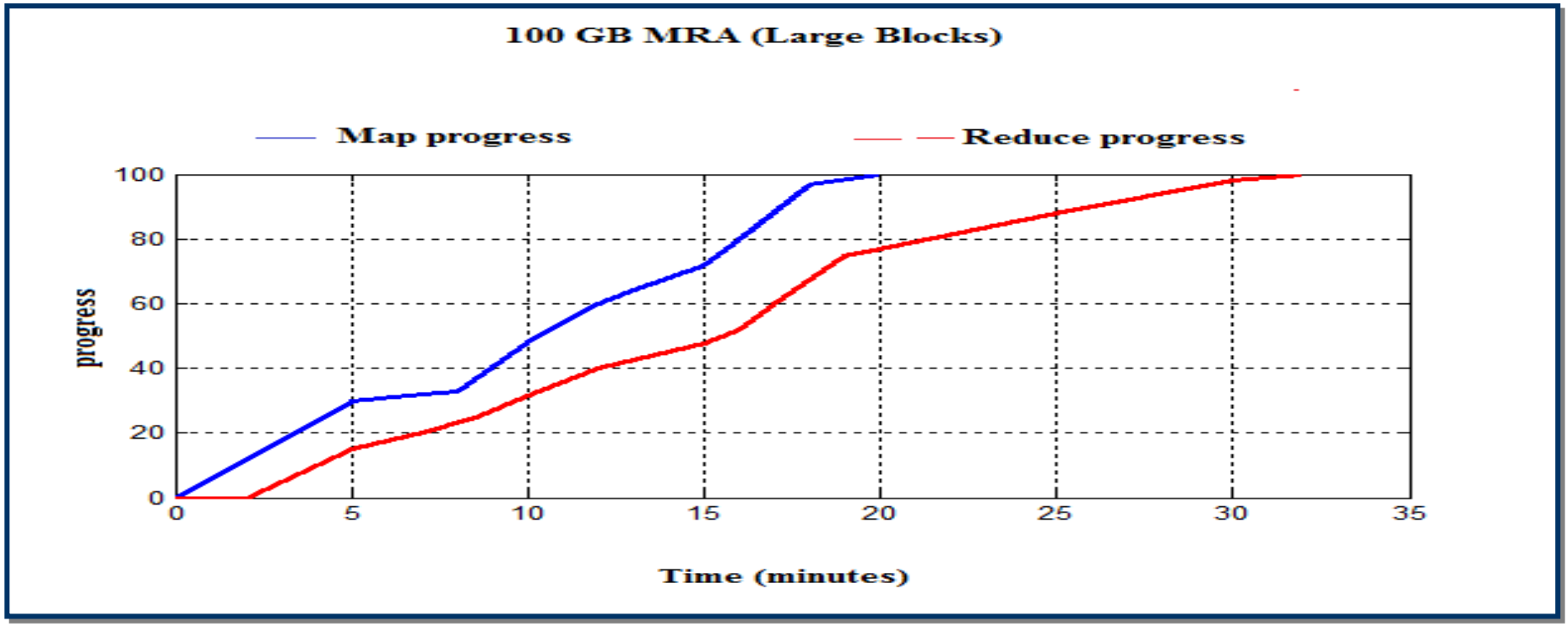
100 GB pipelining



- CDF of map and reduce task completion times for a 100GB wordcount job using 240 map tasks and 60 reduce tasks (512MB block size).
- The total job runtimes were **36 minutes** for pipelining.

Performance Evaluation (6/6)

100 GB MRA



- CDF of map and reduce task completion times for a 100GB wordcount job using 240 map tasks and 60 reduce tasks (512MB block size).
- The total job runtimes were **32 minutes** for MRA.

Conclusion

- In these circumstances, our proposed mechanism (MRA) may prove its better time completion for Hadoop technique.
- Now a days Hadoop technique is used in facebook, yahoo, cloud computing.
- So, our research can play an important role in the distributed file system and network storage area.

Reference

- DEAN, J., AND GHEMAWAT, S. MapReduce: Simplified dataprocessing on large clusters. In OSDI (2004).
- SAM-3 Information Technology – SCSI Architecture Model 3, Working Draft, T10 Project 1561-D, Revision7, 2003.
- S.M.Allayear, Sung Soon Park: iSCSI Multi-Connection and Error Recovery Method for Remote Storage System in Mobile Appliance. The 2006 International Conference on Computational and It's Applications (ICCSA2006), Glasgow- Scotland. Springer-Verlag Berlin Heidelberg 2006, (SCI Indexed) LNCS 3981, pp.641-650.
- Tyson Condie, Neil Conway, Peter Alvaro, Joseph M. Hellerstein UC Berkeley: MapReduce Online. Khaled Elmeleegy, Russell Sears(Yahoo! Research)
- HELLERSTEIN, J. M., HAAS, P. J., AND WANG, H. J. Online aggregation. In SIGMOD (1997).