ROYAL UNIVERSITY OF PHNOM PENH

MASTER OF IT ENGINEERING (MITE)

Subject: Distributed System

Assignment 3: Scenario 1

Lecturer Name: Taing Nguonly

Student Name: **SOK Pongsametrey**
THOU Bunhann
THY Poty
MOUNG Chip Eng

Year: 2011-2012

Outline

I. Introduction
II. Objectives
III. Definition
IV. Methods
V. Lab Specifications
VI. Lab Experiment
VII. Result
VIII. Conclusion

References

Introduction

- Distributed systems are everywhere. The Internet enables users throughout the world to access its services wherever they may be located.

- Resource sharing is the main motivating factor for constructing distributed systems.

Objectives

- Find out the frequency of word.
- Understand multithread, single processor and multiprocessor.
- Evaluation of applying distributed algorithm in a single computer and multiple computers.

Definition

- A distributed system is one in which components located at networked computers communicate and coordinate their actions only by passing messages" - Coulouris, et. al.

- Concurrency is a property of systems in which several computations are executing simultaneously, and potentially interacting with each other.

Definition (Cont.)

- A process is a compiled program that has been loaded into memory, whereupon the CPU may execute its instructions.

- A thread is defined as a path of execution, a collection of statements that execute in a specific order.

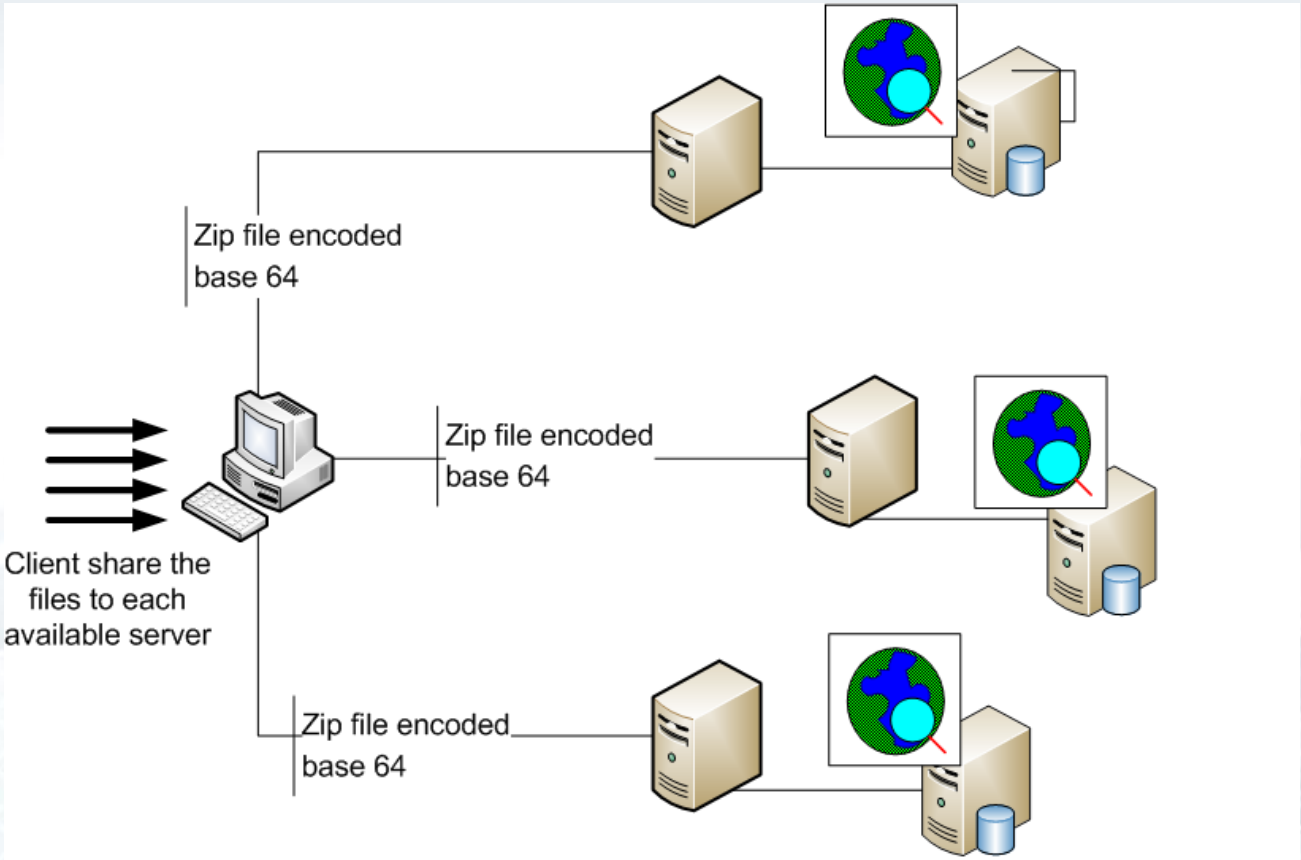
Lab Specifications

- Single PC
  - CPU: Intel Core 2 Duo CPU P8700 @2.53GHz
  - Memory: 3 GB
  - OS: Windows XP Pro SP 2
- Multiple PC
  - CPU: Intel Core 2 Duo CPU P8700 @2.53GHz
  - Memory: 3 GB
  - OS: Windows XP Pro SP 3, Widows 7 Ultimate
  - LAN: YES
- JAVA JDK: 1.6
- JAVA IDE: Eclipse 3.6 JEE
- Apache ANT 1.6 or latest

Methods

- Client – Server: RMI
- File to sent, zip and encoded with Base 64
- Multi-thread / Thread
- Use DB: SQLite for future search
- Client prepare the files to send to each server
- After file sent, it will process mapreducing
- To find word, it will look up in each server’s db

Process Diagram



Serve decode to file and do MapReduce
Save result to SQLite DB

User can search word, it will look up in each db of the server and display the result

Lab Experiment

- Single PC (Multi-thread)

Single PC							
	MS Win XP SP 2, Core 2 Duo CPU P8700 @2.53GHz RAM 3GB						
File (MB)	MapReduce	Sample Search Word					
		from		that		can	
	Time (sec.)	Frequency	Time (sec)	Frequency	Time (sec)	Frequency	Time (sec)
1	17.10994945	1997	1.820162619	3994	6.029178569	3994	0.7826539
4	50.83090169	7988	1.33015509	15976	1.157977341	15976	0.9101889
8	144.514626	15976	1.594184177	31952	1.052706699	31952	1.1618748
16	180.6588022	31952	1.628752817	63904	0.938892792	63904	0.8522935
32	367.941946	63904	1.646371815	127808	2.965521875	127808	2.6721368

Lab Experiment (Cont.)

- Multiple PC (Multi-Thread)

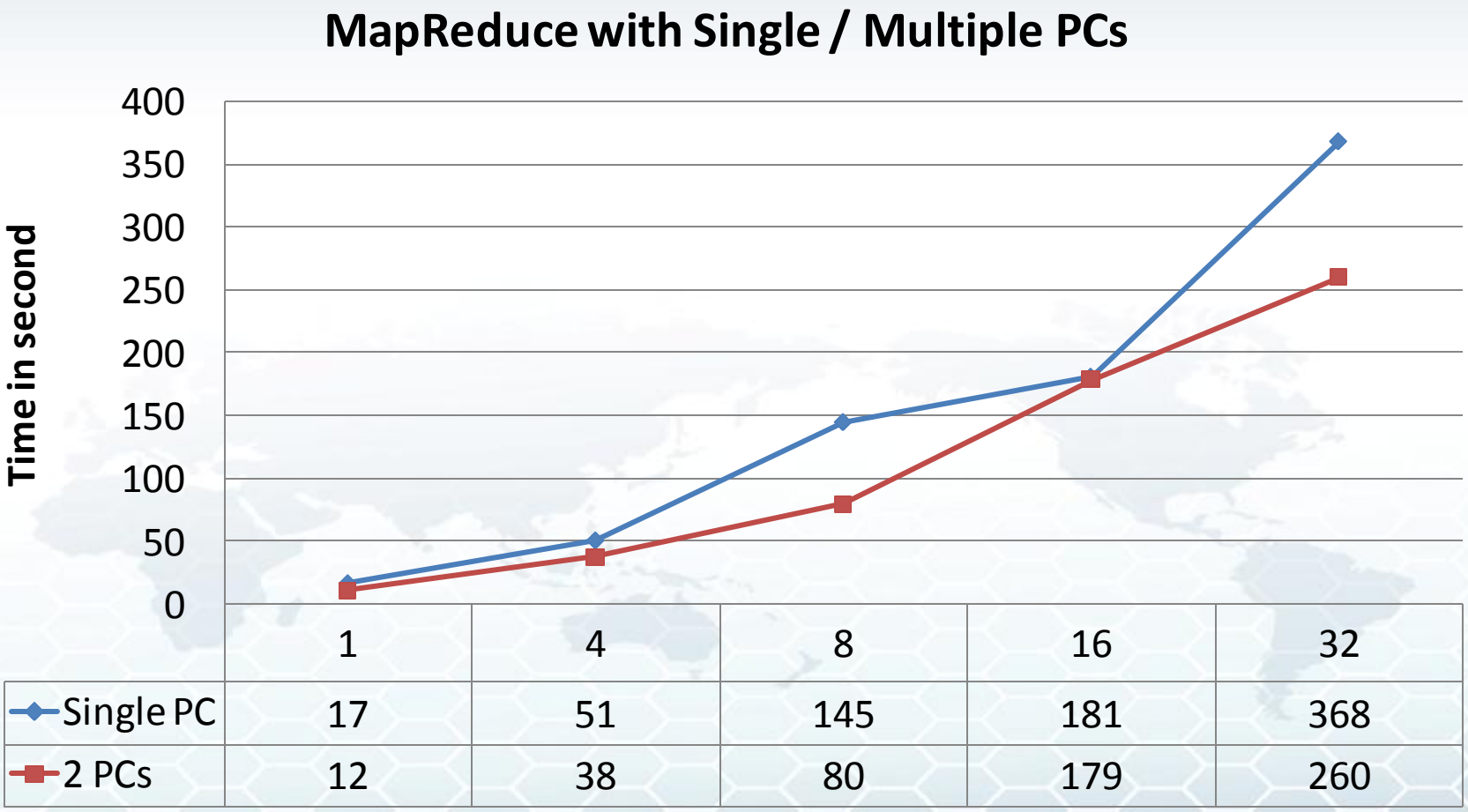
Two PCs

MS Win 7 Utimate, Core 2 Due CPU P8700 @2.53GHz

RAM 3GB

File (MB)	MapReduce	Sample Search Word					
		from		that		can	
	Time (sec.)	**Frequency**	**Time (sec)**	**Frequency**	**Time (sec)**	**Frequency**	**Time (sec)**
1	11.98528053	1997	1.141537262	3994	0.822694505	3994	1.592889
4	37.79924544	7988	1.271142003	15976	2.741015968	15976	2.6712605
8	80.13109152	15976	1.096724305	31952	2.753692439	31952	0.8804127
16	178.8897533	31952	1.230367547	63904	2.318571495	63904	0.8321178
32	259.4999169	63904	1.54650605	127808	1.348409923	127808	2.0576816

Result



	1	4	8	16	32
Single PC	17	51	145	181	368
2 PCs	12	38	80	179	260

Conclusion

- Single PC process the small size file faster or similar to Multiple PCs.
- Multiple PC performance better than Single PC when the file size is large.
- When the file is larger and larger, single process will spend much time in process than multiple process.

References

- DISTRIBUTED SYSTEMS Concepts and Design 5th edition, (Coulouris, et al. 2012)

- Java in 60 Minutes a Day, 1st edition, (Richard F. Raposa, 2003)

- DISTRIBUTED SYSTEMS, Concurrency (Taing Nguonly, 2012)

សូមអរគុណ !!