

HIGH PERFORMANCE WEB SERVERS USING LOAD BALANCING TECHNIQUE

SITI AISHAH BINTI AB RAHMAN

**BACHELOR OF SCIENCE (Hons.) IN
DATA COMMUNICATION AND NETWORKING
FACULTY OF INFORMATION TECHNOLOGY & QUANTITATIVE
SCIENCES
MARA UNIVERSITY OF TECHNOLOGY
SHAH ALAM**

MAY 2006

**HIGH PERFORMANCE WEB SERVERS USING LOAD
BALANCING TECHNIQUE**

By

SITI AISHAH BINTI AB RAHMAN

(2003470986)

A project paper submitted to

**FACULTY OF INFORMATION TECHNOLOGY & QUANTITATIVE
SCIENCES
MARA UNIVERSITY OF TECHNOLOGY**

In partial fulfillment of requirement for the

**BACHELOR OF SCIENCE (Hons.) IN
DATA COMMUNICATION AND NETWORKING**

Major Area: Network/Communication

Approved by the Examining Committee:

.....
En Mohamad Yusof Darus
Supervisor

Project

.....
Pn Rozita Yunos

Examiner

**UNIVERSITI TEKNOLOGI MARA
SHAH ALAM, SELANGOR
MAY 2006**

CERTIFICATE OF ORIGINALITY

This is to certify that I am responsible for the work submitted in this project that the original work is my own except a specified in the references and acknowledgment and the original work contained here have not been taken or done by unspecified sources or persons.

.....
(SITI AISHAH BINTI AB RAHMAN)

ABSTRACT

A World Wide Web (WWW) Server is normally a single machine dedicated to process a HTTP request for a single WWW site. Web server serves web pages to clients across the Internet or an Intranet. The web server hosts the pages, scripts, programs, and multimedia files and serves them using HTTP, a protocol designed to send files to web browsers and other protocols. In order to achieve web server scalability, more servers need to be added to distribute the load among the group of servers, which is also known as a server cluster. The load distribution among these servers is known as load balancing. Load balancing applies to all types of servers, the application server and database server. However, this project will be devoting this section for load balancing of web servers (HTTP server) only. Load balancing is a technique that distributes processing and communications activity evenly across a computer network so that no single device is overwhelmed. In other words, when multiple web servers are present in a server group, the HTTP traffic needs to be evenly distributed among the servers. The purpose of load balancing was done due to the increases of traffic, complexity of the application software and to satisfy the critical online transaction nowadays. The overall exact nature of this project is a setup of load balancer, a cluster of real servers and clients. The architecture of the cluster is transparent to clients outside the cluster. Client applications interact with the cluster as if it were a single high-performance and high available server. Clients will not be affected by interaction with the cluster and do not need modification. The load balancing method used in this project is Round Robin Scheduling that distributes job or request equally among the web servers. The Round Robin method that will be use in this project is dispatcher-based load balancing cluster where the parallel services of servers can appear as a virtual service on a single IP address. For routing, Network Address Translation is a technique in which the source or destination addresses of IP packets are rewritten as they pass through a router or firewall.

TABLE OF CONTENT

TITLE	PAGE
CERTIFICATE OF ORIGINALITY	ii
ACKNOWLEDGEMENT	iii
ABSTRACT	iv
TABLE OF CONTENT	v
LIST OF TABLES	viii
LIST OF FIGURES	ix
LIST OF ABBREVIATION	x
CHAPTER	
1 INTRODUCTION	
1.1 INTRODUCTION	1
1.2 PROBLEM STATEMENTS	4
1.3 PROJECT OBJECTIVES	5
1.4 PROJECT SCOPE	5
1.5 PROJECT SIGNIFICANCE	5
2 LITERATURE REVIEW	
2.1 INTRODUCTION	7
2.2 DEFINITION OF PERTINENT TECHNICAL TERMINOLOGIES	
2.2.1 NETWORK LOAD BALANCING	7
2.2.2 HTTP PROTOCOL	9
2.2.3 WEB SERVER	9
2.2.3.1 APACHE WEB SERVER	10
2.2.3.2 WINDOWS IIS WEB SERVER	11
2.2.4 SCHEDULING ALGORITHM	11
2.2.4.1 ROUND ROBIN SCHEDULING	12
2.2.4.2 WEIGHTED ROUND ROBIN SCHEDULING	12