Performance Report

IBM Netfinity 5500 (550MHz)

Version 2.0
June 1999

IBM ®

Executive Overview

The IBM Netfinity 5500 servers offer new levels of power, scalability and serviceability for creating a reliable foundation for networked, business-critical applications. New models announced worldwide in June 1999 feature single or dual symmetrical multiprocessing (SMP) with the new 500MHz[1] Pentium** III processor, which supports 100MHz operations to memory.

The new 550MHz system (Model 8660-72U) was evaluated using Ziff-Davis' benchmark ServerBench** Version 4.02.

For comparison, the IBM Netfinity server performance laboratory also conducted the benchmark with the IBM Netfinity 5500 configured with a 500MHz Pentium III processor.

All results obtained from this benchmark are presented in this report.

Performance Highlights

Following are highlights of the benchmark results. Please review the more detailed information concerning the results later in this report.

ServerBench 4.02

ServerBench 4.02 was used to measure the performance of the IBM Netfinity 5500 systems as dual-processor application servers running Windows NT Server 4.0 and providing services to Windows NT Workstation 4.0 clients.

The IBM Netfinity 5500 550MHz system achieved a peak level of transactions per second that was **slightly higher** than that of the 500MHz system in both RAID-0 and RAID-5 configurations.

# Test Environments and Results

ServerBench 4.02

The ServerBench 4.02 system test suite SYS_60.TST was used to measure the performance of the IBM Netfinity 5500 550MHz and 500MHz systems, configured as two-way application servers running Windows NT Server 4.0.

ServerBench 4.02 provides an overall transactions-per-second (TPS) score showing how well the server handles client requests for a variety of operations involving the server's processors, disk and network subsystems.

Results Summary

RAID-0 Configuration

The IBM Netfinity 5500 550MHz system achieved a peak level of transactions per second at four clients that was **slightly higher** than that of the 500MHz system.

RAID-5 Configuration

The IBM Netfinity 5500 550MHz system achieved a peak level of transactions per second at four clients that was **slightly higher** than that of the 500MHz system.

ServerBench Version 4.02
RAID-5 Configuration - Windows NT Server 4.0
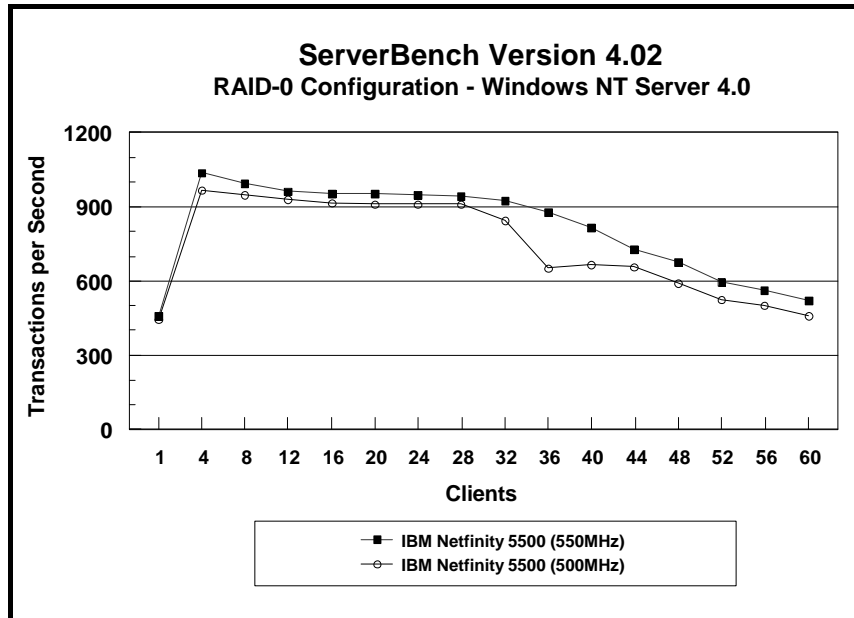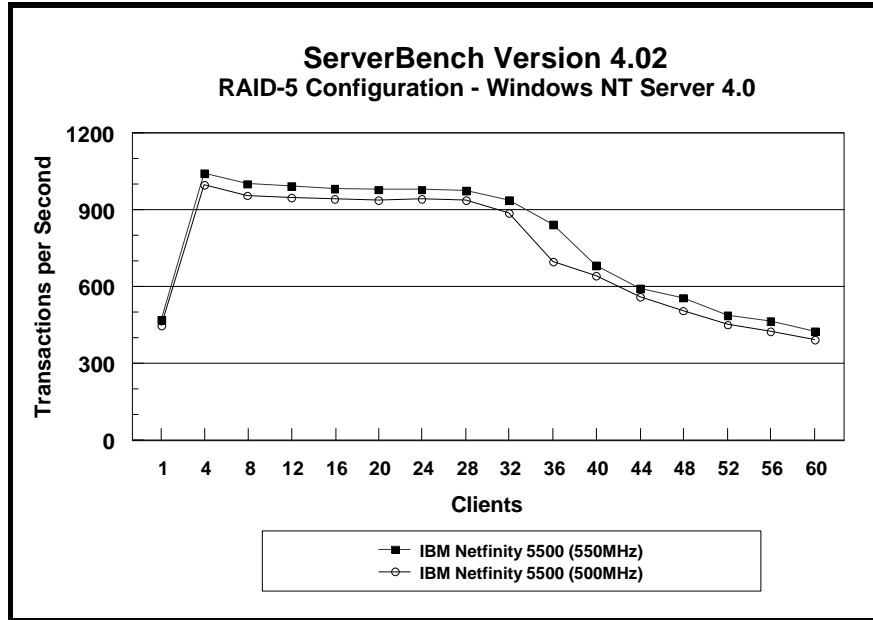
Measurement Methodology

The system test suite was performed using four 100Mbps Ethernet network segments with a total of 60 IBM PC 750 166MHz systems as client workstations attached to the server. Each workstation ran Windows NT 4.0 Workstation and executed the ServerBench 4.02 SYS_60.TST workload, which includes the client/server, processor, server/client, random read, and random write requests typically made in a client/server computing environment. (The default values were used for all NT registry variables. The NT default is 'Max throughput for file sharing'.)

A transaction is a request issued by any one of the 60 clients; the TPS score is the number of transactions per second completed by the server under test. In the ServerBench environment, the server will not service the next request until it has finished the previous one. A higher TPS indicates better performance.

The clients randomly send requests to the server, producing different types of loads on the server. The server performs the work by disk caching if system memory is available, or swapping mapped memory out to paged files if system memory is full.

The SYS_60.TST test suite contains a total of 16 test mixes. Measurements of transactions per second (TPS) were recorded as a weighted harmonic mean of the total TPS obtained by all clients in each test mix as clients were added. Clients were added in increments of four as follows: 1, 4, 8, 12, 16, 20, 24, 28, 32, 36, 40, 44, 48, 52, 56, 60.

Measurement Analysis

ServerBench's server application on Windows NT provides up to 47 service threads with 60 clients, plus one thread for each server processor. For this test, the servers were configured with two processors; therefore, a total of 49 service threads were used. A client workstation generates a request for the server to begin the next phase of a mix or to ask the server to perform some operation. The server creates a new service thread and passes that connection with the client to an I/O completion port.

As clients are added to the network, the I/O workload increases, requiring more service threads to be allocated to the clients. When all the service threads have been allocated, any new client requests cannot be serviced until an I/O completion port becomes available. Using four 100Mbps network adapters provided sufficient bandwidth to the application server.

ServerBench requires a large amount of system memory to produce a meaningful result. When workload increases gradually, the processor subsystem (processor and system memory) provides adequate service to all requests by caching them in the system memory, which is the primary factor affecting the TPS throughput.

As workload continued to increase (i.e., more clients joined the test mixes), system memory was exhausted, and the server had to rely on the disk subsystem for virtual memory. When this happened, the bottleneck shifted to the disk subsystem, and the application became disk-bound. Running ServerBench with Windows NT may result in a low cache-hit ratio because some NT system threads (e.g., cache manager's lazy writer thread, memory manager's mapped page writer thread) will automatically move some mapped memory into paged files. If a client happens to request that paged-out data again, a cache-hit-miss will result.

The exact number of clients required to move the bottleneck from the processor to the disk subsystem depends on the amount of installed system memory. In our measurements, the application was processor-bound when running from 4 to 8 clients; with more than 32 clients, the application became disk-bound.

Server Configurations

ServerBench 4.02

Features	IBM Netfinity 5500 550MHz/512KB	IBM Netfinity 5500 500MHz/512KB
Processor	Two 550MHz Pentium III	Two 500MHz Pentium III
Memory	512MB ECC Synchronous DRAM	512MB ECC Synchronous DRAM
L2 Cache	512KB (Write-Back)	512KB (Write-Back)
RAID Level	0 and 5	0 and 5
Disk Drive	Five 9.1GB 10K Wide Ultra SCSI Drives	Five 9.1GB 10K Wide Ultra SCSI Drives
Disk Drive Adapter	ServeRAID II Ultra SCSI PCI Bus on Planar	ServeRAID II Ultra SCSI PCI Bus on Planar
Disk Driver	IPSRAIDN.SYS	IPSRAIDN.SYS
Network Adapter	Four IBM Ethernet 10/100 PCI Adapters	Four IBM Ethernet 10/100 PCI Adapters
Bus	PCI	PCI
Network Driver	IBMFENT.SYS	IBMFENT.SYS
Network Operating System	Windows NT Server 4.0 with Service Pack 3	Windows NT Server 4.0 with Service Pack 3
System Partition Size	1GB	1GB
File System	NTFS	NTFS
Allocation Unit Size	Predefined Default	Predefined Default
ServerBench Version / Test Suite	ServerBench 4.02 / SYS_60.TST	ServerBench 4.02 / SYS_60.TST

Test Disclosure Information

ServerBench 4.02

The ServerBench measurements were conducted using Ziff-Davis' ServerBench 4.02 running the SYS_60.TST test suite with Windows NT Workstation 4.0 as described below:

Version: ServerBench 4.02

Mixes

- System Test Mixes
- Clients: 1, 4, 8, 12, 16, 20, 24, 28, 32, 36, 40, 44, 48, 52, 56, 60
- Data Segment Size: 16MB
- Segment Access Ratio: 1
- Ramp up: Default setup
- Ramp down: Default setup
- Delay: 0
- Think: 0

Network Operating System: Windows NT Server 4.0 with Service Pack 3

Testbed Disclosure

The IBM Netfinity 5500 550MHz is planned to be available June 30, 1999. All other products are shipping versions available to the general public. All measurements were performed without independent verification by Ziff-Davis.

Network	100Mbps Ethernet
Clients	60
Hubs	Asante 100Mbps Ethernet
Clients per Segment	15
CPU / Memory	166MHz Pentium / 32MB
Network Adapter	IBM 100/10 PCI Ethernet Adapter (Bus 0)
Software	Windows NT 4.0 Workstation
Cache	L2 = 512KB
Controller Software	Microsoft Windows NT Workstation 4.0

Notes

[1] MHz denotes the clock speed of the microprocessor only, not application performance. Many factors affect application performance.

[2] When referring to hard disk capacity, GB, or gigabyte, means one thousand million bytes. Total user-accessible capacity may vary depending on operating environment.