Exchange 5.5 MAPI Messaging Benchmark (MMB) Performance Result

Date of Results: September 29, 1999

Executive Summary

IBM's Netfinity® 3500 M10 server has demonstrated exceptional performance on a low-cost, compact server. Microsoft's LoadSim tool was used to simulate 9,800 Exchange (MAPI) e-mail users running on an IBM Netfinity 3500 M10 server configured with two 600MHz Intel® Pentium® III processor and 1024MB of memory. The canonical "Medium User" profile was used, which represents the tasks typically performed by a corporate e-mail user. During the test period, the Netfinity 3500 M10 provided a weighted 95th-percentile response time of 468ms for 9,800 MMB, with overall CPU utilization of 87%.

Results are based on four hours of steady-state running.

Test Result Highlights

MAPI Messaging Benchmark (MMB)	9,800
Response Time (milliseconds)	468
Messages Submitted (four-hour steady-state period)	72,163
Message Recipients Delivered (four-hour steady-state period)	401,286
Messages Sent (four-hour steady-state period)	19,464

The Netfinity 3500 M10 Server

The IBM Netfinity 3500 M10 server is the new-generation, value-priced foundation for your file-and-print serving, e-mail, shared Internet access and basic application serving needs. Designed for enhanced manageability in enterprise environments, the entry-level Netfinity 3500 M10 is backed by the same extensive service and support provided with Netfinity mainstream and enterprise servers. The results reported in this document demonstrate the outstanding performance that you can expect from the Netfinity line of products.

Key features of the Netfinity 3500 M10 include:

- 90-Day Server Startup Support - Providing free support for system installation, Network Operating System and third-party ServerProven™ products.
- Integrated Environmental Monitoring.
- PFA - Predictive Failure Analysis on select components.
- Netfinity Manager comprehensive system management tool.
- Capacity Manager artificial intelligence tool enabling administrators to identify and eliminate bottlenecks before they occur.
- Support for Alert on LAN advanced monitoring and alerting of system failure.
- ServerGuide™ setup and configuration tool.
- Support for SMA - System Migration Assistant allowing Windows NT® configuration information to be easily transferred from existing machines to the Netfinity 3500 M10 during upgrades.
- Support for LCCM - LANClient Control Manager allowing remote and automatic formatting of drives and deploying of the Operating System and other software across multiple systems over the network using Windows NT.
- Complete 3-year on-site warranty, including all Netfintiy options.

Benchmark vs. Production Configuration Disclosure Note

This test measures the messaging throughput of a single-server, single-site topology. Its purpose is to measure the maximum throughput of a Microsoft® Exchange Server on this hardware configuration. This benchmark can provide a means for comparing hardware and/or software products, but it cannot be used as a deployment guide for production environments. For deployment-specific information, contact a Microsoft or an IBM representative, or visit www.microsoft.com/exchange/support/DeployAdmin/DeployAdmin.htm.

The MMB benchmark does not account for:

- Usage profiles that do not match the Load Simulator MAPI Medium profile

- Per-user storage and per-server backup requirements

- Fault-tolerance requirements

- Workloads other than Messaging Application Programming Interface (MAPI) private folder access, including Public Folder, NNTP, POP3 and other e-mail interfaces

- Multiple Exchange Server deployments, where additional resources are required to forward mail intra-site Connectors, Links and replication to remote Exchange sites

Test Results

Summary	
Supported Benchmark Load	9,800 MMBs
Benchmark Profile	MAPI Messaging Benchmark
Protocol	Exchange MAPI
Length of Steady State	4 hours
Length of Test	8 hours
Unless otherwise noted, values listed are averages over the entire steady-state period.	
Transaction Load (hourly)	
Messages Submitted	18,040
Message Recipients Delivered	100,321
Messages Sent	4,866
Transaction Load (per Second)	
Message Opens/Sec	37,699
Folder Opens/Sec	9.553
Remote Procedure Call (RPC) Read Bytes/Sec	32,332
RPC Write Bytes/Sec	189,250
Transaction Queues	
Information Store (IS) Send Queue Average Length	49.563
Message Transfer Agent (MTA) Work Queue Average Length	3.813
Processor Utilization	
System Processor Utilization (%)	86.790%
System Processor Queue Length	6.021
System Context Switches/Sec	3533.504
Process % CPU Time - Store	100%
Process % CPU Time - Directory Store (DS)	10.496%
Process % CPU Time - MTA	6.515%
Memory Utilization	
Available Bytes	5,076,882
Pages/Sec	0.103
Process Working Set Bytes - Store	962MB
Process Virtual Bytes - Store	1.549GB
Logical Drive Utilization	
IS Database Disk Reads/Sec	0.008
IS Database Disk Writes/Sec	0.032
IS Database Average Disk Queue Length	7.662
IS Log Disk Reads/Sec	0.036
IS Log Disk Writes/Sec	104.646
IS Log Average Disk Queue Length	0.838

Descriptive Terms

Messages Submitted: Submit calls made by clients, which equates to total messages sent by users.

Messages Sent: Messages that the Store sends to the MTA. Normally all the messages submitted by the clients are sent to the MTA except in the case where all the recipients are local

mailboxes. In that case, since all the deliveries can be performed locally, no message is sent to the MTA.

Message Recipients Delivered: Separate mailboxes that messages have been delivered to.

Message Opens/Sec: Messages accessed for reading per second.

Folder Opens/Sec: Folders opened for browsing per second.

RPC Read Bytes/Sec: Bytes read from clients, sent via RPCs.

RPC Write Bytes/Sec: Bytes written to clients, sent via RPCs.

IS Send Queue Average Length: Send Queue Size is the number of messages in the private information store's send queue.

MTA Work Queue Average Length: Work Queue Length is the number of outstanding messages in the Work Queue, which indicates the number of messages not yet processed to completion by the MTA.

Response Times (Latencies)

Client Actions	95th Percentile Response Time (in Milliseconds)
Read	340
Send	711
Delete	651
Move	771
Submit	601

Usage Profile

Summary

Default LoadSim Profile Used	Medium
Length of Day	4 hours
Messages Sent/MMB/Day	14.727
Average Recipients per Message	5.56
Messages Read/MMB/Day	48.45

Configuration

Hardware	
Vendor	International Business Machines Corporation
Model	IBM Netfinity 3500 M10 – Model Type 8655-31Y
Processor	600MHz Pentium III
Number of Processors	Two
Primary Cache	32KB (16KB I+16KB D)
Secondary Cache	512KB
Other Cache	None
Memory	1024MB ECC SDRAM
Disk Subsystem	C: (1) 9.1GB 7200 RPM E: (1) 9.1GB 7200 RPM F: (5) 9.1GB 10K RPM (RAID-0) G: (15) 9.1GB 10K RPM (RAID-0)
Disk Controllers	(2) Netfinity ServeRAID-3HB Ultra2 SCSI Adapter
Other Hardware	(2) Netfinity 10/100 Ethernet Adapter
Hardware Tunings	Block Size set to 16KB , Write-Back enabled
Mail Software	
Vendor	Microsoft Corporation
Mail Server	Exchange Server 5.5 Enterprise Edition
Build/Release Version	5.5 - Build 1960.8
Service Pack/Patch Info	Exchange 5.5 Service Pack 3
Additional Software Tuning	None

OS Software	
Operating System/Version	Microsoft Windows NT Server 4.0 Enterprise Edition
Service Pack/Patch Info	Service Pack 5
File System Type	C: NTFS, E: NTFS, F: NTFS , G: NTFS
Other Software	None
Network	
Type of Network	(3) 100TX Segments
Network Speed	100Mbps Full Duplex
Load Generators	
Number of Load Generators (LG)	40
Total Number of LG Processes	40
Simulated Users/Process	(40) 245 MMB
Model	(40) IntelliStation® E Pro - Model 6893-56U
Processor	500MHz Pentium III
Number of Processors	One
Memory	128MB
Network Controller	Intel Ethernet 10/100 Adapter (on planar)
Operating System	Microsoft Windows NT Workstation 4.0 with Service Pack 5

Load Simulator Medium User Profile

Microsoft Exchange Server Overview

This section defines the default medium user profile as defined by Load Simulator. This profile includes messaging and Schedule+ activity and is designed to simulate an average corporate MAPI user.

All frequencies are based on an eight-hour day.

There are several issues to keep in mind when using or referring to Load Simulator profiles:

1. This standard user profile provides a basis for benchmarking the messaging throughput of different hardware and software configurations.

2. To maintain a consistent benchmark, this profile has not been updated over the years. Trends in user behavior such as larger message and attachment sizes have not been incorporated.

3. Load Simulator cannot exactly simulate the profile of a company's user population. The default profile can be tuned with data from research and live server monitoring, but it is important to note that a simulation tool has a finite amount of complexity and accuracy.

Summary

Message Traffic (per User, per Day)

Logon/Logoff	0 (user remains logged on the entire day)
Total Messages Sent	14.2
Average Recipients per Message	4.68
Messages Received	66.3
Messages Read	81.3
Messages Deleted	33.1
Messages Moved	16.6

Details

Test Initialization (per User)

Number of Messages in Inbox	4
Number of Messages in Deleted Items	1
Number of Private Folders	40
Messages per Private Folder	5

Configuration

Public Folder Post	No
Public Folder Browse	No
Logoff	No
Update Free/Busy Information	No

Schedule+ File Size (KB)

Avg	22
Min	5
Max	40
Use Distribution Lists	Yes
DLs per Site	30
Maximum DL size	20
Recipients per Message	3
Add DL to message	30%
Save copy in Sent Items Folder	Yes
Message Priority	Normal
Delivery Receipts	No
Read Receipts	No

Message Actions (per User, per Day)

Send New Mail	4
Browse Mail (in Private Folders)	15
Check Inbox	12
Reply	3.8 (7% of all reads)
Reply All	2.7 (5% of all reads)
Forward	3.8 (7% of all reads)
Delete	33.1 (40% of all reads)
Move	16.6 (20% of all reads)
Copy	0 (20% of all reads)
Load Message Attachments	2.6 (25% of messages with attachments)
Update Schedule	5

Message Mix

Body	Attachment	Weighting
1KB Rich Text Format (RTF)		60
2KB RTF		16
4KB RTF		4
1KB RTF	10KB text file	6
1KB RTF	43KB Embedded Bitmap Object	2
1KB RTF	16KB Word document	4
1KB RTF	15KB Excel spreadsheet	4
1KB RTF	17KB Embedded Excel Object	4

Testing Methodology

LoadSim

LoadSim is an Exchange client emulator provided by Microsoft to simulate client messaging workload to the server. The server perceives its actions as that of an actual client because it sends the same messages to the server as a client would when the user of the client software is taking some sort of action. Actions performed by LoadSim and the Microsoft Exchange Client are very similar in terms of server load; the user simulation features in LoadSim are designed specifically to mimic the Microsoft Exchange Client's use of Mail API.

Of equal significance, LoadSim is a multi-client emulator in that it allows for simulating large-scale, client-server scenarios without actually involving a multitude of client machines. LoadSim takes advantage of multiprocessing, multithreading, and shared memory features in Windows NT to create and manage up to several hundred simulated users on a single client computer. Our testing relied on LoadSim to:

- Test the performance and capacity of server hardware.

- Determine network bandwidth requirements, given varying server configurations.

- Conduct performance impact tests on various server configurations without affecting numerous actual users.

User-behavior models define how many times per day a user sends a message, how often the user reads a message, and what percentage of messages are deleted, saved in folders, or

forwarded to others. LoadSim allows user behavior models to be customized to a quite detailed level. LoadSim also provides a simple means of characterizing user loads according to predefined usage levels. These predefined user profiles, called canonical user profiles, provide a workload standard that can be used to compare data in common scenarios. LoadSim provides canonical user profiles for light, medium, and heavy users. For our testing, we used the medium canonical user.

As LoadSim runs a scenario, it probabilistically generates user activities according to the model provided. For example, if the model specifies that users send 22 messages in an eight-hour day, LoadSim will have each user in the scenario generate messages at an average rate of 2.75 messages per hour, although the intervals between messages will not be uniform. Random variations are included to simulate the behavioral patterns of actual users. A user action is simulated by LoadSim's generating the same MAPI calls to the server that a client would.

Generating Scores Using LsLog

When LoadSim runs a scenario, it creates two files. One of these, named loadsim.out, contains a history of all generated actions. Any errors that occurred during the run are logged in this file in the form of a screen dump from the LoadSim application window during the scenario run. This file is checked (in addition to NT Performance Monitor results) for errors that might possibly invalidate the entire run. Another file, named lsperf.log, contains a collection of all the individual action timings that result from the LoadSim run. This data is used to generate scores, that is, numerical values (measured in milliseconds) that represent the responsiveness of the server under the specified load. Results measuring the impact of component or other system variations, as well as variations in the number of users, are derived by averaging these scores.

The LoadSim Performance Log Parser is a Windows NT command-line executable file designed to manipulate data from the LoadSim performance log files. After each LoadSim test run, LsLog is used to merge the logged data from all test clients, and then to truncate unwanted data in order to extract the relevant time window. The result is used to generate the 95th-percentile result for all actions. The default 95th-percentile response time for a given action means that 95% of all response times recorded for that action were equal to or below that value. The average response time, measured in milliseconds, provides an indication of how responsive the server is to the user. A small number means quicker (hence better) response. Microsoft recommends 1,000 milliseconds (one second) or less as the suggested response time for Exchange Server.

Windows NT Performance Monitor

LoadSim scores represent the end result of a scenario run. To properly understand what actually occurred during the course of a run, it is necessary to use the Windows NT Performance Monitor (PerfMon) during the test to collect performance data, which must be analyzed when the test is complete. In addition to the operating system-specific counters included as a default with Windows NT, the Exchange Server includes counters specific to Exchange to provide further means to track Exchange's performance. PerfMon's record indicates bottlenecks and problem areas. Studying the results logged by PerfMon ensures that the results reflected in the LoadSim scores are not corrupted by factors outside the scope of the testing methodology.

Performance Criteria

LoadSim was configured to simulate 9,800 medium MAPI e-mail users for eight hours. The test environment reached a stable state after two hours, so the results from the first two hours of the test run were discarded. The readings from the remaining six hours were sampled to provide the average response time over four hours of steady-state period.

Response Time: Obviously, users do not want to experience a long wait for the program to respond to their actions. A response time of about one second is considered reasonable.

CPU Utilization: The greater the amount of processing power used by Exchange, the less efficient the server's performance will be. This ties in to response time; higher CPU utilization generally leads to slower response times.

Send Queue Size: The Exchange Server's send queue is used to store messages that have been sent by a client, but have not yet been processed by the server. This is needed to allow for the random nature of mail use: Statistically, there will be times when many users are accessing Exchange at once, and other times when fewer users than normal access it. Ideally, the send queue should accumulate messages at times of heavy use, and then empty itself when usage falls below the average. Therefore, in a system with sufficient capacity, the send queue should be empty for a significant portion of the time (at these times, the server is processing messages as soon as they come in). This will result in an average send queue volume that is effectively zero.

Additional Information

For more information on IBM Netfinity direction, products and services, refer to the following white papers, available from our Web site at www.pc.ibm.com/netfinity.

Management

Integrating IBM Netfinity Manager with Microsoft Systems Management Server
Integrating IBM Netfinity Manager with Intel LANDesk Server Manager
IBM Netfinity Manager 5.2
IBM Netfinity Manager Plus for Tivoli Enterprise Overview
IBM Netfinity Advanced Systems Management
IBM Netfinity Advanced Systems Management for Servers
IBM ServerGuide for Netfinity and PC Server Systems

Other Topics

IBM Netfinity 3000 Exchange 5.5 MAPI Messaging Benchmark (MMB) Performance Result
Capacity Planning for Netfinity on Windows Terminal Server
Enterprise Storage Solutions
Fibre Channel Solutions for Enterprise Storage
IBM Chipkill Memory
IBM Netfinity X-architecture
IBM ClusterProven Program on Netfinity
IBM Netfinity Predictive Failure Analysis
IBM Netfinity Cluster Directions
IBM Netfinity Web Server Accelerator
Implementing Microsoft IIS on Netfinity 5500 M10
IBM Netfinity Availability Extensions for Microsoft Cluster Server
IBM Netfinity ESCON Adapter
IBM Netfinity Hot-Plug Solutions
IBM Netfinity Storage Management Solutions Using Tape Subsystems
IBM Netfinity Storage Area Networks
IBM Netfinity 8-Way SMP Directions
IBM Netfinity Server Ultra2 SCSI Directions
IBM Netfinity Server Quality
IBM Netfinity 3500 M10 Server
At Your Service...Differentiation beyond technology

12