



MailMarshal 6.0 SMTP Sizing Guide

White Paper

June 2004

Contents

MailMarshal Sizing Guidelines.....	1
Minimum Hardware and Software Requirements.....	2
Performance Matrix.....	4
Performance Tuning Recommendations	8
Server Sizing Summary	10

The purpose of this whitepaper is to present some detail that can be used when planning the hardware specifications of a MailMarshal 6.0 SMTP Server.

The minimum recommended specifications will be given as well as information on what hardware components affect the performance of MailMarshal the most.

There is also a performance matrix that can be used as a guide to determine the throughput of different server specifications; this will assist to determine what the specification of your MailMarshal server should be.

Finally there is section on tuning recommendations

Please remember that the specifications detailed and the performance data provided make the assumption that MailMarshal is on its own dedicated server, apart from where stated.

Authored by Bradley Anstis

THIS DOCUMENT AND THE SOFTWARE DESCRIBED IN THIS DOCUMENT ARE FURNISHED UNDER AND ARE SUBJECT TO THE TERMS OF A LICENSE AGREEMENT OR A NON-DISCLOSURE AGREEMENT. EXCEPT AS EXPRESSLY SET FORTH IN SUCH LICENSE AGREEMENT OR NON-DISCLOSURE AGREEMENT, NETIQ CORPORATION PROVIDES THIS DOCUMENT AND THE SOFTWARE DESCRIBED IN THIS DOCUMENT "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. SOME STATES DO NOT ALLOW DISCLAIMERS OF EXPRESS OR IMPLIED WARRANTIES IN CERTAIN TRANSACTIONS; THEREFORE, THIS STATEMENT MAY NOT APPLY TO YOU.

This document and the software described in this document may not be lent, sold, or given away without the prior written permission of NetIQ Corporation, except as otherwise permitted by law. Except as expressly set forth in such license agreement or non-disclosure agreement, no part of this document or the software described in this document may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, or otherwise, without the prior written consent of NetIQ Corporation. Some companies, names, and data in this document are used for illustration purposes and may not represent real companies, individuals, or data.

This document could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein. These changes may be incorporated in new editions of this document. NetIQ Corporation may make improvements in or changes to the software described in this document at any time.

© 1995-2004 NetIQ Corporation, all rights reserved.

U.S. Government Restricted Rights: If the software and documentation are being acquired by or on behalf of the U.S. Government or by a U.S. Government prime contractor or subcontractor (at any tier), in accordance with 48 C.F.R. 227.7202-4 (for Department of Defense (DOD) acquisitions) and 48 C.F.R. 2.101 and 12.212 (for non-DOD acquisitions), the government's rights in the software and documentation, including its rights to use, modify, reproduce, release, perform, display or disclose the software or documentation, will be subject in all respects to the commercial license rights and restrictions provided in the license agreement.

Check Point, FireWall-1, and Provider-1 are trademarks or registered trademarks of Check Point Software Technologies Ltd.

ActiveAgent, ActiveAnalytics, ActiveAudit, ActiveKnowledge, ActiveReporting, ADcheck, AppAnalyzer, Application Scanner, AppManager, AuditTrack, AutoSync, Chariot, ClusterTrends, CommerceTrends, Configuration Assessor, ConfigurationManager, the cube logo design, DBTrends, DiagnosticManager, Directory and Resource Administrator, Directory Security Administrator, Domain Migration Administrator, End2End, Exchange Administrator, Exchange Migrator, Extended Management Pack, FastTrends, File Security Administrator, Firewall Appliance Analyzer, Firewall Reporting Center, Firewall Suite, Ganymede, the Ganymede logo, Ganymede Software, Group Policy Administrator, Intergreat, Knowledge Scripts, Migrate.Monitor.Manage, Mission Critical Software, Mission Critical Software for E-Business, the Mission Critical Software logo, MP3check, NetIQ, the NetIQ logo, the NetIQ Partner Network design, NetWare Migrator, OnePoint, the OnePoint logo, Operations Manager, PentaSafe, PSAudit, PSDetect, PSPasswordManager, PSSecure, Qcheck, RecoveryManager, Security Analyzer, Security Manager, Server Consolidator, SQLcheck, VigilEnt, Visitor Mean Business, Vivinet, W logo, WebTrends, WebTrends Analysis Suite, WebTrends for Content Management Systems, WebTrends Intelligence Suite, WebTrends Live, WebTrends Log Analyzer, WebTrends Network, WebTrends OLAP Manager, WebTrends Report Designer, WebTrends Reporting Center, WebTrends Warehouse, Work Smarter, WWWorld, and XMP are trademarks or registered trademarks of NetIQ Corporation or its subsidiaries in the United States and other jurisdictions. All other company and product names mentioned are used only for identification purposes and may be trademarks or registered trademarks of their respective companies.

MailMarshal Sizing Guidelines

When sizing a MailMarshal server the most important factor is to determine the expected through-put of the solution and the volumes of Internet SMTP traffic that a particular company receives.

Email today is required to be a highly available application and so it is always recommended to build redundancy into the MailMarshal installation by having more than a single MailMarshal server available at all times. Most of the scenarios that are defined in this Whitepaper are based on what a single server might be capable of processing so when deciding on the final hardware specifications it is recommended that you deploy at least two individual MailMarshal servers to meet your required throughput.

Basically, the higher specification of your server then the more mail Traffic you can expect to process.

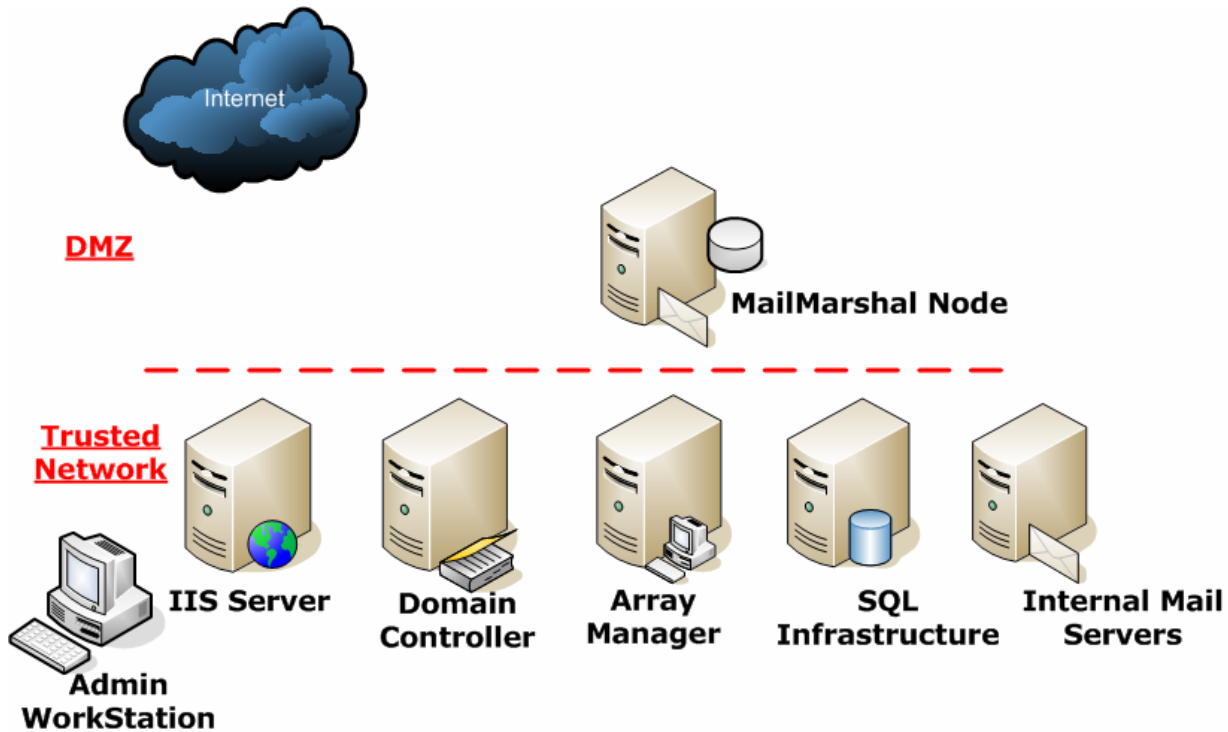
In this scenario, server specifications are seen as a combination of:-

- Hard drive subsystem speed
- Processor Speed
- Installed Memory

Overall the Hard Drive subsystem speed will have the most effect on the expected through-put of the MailMarshal server; this is then followed by the speed of the installed processor and finally the amount of installed memory.

Minimum Hardware and Software Requirements

What Components do I need?



Required Components

- MailMarshal Node
 - Mail Processing Server
- Array Manager
 - New MailMarshal Management Server
- SQL Infrastructure
 - Stores configuration data as well as logging info

Optional Components

- IIS Server
 - Used for the new WebConsole and Spam Quarantine Management tool (SQM)

Deciding how many servers to deploy, and what components will be on which servers would be based on the customer's current environment and their requirements for the new implementation. All the performance statistics noted below are for a MailMarshal node server running just as the mail processing node, the Array Manager, SQL Server etc all being on a single separate server.

Consult the *MailMarshal 6.0 SMTP Architecture guide* for more information on deployment options.

MailMarshal Node Server

Hardware

An Intel Pentium III 500MHz Class PC with 256MB RAM and 1GB Available Hard Disk

Software

Operating System: Microsoft Windows 2000

Windows Server 2003

Windows XP

Microsoft SQL Server 2000 SP3 or greater, or Microsoft Database Engine 2.0 SP3¹

MailMarshal Web Components – Internet Information Server

Hardware

An Intel Pentium III 500MHz Class PC with 256MB RAM and 500MB Available Hard Disk

Software

Operating System: Microsoft Windows 2000 SP4 or greater

Windows Server 2003

Windows XP

Microsoft ASP.Net 1.1 is required²

MailMarshal Administration Tool / Console

Hardware

An Intel Pentium III 500MHz Class PC with 128MB RAM and 5MB Available Hard Disk

Software

Operating System:

Microsoft Windows 2000

Windows Server 2003

Windows XP

Note 1: Microsoft Database Engine 2.0 (MSDE 2000) with SP3 pre-applied is supplied with MailMarshal. SP3 for SQL Server 2000 is also supplied with MailMarshal.

Note 2: Microsoft .Net Framework 1.1 is supplied with MailMarshal.

NB: All pre-requisites for installing MailMarshal on Windows 2000 or greater are supplied on the MailMarshal CD. No other installation software should be required other than that supplied on the CD.

Performance Matrix

MailMarshal Server Sizing

These guidelines set out typical email usage patterns and associated minimum hardware specifications for sizing and scoping a new MailMarshal SMTP installation.

MailMarshal is a fast, multi-threaded, high-performance SMTP content scanning relay that performs equally well in small or large enterprises. It fully supports multiple processor servers allowing it to scale to the largest of organizations.

All these estimations are based on having a single MailMarshal server. It is however recommended to have more than a single server for redundancy considerations.

MailMarshal SMTP Hardware Sizing Table

Email Users	Messages Per day	Text Log Size Per day in MB	Quarantine Directory & Spam Logs per day in MB	CPU GHz	Memory MB	H/Drive GB
100	2,000	21	9	0.5	256	1
250	5,000	53	23	0.7	256	2
500	10,000	106	46	1.0	512	5
1,000	20,000	212	93	1.8	512	10
5,000	100,000	1,061	464	2.0	768	32
10,000	200,000	2,122	927	2.8	1000	64
25,000	500,000	5,305	2,318	2 x 2.8	2000	120
50,000	1,000,000	10,610	4,635	6 x 2.8	2000	240

Notes:

1. Hardware specifications are based on email usage for a typical corporate environment. Your environment may differ in email volume or content. If unsure, please contact your NetIQ partner for further sizing recommendations.
2. The server sizing depicted above are assuming that the MailMarshal server is just a MailMarshal node, the Array Manager, SQL Server and so on are on separate servers
3. Number and volume of email based on 20 messages per day to/from the Internet for each user, an average email message size of 50Kb, spread over a ten-hour business day. Spam is estimated at 50% and the average size of Spam is 4.5 KB
4. Text logs are kept for 5 days, default setting for Quarantine folders retention is 7 days
5. When logging to a SQL database, sites of 1000 users and over are recommended to use the full SQL Server version, as the MSDE version is subject to database size and other constraints. These sites are also recommended to run SQL Server on a separate server.
6. It is assumed MailMarshal will be operating on a dedicated server. Ensure there is enough free resource if other applications are also installed on the same server (e.g. another email application or proxy server).
7. In all cases, typical rules are assumed, including integration with a dll-based anti-virus scanner. Extra resources may be needed if using virus scanners invoked from a command line.
8. Sites with more than 50,000 users may require enhanced hardware. Please contact your NetIQ partner for a recommended configuration.

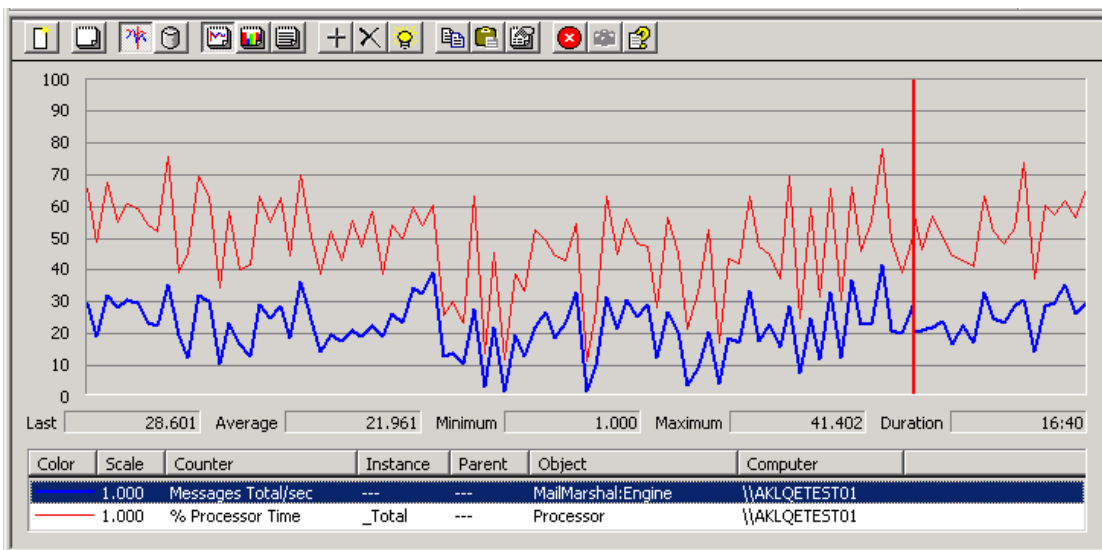
Single MailMarshal Server vs. Components deployed on separate servers

All the Server sizing information defined above are assuming that you have a separate MailMarshal Node server. However if you are wanting to have all the MailMarshal components deployed on the same server, that is the MailMarshal Node, The Array Manager and the SQL server all on the same server then the comparisons below can be used.

Test Scenario 1

Single DL 360 Server, Dual 2.6 GHz Xeon CPU, 2 GB Ram & Dual 146GB Drives

As a single server, all components deployed on the same server:-

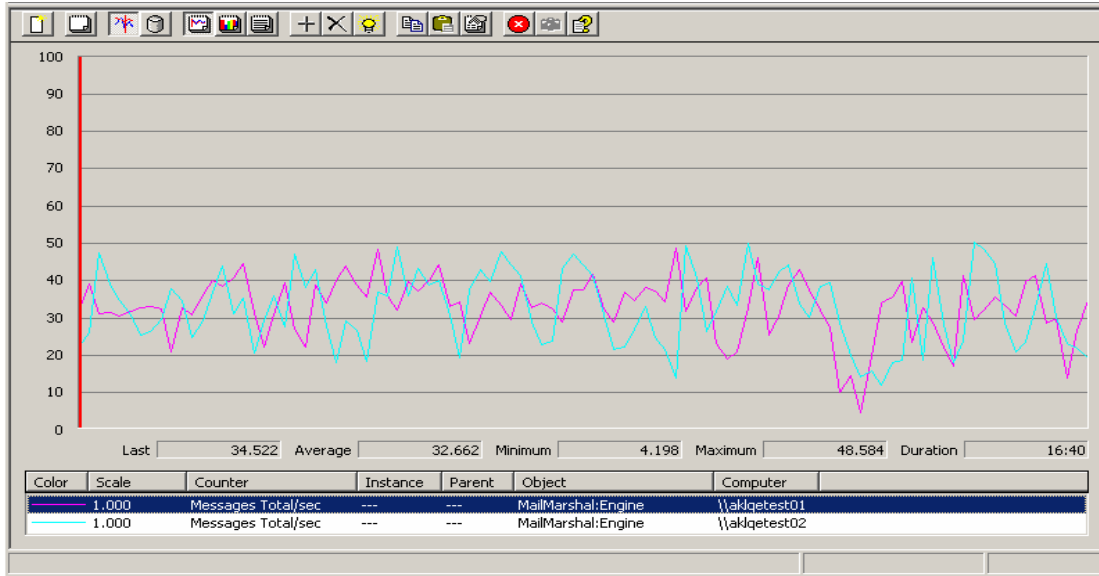


Average 21 messages / second, 5 GB Mail data processed per hour

Test Scenario 2

Single DL 360 Server, Dual 2.6 GHz Xeon CPU, 2 GB Ram & Dual 146GB Drives

As a MailMarshal Node, all other components deployed on separate server:-



Average 33 messages / second, 7 GB Mail data processed per hour

SQL Database Sizing

To work out how much space to allow for the MailMarshal 6.0 database, as a general rule allow 1024 bytes per message that will be logged. If we also assume an average user sends and receives 20 SMTP messages / day then this provides the following data.

Users	Messages / Day	SQL Data/day	SQL Data over 100 Days*	Database Type
100	2,000	2 MB	200 MB	MSDE or SQL
250	5,000	5 MB	500 MB	MSDE or SQL
500	10,000	10 MB	1 GB	MSDE or SQL
1,000	20,000	20 MB	2 GB	SQL
5,000	100,000	100 MB	10 GB	SQL
10,000	200,000	200 MB	20 GB	SQL
25,000	500,000	500 MB	50 GB	SQL
50,000	1,000,000	1 GB	100 GB	SQL

*100 Days is the default period for logging information to be kept. This is also not allowing for any non-working days so in reality this figure is generous.

MailMarshal 5.5 Comparison

As a comparison test MailMarshal 5.5 was installed onto one of the nodes that was used for the 6.0 testing, its default configuration was used with Sophos Anti-Virus enabled. Also the unpacking directory was set to the second disk to spread disk i/o and logging was directed to a separate SQL server.

MailMarshal 6.0 clearly has improved on the message throughput capabilities of MailMarshal 5.5. MailMarshal 6.0 (32.66 messages/sec) is approximately 40% quicker than MailMarshal 5.5 (23.45msgs/sec) while offering a significantly improved console functionality.

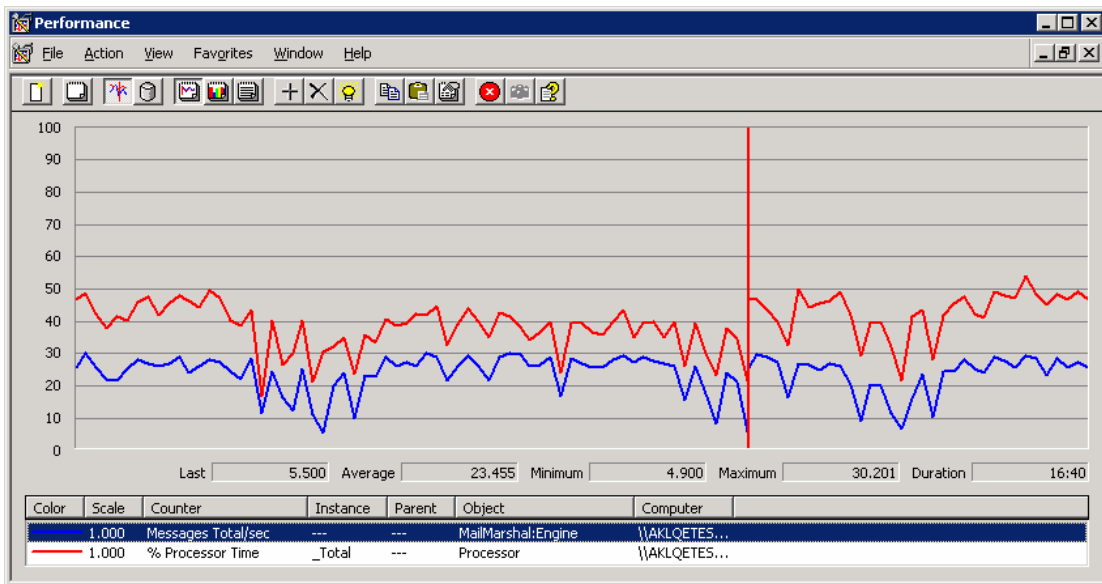


Figure 1 MailMarshal 5.5 message throughput: 16 min window, 10 sec sample

Performance Tuning Recommendations

Unpacking Directory

During the initial testing the DL 360 servers were configured so the 2 disks were mirrored – this would be a typical deployment for a 1U rack mount server. It quickly became apparent that disk queue length, which was averaging around 12, was a problem.

The servers were then reconfigured to provide 2 separate disks and the Unpacking directory was moved onto the second disk. This promptly increased the message throughput by close to 50% together with making the machine generally more responsive.

However this would make a DL360 unsuitable as a node since there is no data protection, a 2U rack mount server chassis that supports 4 disks so they could be setup as 2 mirrored pairs would be ideal.

Location of the Array Manager

The location of the Array Manager can affect the performance of the administration and configuration tools used in MailMarshal, but will not necessarily affect the mail processing performance.

The recommended locations for the Array Manager are listed below, from most to least preferred:-

1. On the same server as the SQL Server holding the MailMarshal database. This is because the Array Manager is the only component that actually communicates directly to the MailMarshal database. All the other components communicate to the database through the Array Manager.
2. On a Server that is as close as possible to the SQL Server holding the MailMarshal database, and has a LAN speed network connection to it.
3. On the same server as an Active Directory Global Catalog. The other component that the Array Manager communicates to on a regular is a Global Catalog if you are running Active directory, or through LDAP to another existing Directory server.
4. On a smaller site, the Array Manager, MailMarshal Node and the SQL Database (MSDE supported) could all be on the same server.

Dual CPU's vs. Quad CPU's

Observation of the Nodes showed both a high CPU usage and high Disk Queue, this indicates a reasonable balance between the I/O and processor performance. It is suspected that going to a quad CPU machine would not show significant improvement in message throughput unless it was also accompanied by a suitable improvement in disk I/O capability.

Strictly from a price performance basis adding more dual processor boxes would be cheaper than going to quads.

Memory Requirements

The memory footprint of the Nodes was relatively light even under full load, rarely getting above 256M, however this can vary substantially with some large configurations and large user groups, 1G of memory is recommended in these situations.

Database File groups

A SQL Database can be comprised of either a single file, or can be split across multiple files; these multiple files are called file groups.

The MailMarshal database was created which defaults to placing everything in the PRIMARY file group, the IX_From and IX_To indexes were then moved to a second file group that was on a different physical disk to the PRIMARY file group.

Above 500 – 1000 users the memory footprint is quite high. SQL server will take everything it can get; 3G of memory is recommended so SQL can take its maximum of 2G and still leave some for the Array Manager and OS should these be installed on the same server.

Normal operation of the Nodes puts only a light load on the SQL server, using the Consoles and reports places a heavier burden on it.

Server Sizing Summary

Observation of several Email servers has shown that it is not unusual for 80% of email traffic to arrive in a 10 hour period. The same observation also has shown that the flow of email traffic is fairly linear during that 10 hour window. Larger sites tend to have more geographically spread offices which can increase this time window substantially, but let's assume our users must send their email in a 10 hour window.

Therefore for 100,000 users we need to ship $100,000 * 20 \text{ msgs/person} * 80\%$ in 10 hours, this is 45 msgs/sec.

Based on the test scenarios above, this would require 2 modern dual processor servers and a separate SQL / Array Manager Server. It would be prudent to then add a third server to ensure that there is no latency in email traffic even in extreme peaks and to provide sufficient capacity should one server fail.

Another measure for server sizing could be done by bandwidth, each Node can deal with a peak of 10Mbit/sec of email, if you only have a 10Mbit/sec link to the Internet then 2 Nodes would be sufficient to give some leeway and cope with a Node failure.

Based on bandwidth, a full-duplex E1 connection should support $2\text{M} / 10\text{bits/byte} / 30\text{K} = 6.6\text{msgs/sec}$ inbound typically this would mean 3.3msgs/sec outbound for a total of 10 msgs/sec. Relating this back to an average user count we get $10\text{msgs/sec} * 10\text{hours} * 60\text{mins/hour} * 60\text{secs/min} / (20\text{msgs/user} * 80\% \text{ in } 10\text{hours}) = 22,500 \text{ users}$.

Although the test scenarios showed that a MailMarshal Node could sustain 32Msgs/sec it would generally be unwise to design a system where the expected load per Node averaged above 20Msgs/sec. A connectivity fault could result in large volumes of email being held and then delivered to MailMarshal in a hurry; at that point the extra capacity will be desirable.

It is readily possible to build MailMarshal environments that will support well in excess of 100,000 users. Over the period of the week spent testing the various scenarios MailMarshal processed over 15 million messages without fault.

Key findings, based on 20 messages per user per day with an average size of 30K where 80% of these are sent in a 10 hour window:

- 1 standalone MailMarshal can sustain 20msgs/sec.
this is easily sufficient to cope with a dedicated E1 or T1 email connection or around 22,500 users.
- 1 MailMarshal Node with separate Array Manager can sustain 32 msgs/sec. This is the generally recommended deployment option for MailMarshal as reporting and console actions do not impact Node throughput.
This is sufficient for a dual E1 connection or around 45,000 users.
- 2 MailMarshal Nodes with separate Array Manager can sustain 64 msgs/sec.
This solution would cope with up to 90,000users

Even under high load the console was responsive and messages could easily be located and appropriate actions taken.