

WebMarshal Performance and Scalability Whitepaper

October, 2007

Contents

WebMarshal Server Specifications	2
Database Sizing	3
Load Balancing	4
Supported Environments	4
Testing Methodology	5

This document is intended to give administrators sizing information for WebMarshal 6.0.

WebMarshal Server Specifications

WebMarshal 6.0 was tested on a server with the specifications below. Servers with different specifications will have different results – increasing the speed of the disk subsystem or processor will significantly impact the maximum number of users WebMarshal can support.

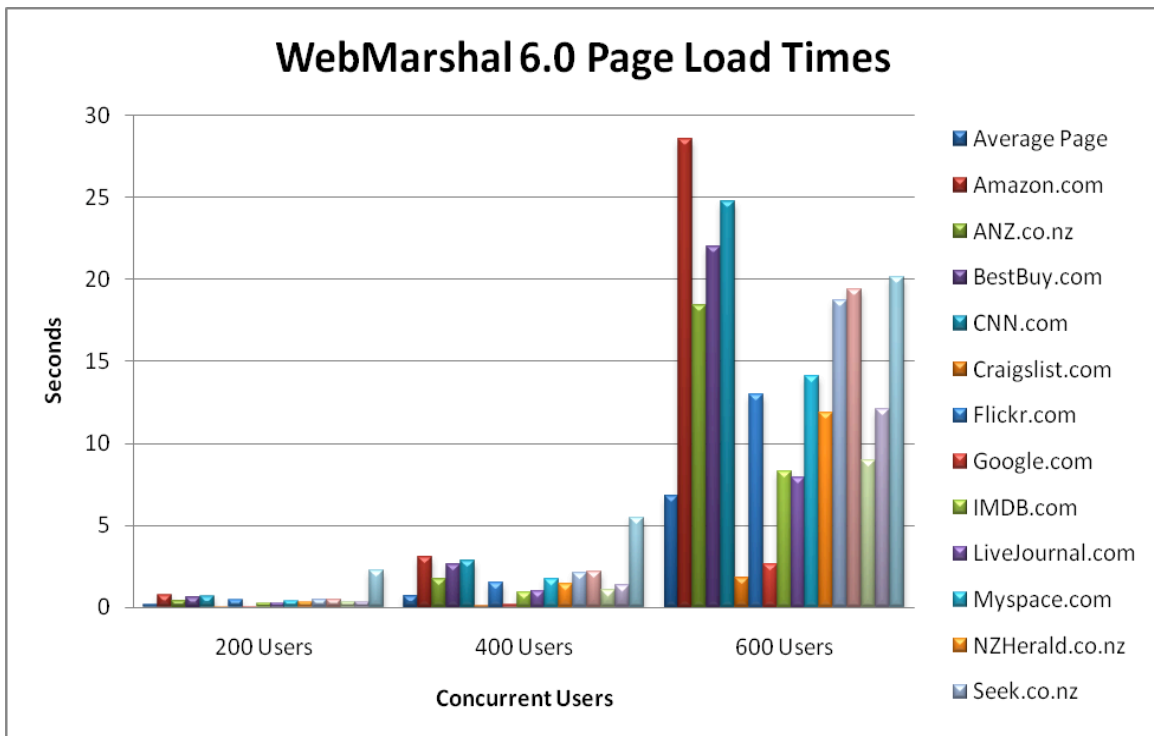
In addition, these results were obtained using the methods described in the Testing Methodology section of this whitepaper; real life results will vary based on ruleset, browsing habits, and Internet connection speed. This server was used as a dedicated Content Processing Node joined to an array. Its Array Manager and database server were located on another server. No other applications were active on the server.

*3.4 Ghz Pentium 4 CPU
2 GB SDRAM
250 GB SATA hard drives (Mirrored)
Windows Server 2003 with Service Pack 2
WebMarshal 6.0.2.2865*

The graph below shows the page load times for different numbers of concurrent simulated users. At 200 users, the page load times are approximate to the load times users would see if they were accessing the page directly, without WebMarshal filtering their web requests.

At 400 users, the page load times are slightly over double what they were on at 200, but are still well within acceptable levels. For example, an average-sized page still loads in under a second, and a large page, such as Amazon.com, loads within 3 seconds.

At 600 users, the page response times become unacceptable, with some pages taking nearly 30 seconds to load.



Marshal recommends that a server of this specification be used to support no more than around 400 concurrent users. Depending on your environment’s web usage, a box of this specification can be used to support different total user counts.

	Low Usage	Medium Usage	High Usage
% of Total Users Concurrently Browsing	10%	20%	30%
Number of Total Users	4000 Users	2000 Users	~1300 Users

Database Sizing

WebMarshal’s database is used to store both user browsing history and parts of its configuration data for reporting purposes. The configuration data stored within the database is small in size and relatively static, at about 3 megabytes. Most of the space used in the WebMarshal database contains user browsing information.

Each user browsing session logged to the database takes a variable amount of space, depending on the length of the session, the number and content of the websites visited, and whether sites or files were blocked by WebMarshal.

Marshal’s testing showed that an average session, consisting of a period about 5 minutes of continuous browsing over 4-5 websites, logged about 5.5k worth of data to the database. Our testing also revealed that an average user initiates about 10 sessions per day.

Therefore, using the formula *(average session database log size) * (average number of sessions per day) * (number of users)*, we would expect the database for a 1000 user organization to grow by approximately 54 megabytes a day. Over a period of 30 days (allowing for weekends) of history retention, the database would be approximately 1.2 gigabytes.

User Count	Database Growth per Day	Total Size for 30 Days of Logging	Total Size for 100 Days of Logging
100	~5.4 MB	~120 megabytes	~408 megabytes
500	~26.8 MB	~591 megabytes	~2 gigabytes
1000	~54 MB	~1.2 gigabytes	~4 gigabytes

It is important to note that SQL Server 2005 Express Edition has a 4 gigabyte cap its databases. Therefore, SQL Server 20005 Express Edition is not appropriate for customers with large numbers of users, or for customers who need to retain logging data for long periods of time.

Load Balancing

If multiple WebMarshal Content Processing Nodes are required to service a given environment, they should be joined together in a WebMarshal Array. Once joined in an Array, the Nodes' policy can be centrally managed and consolidated reports become available.

However, while Nodes in an Array share policy, WebMarshal does not provide any load balancing features. In order to distribute load across multiple nodes, one of the following strategies can be used:

- Manually configuring proxy settings to point to specific nodes
- Round-robin DNS records.
- Microsoft Network Load Balancing Services
- Dedicated hardware load balancer (*i.e., F5 BIG-IP Load Balancers*)

Detailed instructions on configuring Microsoft's Load Balancing Services with WebMarshal can be found in Marshal Knowledge Base Article [Q11845](#).

Supported Environments

Operating System:

Windows Server 2003 (all versions)

Windows XP Professional [not recommended for production use]

Databases:

Microsoft SQL Server 2005 Express Edition

Microsoft SQL Server 2005 (all versions)

Authentication Services:

Microsoft NT Domains

Microsoft Active Directory

Novell Directory Server

IP Address

Testing Methodology

Marshal used a series of tools and custom scripts to simulate large numbers of users browsing through WebMarshal. A single node received thousands of requests from several load-generation servers, accessing sample Web pages.

These sample pages were downloaded from popular websites such as Google, MySpace, and CNN, and mirrored on a web server in the local test network. In addition, an “average” page was created which contained an average-sized HTML page containing several images.

The simulated users’ web browser requested Web pages using multiple threads to load components in real time as the page was being “rendered.” After the page had finished loading, each simulated user would pause for a random amount of time to reflect the page being read by the user. Typically, the time between page loads was around 17 seconds.

It is important to make two points about Marshal’s testing. Firstly, because the pages were hosted on a local Web server, the load times are much lower than what one would expect to see in real life, where a request must traverse the Internet and back.

Secondly, Marshal used a ruleset that was very close to the default WebMarshal ruleset, but it did differ in several ways. Several TextCensor rules and URLCensor rules were disabled, which impacted performance positively. Performance with the default ruleset would be slightly lower than what is reflected in this paper.

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, MARSHAL LIMITED 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 JURISDICTIONS 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 Marshal, 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 Marshal. 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. Marshal may make improvements in or changes to the software described in this document at any time.

© 2007 Marshal Limited, all rights reserved.

U.S. Government Restricted Rights: The software and the documentation are commercial computer software and documentation developed at private expense. Use, duplication, or disclosure by the U.S. Government is subject to the terms of the Marshal standard commercial license for the software, and where applicable, the restrictions set forth in the Rights in Technical Data and Computer Software clauses and any successor rules or regulations.

Marshal, MailMarshal, the Marshal logo, WebMarshal, Security Reporting Center and Firewall Suite are trademarks or registered trademarks of Marshal Limited or its subsidiaries in the United Kingdom 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.



Marshal's Worldwide and EMEA HQ
Marshal Limited,
Renaissance 2200,
Basing View,
Basingstoke,
Hampshire RG21 4EQ
United Kingdom

Phone: +44 (0) 1256 848080
Fax: +44 (0) 1256 848060

Email: emea.sales@marshal.com

Americas
Marshal Inc.
5909 Peachtree Dunwoody Road NE,
Suite 770,
Atlanta,
GA 30328
USA

Phone: +1 404 564-5800
Fax: +1 404 564-5801

Email: americas.sales@marshal.com
info@marshal.com | www.marshal.com

Asia-Pacific
Marshal Software (NZ) Ltd
Suite 1, Level 1, Building C
Millennium Centre
600 Great South Road
Greenlane, Auckland
New Zealand

Phone: +64 9 984 5700
Fax: +64 9 984 5720

Email: apac.sales@marshal.com