Rosetta Sizing Options

Note: This document should be used for reference only. Actual Rosetta environment preparation should not be done based solely on this document.
Actual sizing recommendations will be given by the Ex Libris Sizing Manager following a thorough analysis of the customer environment and its needs.

CONFIDENTIAL INFORMATION

The information herein is the property of Ex Libris Ltd. or its affiliates and any misuse or abuse will result in economic loss. DO NOT COPY UNLESS YOU HAVE BEEN GIVEN SPECIFIC WRITTEN AUTHORIZATION FROM EX LIBRIS LTD.

This document is provided for limited and restricted purposes in accordance with a binding contract with Ex Libris Ltd. or an affiliate. The information herein includes trade secrets and is confidential

DISCLAIMER

The information in this document will be subject to periodic change and updating. Please confirm that you have the most current documentation. There are no warranties of any kind, express or implied, provided in this documentation, other than those expressly agreed upon in the applicable Ex Libris contract. This information is provided AS IS. Unless otherwise agreed, Ex Libris shall not be liable for any damages for use of this document, including, without limitation, consequential, punitive, indirect or direct damages.

Any references in this document to third-party material (including third-party Web sites) are provided for convenience only and do not in any manner serve as an endorsement of that third-party material or those Web sites. The third-party materials are not part of the materials for this Ex Libris product and Ex Libris has no liability for such materials.

TRADEMARKS

"Ex Libris," the Ex Libris Bridge to Knowledge, Primo, Aleph, Voyager, SFX, MetaLib, Verde, DigiTool, Rosetta, bX, URM, Alma , and other marks are trademarks or registered trademarks of Ex Libris Ltd. or its affiliates.

The absence of a name or logo in this list does not constitute a waiver of any and all intellectual property rights that Ex Libris Ltd. or its affiliates have established in any of its products, features, or service names or logos.

Trademarks of various third-party products, which may include the following, are referenced in this documentation. Ex Libris does not claim any rights in these trademarks. Use of these marks does not imply endorsement by Ex Libris of these third-party products, or endorsement by these third parties of Ex Libris products.

Oracle is a registered trademark of Oracle Corporation.

UNIX is a registered trademark in the United States and other countries, licensed exclusively through X/Open Company Ltd.

Microsoft, the Microsoft logo, MS, MS-DOS, Microsoft PowerPoint, Visual Basic, Visual C++, Win32, Microsoft Windows, the Windows logo, Microsoft Notepad, Microsoft Windows Explorer, Microsoft Internet Explorer, and Windows NT are registered trademarks and ActiveX is a trademark of the Microsoft Corporation in the United States and/or other countries.

Unicode and the Unicode logo are registered trademarks of Unicode, Inc.

Google is a registered trademark of Google, Inc.

Copyright Ex Libris Limited, 2023. All rights reserved.

Document released: June 2023

Web address: http://www.exlibrisgroup.com

Table of Contents

1 Introduction 4

2 Rosetta Sizing 4

Table 1 – Main Characteristics for the Various Topologies 5

3 Rosetta Main Topology Options 6

Option 1: 3-Tier Topology – High Throughput 6

High availability, separated delivery/deposit 6

Option 2: 3-Tier Topology- Moderate Throughput 8

High availability, separated delivery/deposit 8

Option 3: 2-Tier Topology with High Availability - Moderate Throughput 9

High availability 9

Option 4: 2-Tier Topology, Standard Throughput 10

Minimal sizing for Rosetta 10

Sandbox Environment 11

1

Introduction

The purpose of this document is to help customers and Ex Libris experts determine the best Rosetta topology that will suit the institution's requirements.

Rosetta Sizing

The initial sizing is based on Ex Libris best practices. Once the customer is operational, they might decide to add more hardware to the topology. Rosetta supports adding more servers to the topology after the initial installation, to improve performance.

It is highly recommended to read Rosetta's Ingesting Digital Content at Scale whitepaper to understand the various factors affecting performance, among them sizing.

A separation between Operational servers and Delivery/Deposit servers is not mandatory. All functionalities can run as part of the Application servers. It is recommended to segregate between Front End servers and Operational servers if the customer wants either to allow different network zones or to improve performance.

For example, for an institution that requires high availability of delivery and that needs to eliminate any delay on delivery due to high processing levels on the operational servers, it is recommended to have a separate delivery server(s) solely in charge of processing delivery calls.

Another example is an institution that allows the general public to deposit material into Rosetta. In this scenario, the customer would probably need to locate the deposit server in a dedicated network zone (e.g. a DMZ) in order to quarantine the files deposited by the public until the files are marked safe by the validation checks.

Rosetta storage architecture is based on the OAIS model, with several storage layers used in different stages that the deposited content goes through. It is recommended to read about the Rosetta storage management in the Rosetta System Administration Guide.

The size of the Deposit Network Storage may vary according to the files' retention period in the deposit storage. The numbers provided below assume a retention period of one month.

Note: Final required storage of Permanent and Access copy storage depends on the customers load. This should be calculated according to total planed TB's that are planned to be processed and stored on the system.

Table 1 – Main Characteristics for the Various Topologies

Topology	Number of Application Servers	High Availability?	Dedicated Delivery Server?	Dedicated Deposit Server?	Separated network Zones for Delivery / Deposit ?	Dedicated UI Management and Processes Server?
3-Tier, High Throughput	5 +	Yes	Yes	Optional	Optional	Yes
3-Tier, Moderate Throughput	2 +	Yes	Yes	Optional	Optional	Optional
2-Tier, Moderate Throughput	2 +	Yes	No	No	No	Optional
2-Tier- Standard Throughput	1	No	No	No	No	No

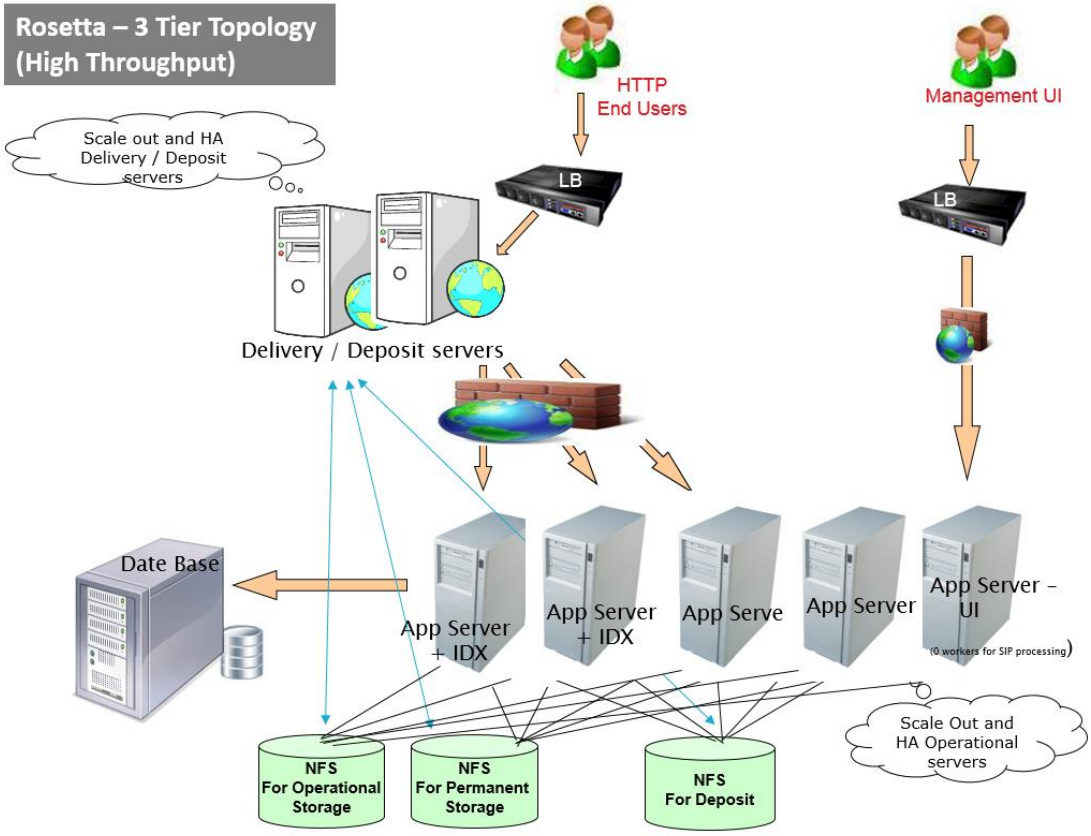
2

Rosetta Main Topology Options

Option 1: 3-Tier Topology – High Throughput

High availability, separated delivery/deposit

This sizing is intended mainly for institutions with large amounts of managed data. Its main advantages are high availability of the system, high throughput, and options for dedicated delivery servers and for separated network zones for the delivery and / or deposit servers. The sizing can be tweaked by adding or reducing hardware according to actual performance and required throughput.



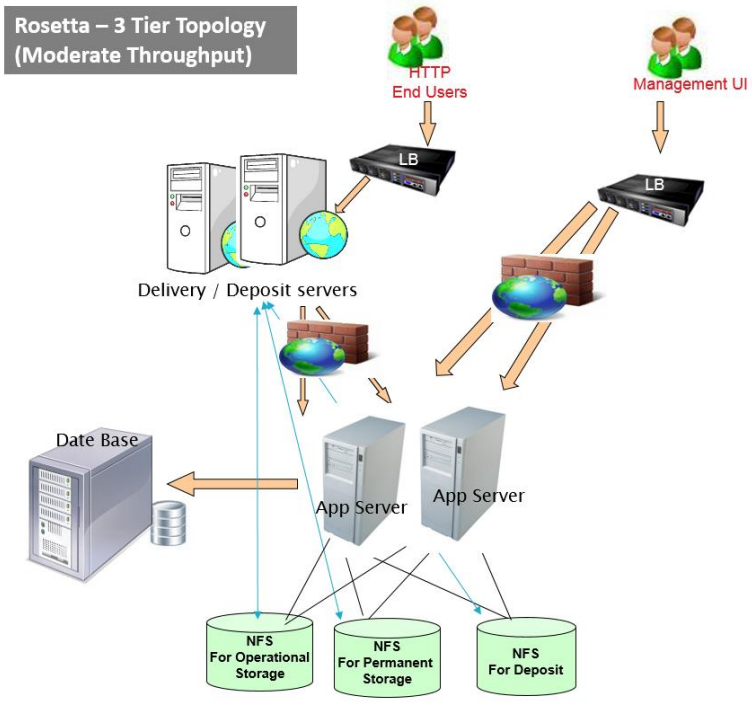
3-Tier Topology (2 Delivery/Deposit Servers + 5 Operational Servers + 1 DB Server)					
Machine Type (Linux 64bit, 2.0Ghz Intel E5 Ivy Bridge / equivalent AMD 22nm Cores)	**Number of Servers**	**Cores**	**Memory (GB)**	**Disk Space**	**Additional Info**
DB Server	1	8	16	50GB	Network / Local
DB Storage (data files)				150GB	Network
Deposit / Delivery Server	2	8	24	100GB	For each server
Operational / Permanent Server	5	8	24	100GB	For each server
Deposit Storage (/deposit_storage)				40TB	Network
Permanent Storage (/permanent storage)				According to customer's calculation (Minimum 50GB)	
Operational Storage (/operational_storage)				According to customer's calculation (Minimum 50GB)	
Operational Shared (/operational_shared)				Minimum 50GB	

Note: All NFS drives should have shared R/W permissions on all servers.

Option 2: 3-Tier Topology- Moderate Throughput

High availability, separated delivery/deposit

This sizing is intended mainly for institutions with a moderate amount of managed data. Its main advantages are high availability of the system, and options for dedicated delivery servers and for separated network zones for the delivery and / or deposit servers. The sizing can be tweaked by adding or reducing hardware according to actual performance and required throughput.



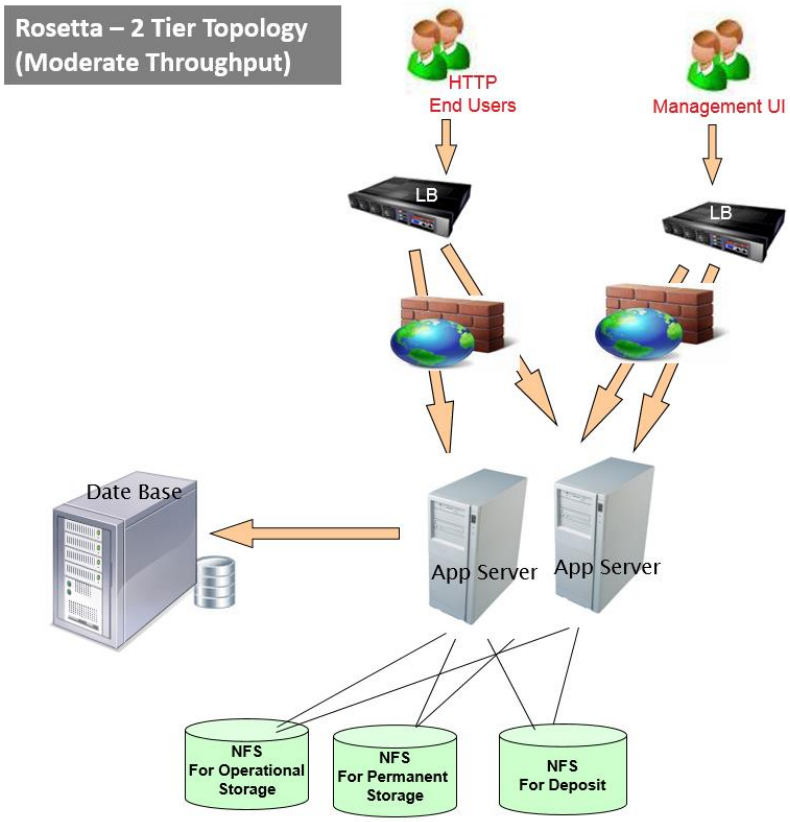
3-Tier Topology (2 Delivery/Deposit Servers + 2 Operational Servers + 1 DB Server)					
Machine Type (Linux 64bit, 2.0Ghz Intel E5 Ivy Bridge / equivalent AMD 22nm Cores)	**Number of Servers**	**Cores**	**Memory (GB)**	**Disk Space**	**Additional Info**
DB Server	1	8	16	50GB	Network / Local
DB Storage (data files)				150GB	Network
Deposit / Delivery Server	2	8	24	100GB	For each server
Operational / Permanent Server	2	8	24	100GB	For each server
Deposit Network Storage				20TB	Network
Permanent Network Storage				According to customer's calculation	
Access Copy Network Storage				According to customer's calculation	

Note: All NFS drives should have shared R/W permissions on all servers.

Option 3: 2-Tier Topology with High Availability - Moderate Throughput

High availability

This sizing is intended mainly for institutions with a moderate amount of managed data. Its main advantages is high availability of the system. The sizing can be tweaked by adding or reducing hardware according to actual performance and required throughput.



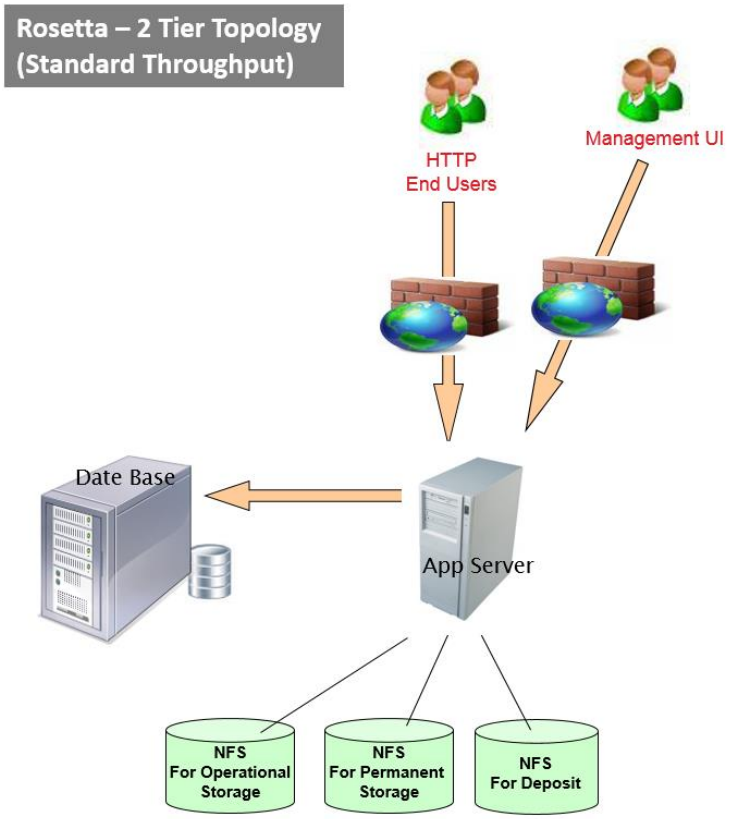
2 Tier Topology (2 Operational Servers + 1 DB Server)					
Machine Type (Linux 64bit, 2.0Ghz Intel E5 Ivy Bridge / equivalent AMD 22nm Cores)	**Number of Servers**	**Cores**	**Memory (GB)**	**Disk Space**	**Additional Info**
DB Server	1	8	16	50GB	Network / Local
DB Storage (data files)				150GB	Network
Operational / Permanent Server	2	8	24	100GB	For each server
Deposit Network Storage				Up to 20 TB	Network
Permanent Network Storage				According to customer's calculation	
Access Copy Network Storage				According to customer's calculation	

Note: All NFS drives should have shared R/W permissions on all servers.

Option 4: 2-Tier Topology, Standard Throughput

Minimal sizing for Rosetta

This sizing is intended mainly for institutions with a standard amount of managed data. It is a minimal-supported Rosetta topology. The sizing can be tweaked by adding hardware according to actual performance and required throughput.



2 Tier Topology (1 Operational Server + 1 DB Server)					
Machine Type (Linux 64bit, 2.0Ghz Intel E5 Ivy Bridge / equivalent AMD 22nm Cores)	**Number of Servers**	**Cores**	**Memory (GB)**	**Disk Space**	**Additional Info**
DB Server	1	8	16	50GB	Network / Local
DB Storage (data files)				120GB	Network
Operational / Permanent Server	1	8	24	100GB	For each server
Deposit Network Storage				10TB	Network
Permanent Network Storage				According to customer's calculation	
Access Copy Network Storage				According to customer's calculation	

Note: All NFS drives should have shared R/W permissions on all servers.

Sandbox Environment

Standard All-in-One Staging Server					
Machine Type (Linux 64bit, 2.0Ghz Intel E5 Ivy Bridge / equivalent AMD 22nm Cores)	**Number of Servers**	**Cores**	**Memory (GB)**	**Disk Space**	**Additional Info**
Basic Staging/Testing Server	1	8	24	100GB	Network / Local
Network Storage				1TB	For staging environment

Note: All NFS drives should have shared R/W permissions on all servers.