



Alfresco Content Services 5.2 On Premises

Reference Architecture

Copyright 2017 by Alfresco and others.

Information in this document is subject to change without notice. No part of this document may be reproduced or transmitted in any form or by any means, electronic or mechanical, for any purpose, without the express written permission of Alfresco. The trademarks, service marks, logos, or other intellectual property rights of Alfresco and others used in this documentation ("Trademarks") are the property of Alfresco and their respective owners. The furnishing of this document does not give you license to these patents, trademarks, copyrights, or other intellectual property except as expressly provided in any written agreement from Alfresco.

The United States export control laws and regulations, including the Export Administration Regulations of the U.S. Department of Commerce, and other applicable laws and regulations apply to this documentation which prohibit the export or re-export of content, products, services, and technology to certain countries and persons. You agree to comply with all export laws, regulations, and restrictions of the United States and any foreign agency or authority and assume sole responsibility for any such unauthorized exportation.

You may not use this documentation if you are a competitor of Alfresco, except with Alfresco's prior written consent. In addition, you may not use the documentation for purposes of evaluating its functionality or for any other competitive purposes.

This copyright applies to the current version of the licensed program.

Document History

DOCUMENT VERSION	DESCRIPTION OF CHANGE	DATE
1.0	First version of the Alfresco Content Services 5.2 On-Premises Reference Architecture	19 June 2017

Table of Contents

EXECUTIVE SUMMARY	1
DOCUMENT PURPOSE	1
ALFRESCO CONTENT SERVICES (v5.2) AND ALFRESCO ONE (v5.1)	1
INTENDED AUDIENCE	1
ASSUMPTIONS	1
GLOSSARY OF TERMS	3
ALFRESCO CONTENT SERVICES OVERVIEW	4
ALFRESCO CONTENT SERVICES ARCHITECTURE	4
ARCHITECTURE DIAGRAMS FOR THE STANDARD DEPLOYMENTS	5
DEPLOYMENT SIZE DEFINITIONS	5
<i>Diagram Key</i>	6
<i>Starter</i>	7
<i>Business</i>	8
<i>Enterprise</i>	9
<i>Enterprise – Scaled</i>	10
KEY COMPONENTS	11
<i>Repository and Share servers</i>	11
<i>Database server</i>	11
<i>Content Store</i>	11
<i>Alfresco Search Services (Index Server)</i>	11
<i>Load Balancing</i>	16
EXTENDED COMPONENTS	16
<i>Dedicated Ingestion Server</i>	16
<i>Dedicated Jobs Server</i>	18
<i>Alfresco Document Transformation Server</i>	20
<i>Desktop Synchronization</i>	21
<i>Alfresco Governance Services (Records Management or RM)</i>	22
HARDWARE SPECIFICATIONS	22
STARTER	22
BUSINESS	23
ENTERPRISE	23
ENTERPRISE - SCALED	24
SOFTWARE SPECIFICATIONS	24
SUPPORTED SOFTWARE STACKS	24
CONFIGURATION AND VALIDATION	25
REQUIRED PORTS	26
VALIDATION	26
BENCHMARKING	26
TUNING	27
SCALING	27
SIZING	27
RESOURCES	28

Executive Summary

This document is a high-level best practices guide for deploying Alfresco 5.x in an on-premises installation. This document primarily describes the recommendation for the Alfresco Content Services 5.2 deployment architectures.

Document purpose

The purpose of this document is to describe reference architectures for Alfresco Content Services 5.2 in an on-premises installation.

Reference architectures define a blueprint for customers to build systems that are capable of supporting a specific size of deployment. The deployment sizes are included in this document (see **Deployment size definitions** section).

Once the reference architecture for the required deployment size is installed on premises, customers can then benchmark, scale, and tune the system.

The purpose of this reference architecture document is to:

- Reduce risk for customers by offering a set of proven deployment architectures
- Enable high performance implementations
- Demonstrate how to scale
- Provide information to enable deployment in a variety of technical situations

This document endeavors to specify the hardware and software requirements for building an Alfresco Content Services system for several sizes of deployment (see **Deployment size definitions** section).

This reference architecture document is intended for use in conjunction with other Alfresco technical white papers, such as the **Sizing Guide** and the **Scalability Blueprint**. For more information about these documents, and other resources, see the **Resources** section.

Alfresco Content Services (v5.2) and Alfresco One (v5.1)

Alfresco has recently renamed parts of the product suite. Alfresco versions 5.0 and 5.1 are referred to as Alfresco One. Version 5.2 is referred to as Alfresco Content Services.

The reference architectures defined in this document are applicable to both products. Where an architectural recommendation is specific to or relies upon a feature that exists only in one version, this will be called out in the diagram or the accompanying text.

For the sake of brevity, both are referred to as Alfresco 5.x throughout this document.

Intended audience

This document assists anyone who has responsibility for building the infrastructure of an Alfresco Content Services 5.2 on-premises deployment.

The typical audience includes system designers, implementers, and administrators who will install and administer Alfresco products.

Assumptions

This document assumes a general knowledge of Alfresco products.

Those implementing the chosen reference architecture will require an in-depth knowledge of Alfresco Content Services administration and ideally be an Alfresco Certified Engineer (ACE) or an Alfresco Certified Administrator (ACA) (<https://university.alfresco.com/certification>). An in-depth knowledge of Alfresco Content Services clustering is also highly recommended.

For the specific supported platforms for Alfresco One and Alfresco Content Services, see the Supported Platforms page on the Alfresco website:

<https://www.alfresco.com/services/subscription/supported-platforms>

Supported software stacks for the reference architectures are described later in this document in the **Software specifications** section.

Hardware requirements for the different deployment sizes are covered in the **Hardware specifications** section.

Note: It is assumed that you can install and configure for scale your chosen operating system and DBMS.

It is assumed that the ratio of concurrent / total users is 1/10. That is, for every 10 users defined in Alfresco, 10% are using the system at any given time.

Glossary of terms

The following table provides a list of terms and acronyms used in this document:

Acronym/term	Meaning
ACS	Alfresco Content Services
ACA	Alfresco Certified Administrator
ACE	Alfresco Certified Engineer
AWS	Amazon Web Services
DBMS	Database Management System
EBS	Elastic Block Store
ECM	Enterprise Content Management
LB	Load Balancer
HA	High Availability
HDD	Hard Disk Drive
JVM	Java Virtual Machine
JDBC	Java Database Connectivity
SSD	Solid State Drive
vCPU	Virtual CPU
EVT	Environment Validation Tool

The main glossary, with an extensive list of acronyms and terms, can be found here:

Glossary of Terms
5.1: http://docs.alfresco.com/5.1/glossary/glossary.html 5.2: http://docs.alfresco.com/5.2/glossary/glossary.html

Alfresco Content Services overview

Alfresco Content Services is an Enterprise Content Management (ECM) platform that allows you to manage any type of content: from office documents, to scanned images, photographs, and even large video files. The built-in workflow feature allows companies to automate document-intensive business processes, saving time and money. The native collaboration features enable secure collaboration, inside or even outside the firewall.

Alfresco Content Services includes hybrid, on-premises, cloud, and mobile deployment options. As a hybrid ECM solution, it combines the security of an on-premises ECM platform with the agility and flexibility of the cloud.

With the hybrid functionality, document management, collaboration, and process automation can be synced to the cloud, avoiding the risks associated with uncontrolled consumer file sharing and enabling secure collaboration outside of the firewall.

Further information on Alfresco Content Services can be found on the Alfresco website (<http://www.alfresco.com>) and in the Alfresco documentation (<http://docs.alfresco.com>).

Alfresco Content Services architecture

At the core of the Alfresco Content Services architecture is a repository for storing content supported by a database server that persists metadata. Out-of-the-box applications provide standard solutions such as document management, records management, and web content management. Programming interfaces support multiple languages and protocols upon which developers can create custom applications and solutions.

Alfresco Content Services is a Java application that runs in an application server, such as Tomcat.

There are three main tiers:

- Repository tier - Repository used to store content and metadata
- Web tier – Alfresco Share or other user interface
- Index tier - Provides advanced search services based on Apache Solr

Depending on the use case and selected products, there may be other application tiers or components to service specific types of needs.

Other application tiers may include:

- Transformation tier - Handles certain types of document transformations
- Sync tier - Used to support the Alfresco Desktop Synchronization service
- Ingestion tier - Separates high volume content ingestion from user facing services
- Alfresco System Receiver - Allows replication of Alfresco managed content to a target file system

The reference architectures for larger deployment sizes use clustering of nodes to ensure that scaling capability. In the reference architectures the two main tiers, repository and Alfresco Share, run on the same node, although there can be multiple clustered nodes.

See the Alfresco documentation for more information on the main system components:

System Components
5.1: http://docs.alfresco.com/5.1/concepts/alfresco-arch-about.html
5.2: http://docs.alfresco.com/5.2/concepts/alfresco-arch-about.html

Note that although Alfresco Content Services supports a wide variety of content store options, the reference architectures assume the default content store (the File Content Store) is used.

Alfresco Content Services supports the clustering of nodes. This offers an option for horizontally scaling your system to meet your requirements. Further information on setting up **clustering** can be found in the Alfresco documentation:

High Availability
5.1: http://docs.alfresco.com/5.1/concepts/ha-intro.html
5.2: http://docs.alfresco.com/5.2/concepts/ha-intro.html

Architecture diagrams for the standard deployments

The following sections describe the architecture diagrams for the standard deployments.

Deployment size definitions

Alfresco defines a number of standard deployment sizes. These deployments are known as:

- . Starter
- . Business
- . Enterprise
- . Enterprise - Scaled

The following table provides the size definitions for the Starter, Business, Enterprise, and Enterprise - Scaled deployments. For the purposes of testing, Alfresco typically assumes 10% of the named users in a system will be active at any given time.

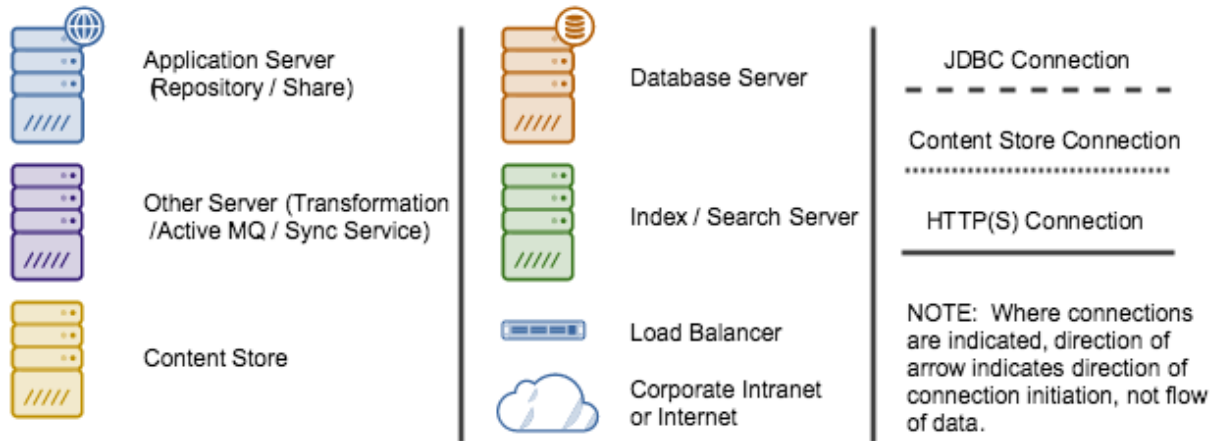
	Starter	Business	Enterprise	Enterprise - Scaled
Named Users	300	500	1000	5000

Other factors should be considered for the size of each deployment, such as the predominant type of document stored in the repository. For example, does the repository primarily store office documents, images, or XML content?

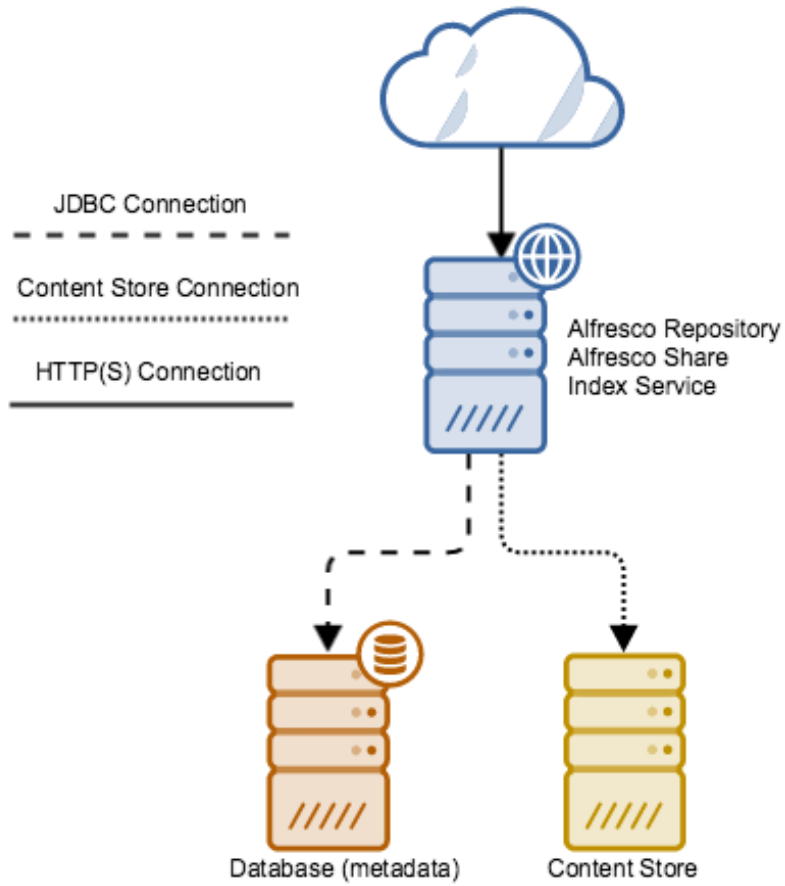
These additional factors are covered in the **Sizing Guide** and **Scalability Blueprint**. See the **Resources** section.

Diagram Key

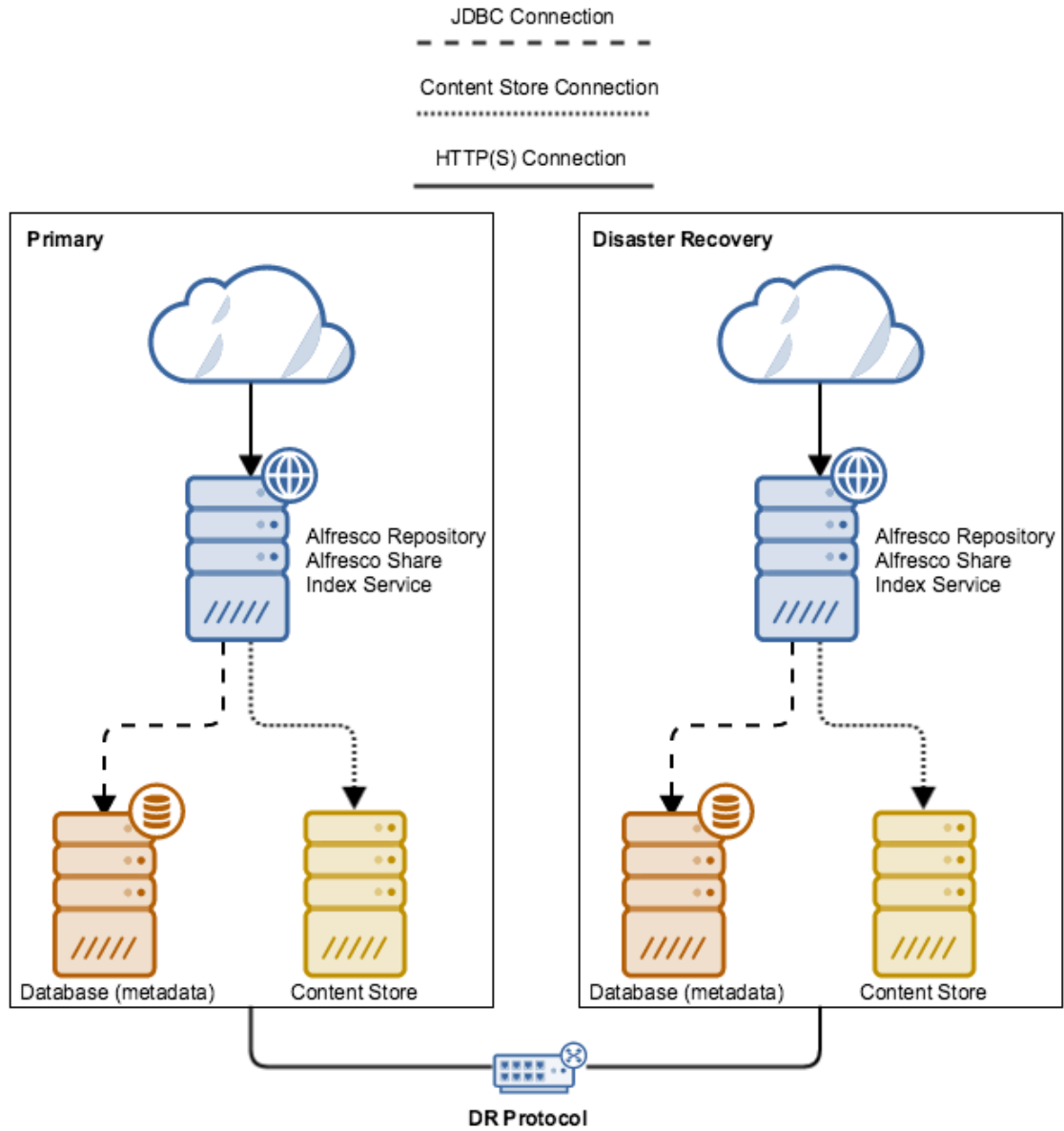
The diagram key defines the icon, color, and other conventions used across the Alfresco reference architectures. Where a specific icon or connection type is only used in a single diagram, that convention will be defined in the diagram itself.



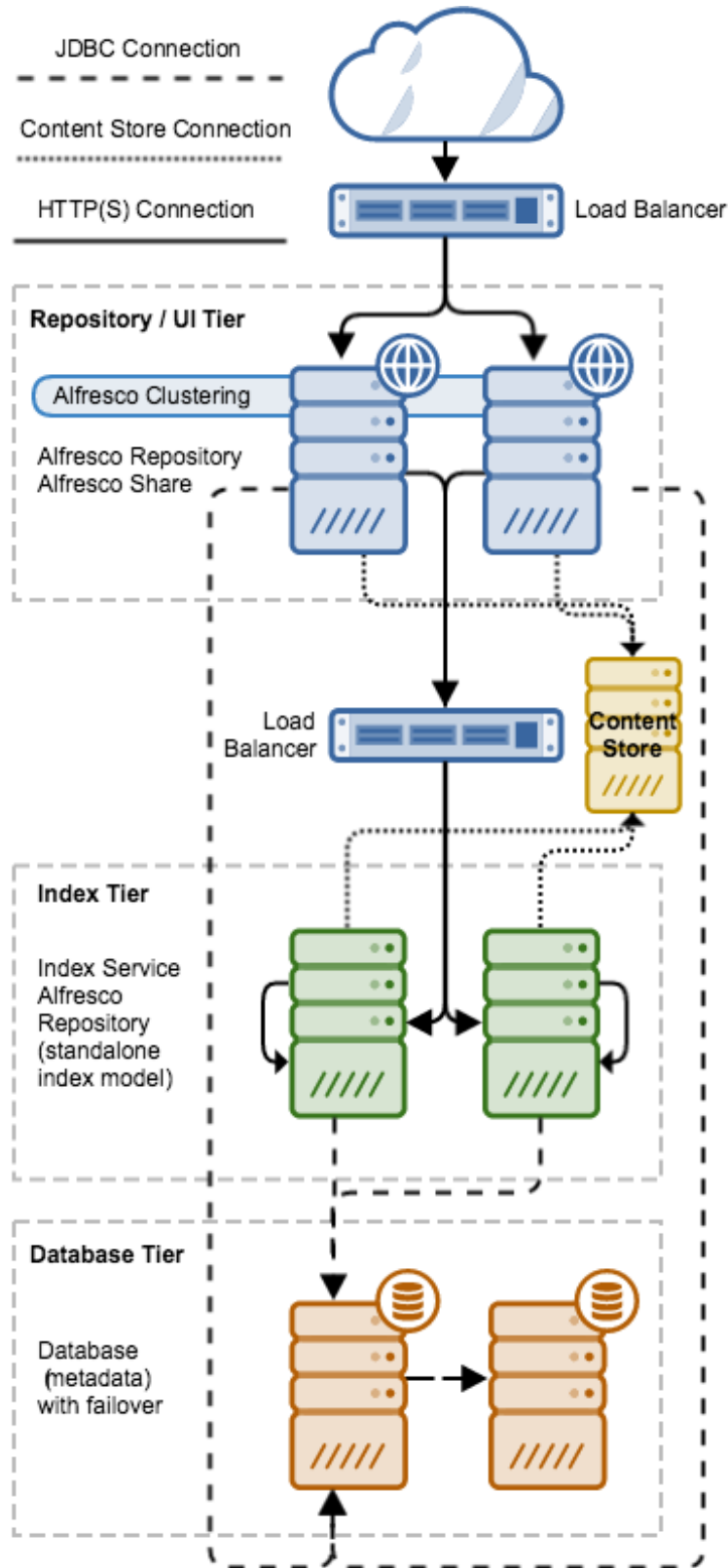
Starter



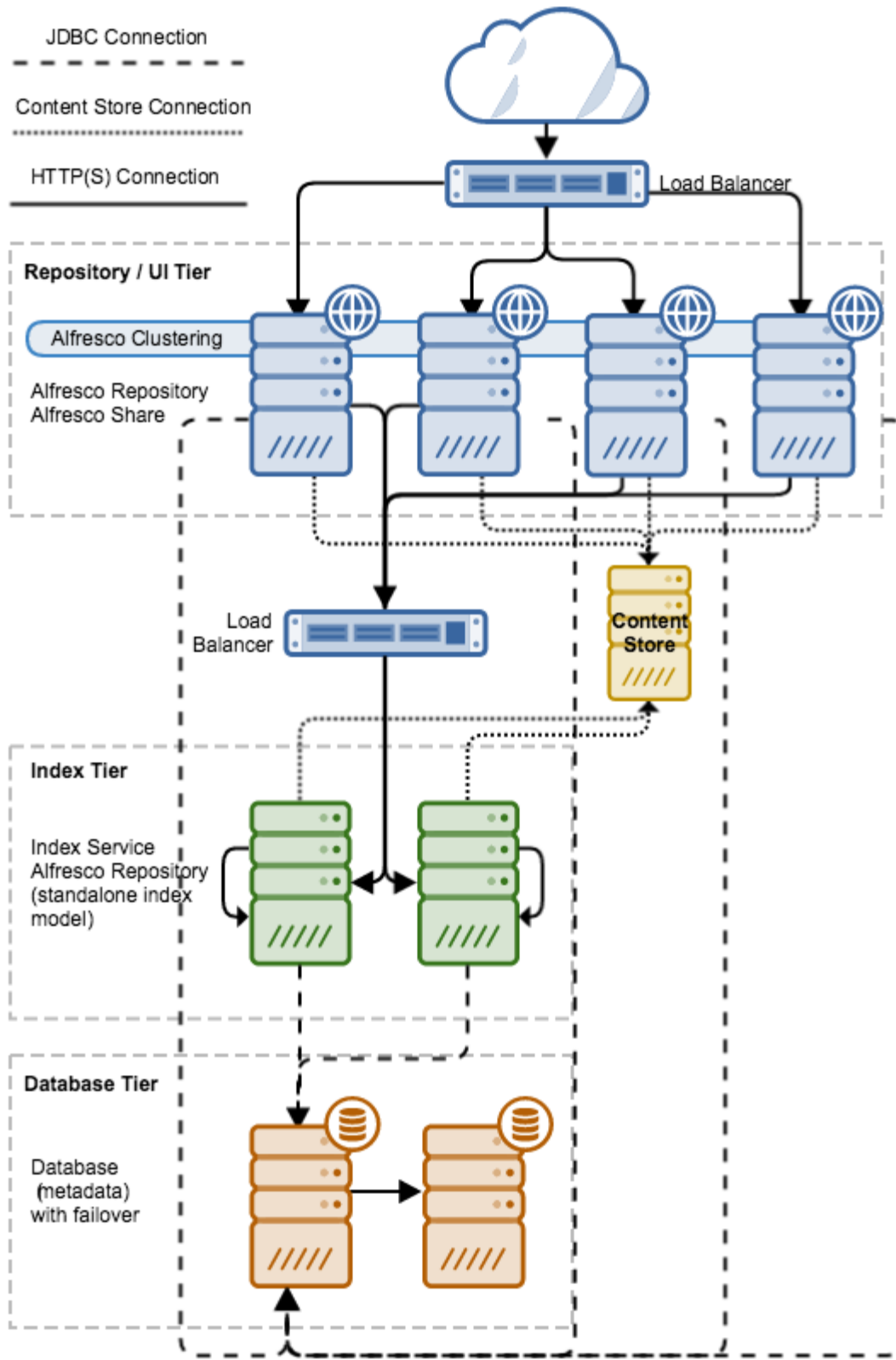
Business



Enterprise



Enterprise – Scaled



Key components

Each reference architecture diagram comprises the key components of an Alfresco Content Services deployment. This section outlines the specific implementations of each of the components.

Repository and Share servers

In each reference architecture, there is at least one server running the Alfresco application with Share. Where the size of deployment requires multiple nodes, access to the nodes are handled by a load balancer. This ensures that network requests (such as HTTP(s), FTP, and CIFS) are evenly distributed between the multiple nodes.

Note: The specific configuration of this load balancer is dependent on the protocols it is expected to handle.

Database server

The Alfresco and Share nodes, and index servers, are configured to communicate with a standalone database server. All metadata for repository content is stored in this database. The database server can represent a single point of failure, so you should have suitable failover systems in place.

Alternatively, advanced features of the selected database such as **replication** and **sharding** can be used, as required.

Configuration of advanced database features is vendor dependent and is out of the scope of this document.

Content Store

For the reference architecture, the default content store is assumed. This is the File Content Store. Alfresco supports many content store types, which are detailed in the official Alfresco documentation.

The content store is an abstraction, and using a different content store has little impact on the reference architecture.

Alfresco also supports multiple content stores. Where multiple content stores are involved, each store needs to be attached to all of the repository nodes in the cluster, as well as any deployed index servers.

Alfresco Search Services (Index Server)

Alfresco Search Services (based on Alfresco Content Services plus Apache Solr 6) or the Alfresco Index Server (based on Alfresco Content Services plus Apache Solr 4) is responsible for indexing content, processing search requests, and returning search results.

Detailed information about Solr and how it is used in an Alfresco Content Services deployment is available in the official Alfresco documentation:

Alfresco and Solr 4/6

5.1: <http://docs.alfresco.com/5.1/concepts/solr-home.html>

5.2: <http://docs.alfresco.com/5.2/concepts/search-home.html>

Alfresco support several architecture variations for the index and search subsystem. At its simplest, the index may be one or more Alfresco Index Servers that maintain separate, independent copies of the index.

Alfresco also supports index replication and sharding with Solr 4, and sharding on Solr 6. The illustrations shown in the architecture diagram section assume that one or more independent indexes will be deployed.

The next sections describe more advanced index subsystem architectures.

Replicated Index (Solr 4)

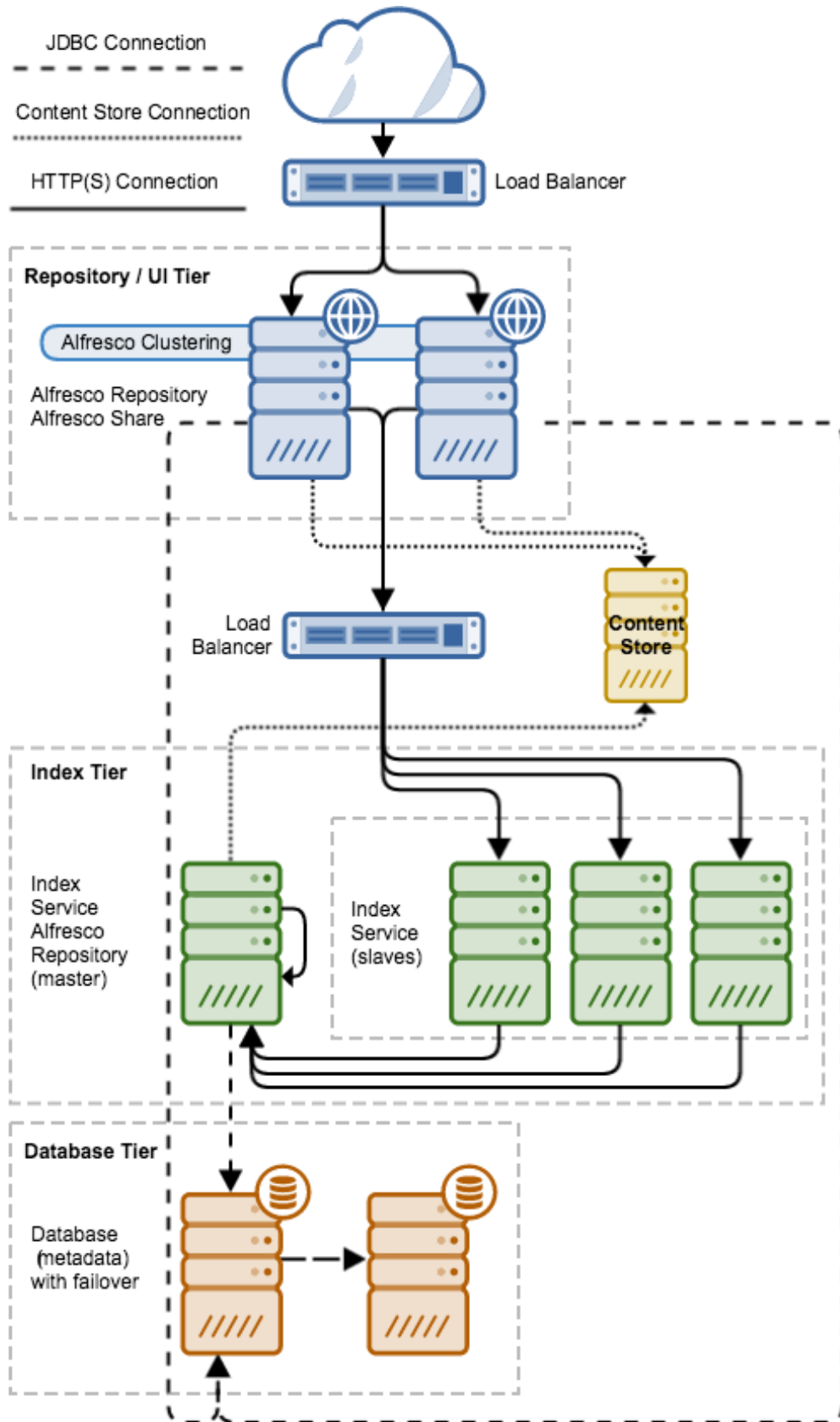
A replicated index can offer some significant advantages for small to medium repositories. In this architecture, a single search master builds the index. This master index is monitored by a number of slave indexes that update themselves as index changes occur. The advantages of this index architecture include lower pressure on the database, as only one index server is tracking changes, and the ability to handle large volumes of concurrent search requests quickly.

The primary disadvantage is scalability limitations that will be hit once the index size exceeds that which can be reasonably handled by a single large index. Information on configuring a replicated index is available in the Alfresco documentation.

Solr 4 Replication

5.1: <http://docs.alfresco.com/5.1/concepts/solr-replication.html>

5.2: <http://docs.alfresco.com/5.2/concepts/solr-replication.html>



Sharded Index (Solr 4 or Solr 6)

A sharded index offers many advantages over a single index, and is Alfresco's recommendation for large repositories. In a sharded index, the index is broken up into segments called shards. These shards are distributed across a number of servers, each of which is capable of servicing search requests for the shards it hosts. Redundancy is provided by hosting the same shard on multiple hosts. This option provides high scalability and performance up to the largest repositories that Alfresco has tested. Alfresco offers several options for slicing the index into shards, including by DBID, ACL, and properties. The exact options will differ between Solr 4 and Solr 6.

Alfresco Search Services with Solr 6 offers more options, and is recommended. Information on configuring a sharded index is available in the Alfresco documentation.

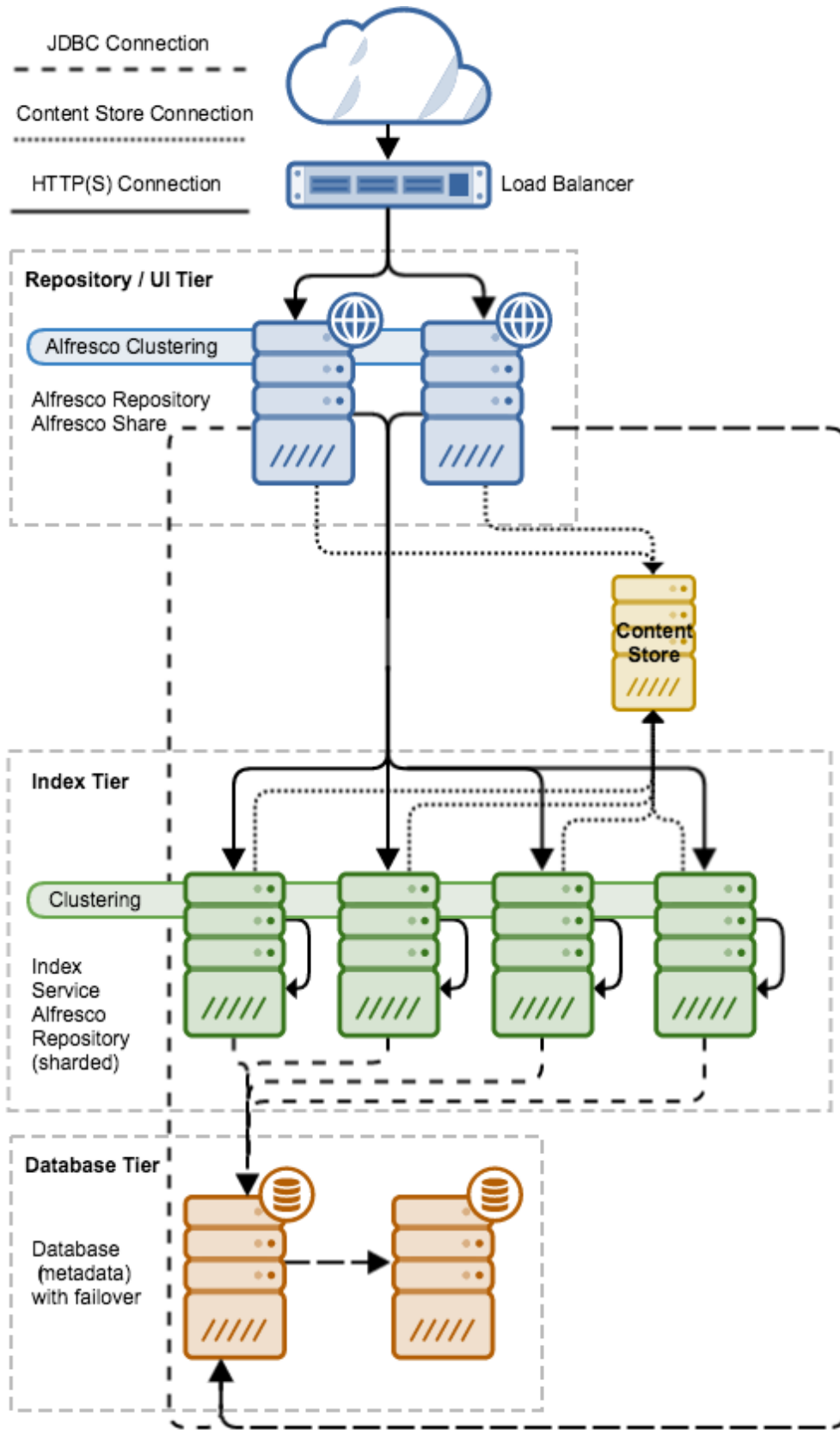
Note that when using a sharded index, all instances used for index tracking and content transformation (even those that are local to the index and search services instance) must be joined to the Alfresco cluster. Clustering is used for index service discovery and dynamic sharding.

Sharding in Alfresco

5.1: <http://docs.alfresco.com/5.1/concepts/solr-shard-overview.html>

5.2 (Solr 4): <http://docs.alfresco.com/5.2/concepts/solr-shard-overview.html>

5.2 (Solr 6): <http://docs.alfresco.com/5.2/concepts/solr6-shard-approaches.html>



Load Balancing

The choice of load balancer will depend on your preferred vendor and anticipated system loads. A range of possibilities may be deployed including dedicated hardware, or generic servers running load balancing software such as Nginx, Apache HTTP, and so on.

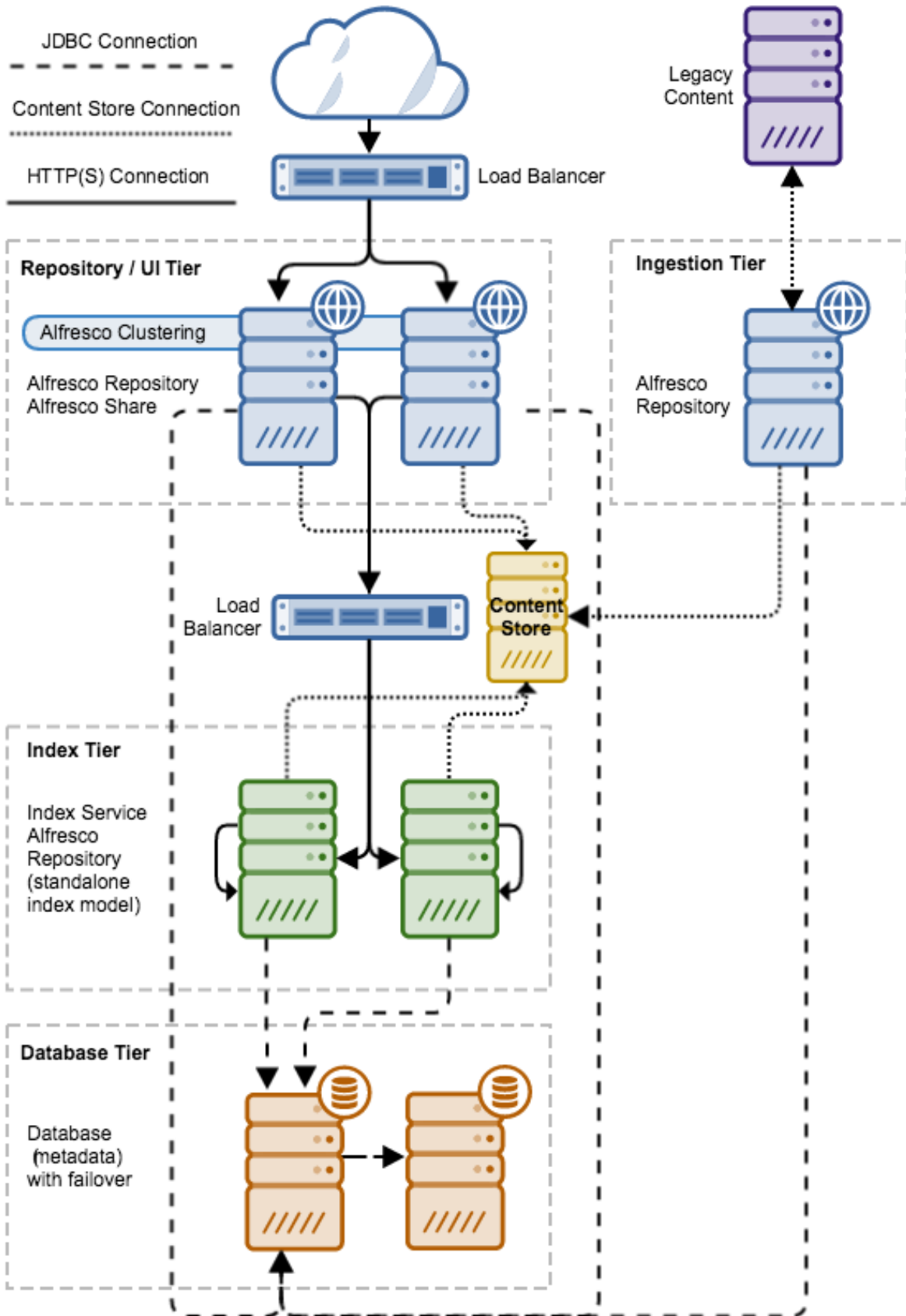
See the **Sizing Guide** for additional information on estimating likely loads.

Extended Components

This section describes the additional components that can be applied to Alfresco Content Services architectures.

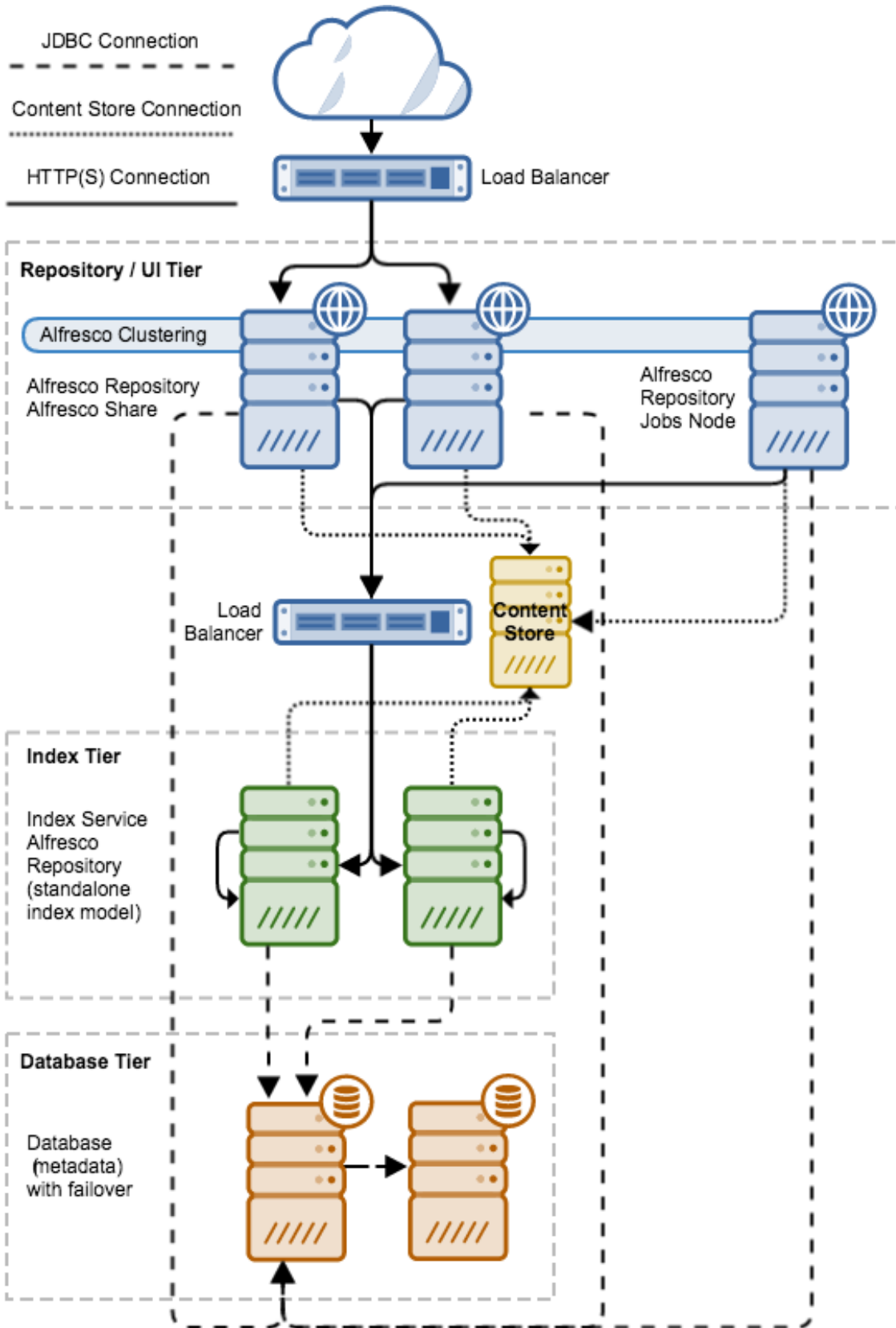
Dedicated Ingestion Server

High volume content ingestion can generate large loads on an Alfresco repository server. If this load is to be ingested into a system that is also servicing user requests, a separate ingestion server can be deployed to handle this load. Example use cases include migrations from legacy systems or repository consolidation. The ingestion server should not be added to the Alfresco cluster, but will write to the same underlying database and content store. The diagram below shows the Alfresco Enterprise reference architecture with the addition of a dedicated ingestion tier, containing one ingestion server.



Dedicated Jobs Server

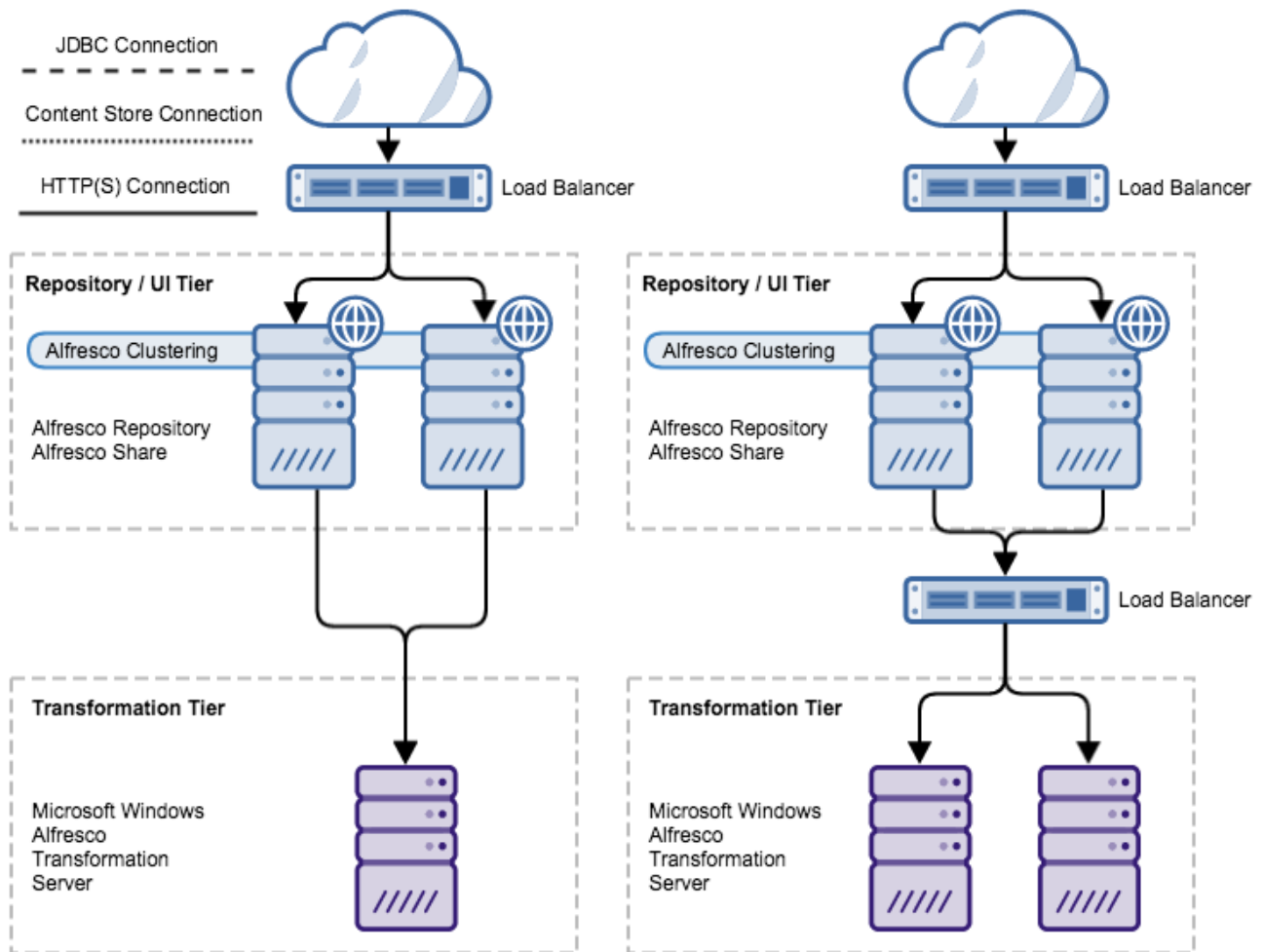
Alfresco runs a series of scheduled jobs to carry out tasks such as repository cleanup, temporary file cleanup, building activities feeds, and so on. On occasion, customers may also need to create custom scheduled jobs based on specific requirements. Depending on the weight and complexity of these jobs, it may be desirable to offload the job execution to a separate server to isolate the load created by those jobs. This can be accomplished by deploying a separate jobs server. If this approach is used, it is important that the jobs server be joined to the Alfresco cluster so that the job lock service can create the locks it needs.



Alfresco Document Transformation Server

The Alfresco Document Transformation Server is an optional component that moves certain types of Microsoft Office document transformations to an external server. The Alfresco repository communicates with this transformation server via HTTP calls. The Document Transformation Server can be deployed singly, or multiple instances can be located behind a load balancer to provide a more scalable transformation solution. Additional information on installing and configuring the Document Transformation Server is available in the Alfresco documentation.

Alfresco Transformation Server	
5.1:	http://docs.alfresco.com/5.1/concepts/transerv-intro.html
5.2:	http://docs.alfresco.com/5.2/concepts/transerv-intro.html

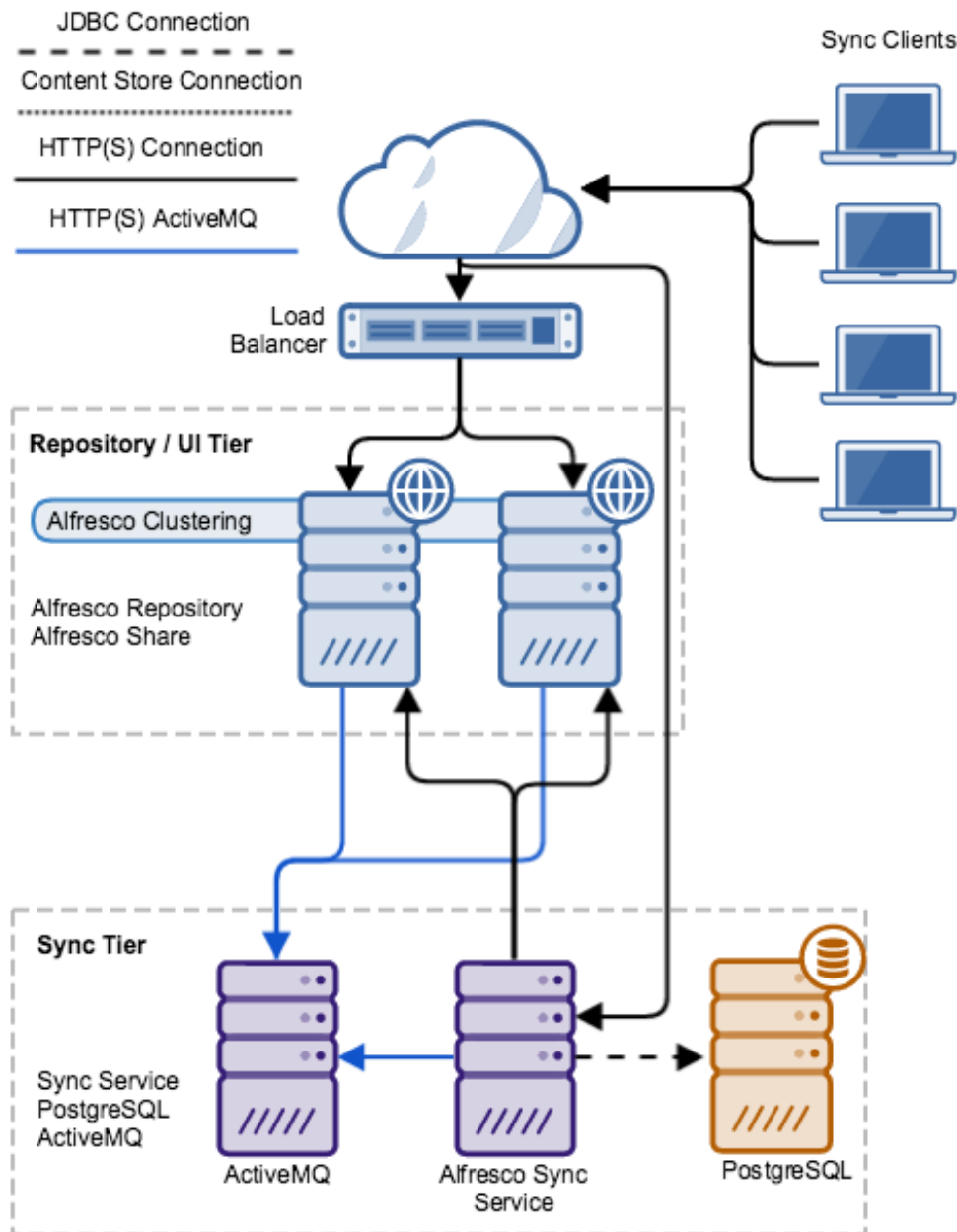


NOTE: Index and DB tiers omitted to simplify diagram

Desktop Synchronization

Alfresco Desktop Sync offers a solution for use cases that require local file sync between an Alfresco repository and multiple clients. This component requires the addition of an ActiveMQ server, the Alfresco Sync Service, and a PostgreSQL database.

Full details around the implementation of Alfresco Desktop Sync can be found in the Alfresco documentation at <http://docs.alfresco.com/desktopsync/concepts/desktop-sync-overview.html>



NOTE: Index and DB tiers omitted to simplify diagram
 NOTE: ActiveMQ, Sync Service and PostgreSQL may be installed on the same server

Alfresco Governance Services (Records Management or RM)

Alfresco Governance Services provides records management functionality for Alfresco. This component does not change the deployment architecture of Alfresco Content Services. It may, however, affect storage requirements, permissions models and other aspects of your usage of the platform. Exactly how it affects storage, index, and database sizing is dependent on how Governance Services is configured and used.

Hardware specifications

This section describes the hardware requirements for each reference architecture.

Note: If you are going to use cloud-based computing resources such as AWS or Rackspace to build a reference architecture, then refer to the **AWS Reference Architecture**, the **Scalability Blueprint**, and the **Sizing Guide** documents for information on specifying resources such as vCPUs, RAM, and storage.

Starter

The following table specifies the hardware requirements for building the **Starter** reference architecture.

Server	CPU	RAM (GiB)	Storage	Notes
Alfresco / Share / Index Server	Quad core, Intel Xeon @2.6Ghz (exact model may vary) or better	16	500GB HDD	Local disk is used for application files, temporary files, and index.
Database Server	Eight core, Intel Xeon @2.6Ghz (exact model may vary) or better	64	Variable	Size and performance of storage dependent on size of repository. See the Sizing Guide for formulas.
Content Store	N/A	N/A	Mounted storage, local, SAN, or NAS.	Size and performance depends on size of repository. See the Sizing Guide for formulas

There is no High Availability (HA) system in place in the **Starter** architecture.

Business

The following table specifies the hardware requirements for building the **Business** reference architecture.

Server	CPU	RAM (GiB)	Storage	Notes
Alfresco / Share / Index Server	Quad core, Intel Xeon @2.6Ghz (exact model may vary) or better	16	500GB HDD	Local disk is used for application files, temporary files, and index.
Database Server	Eight core, Intel Xeon @2.6Ghz (exact model may vary) or better	64	Variable	Size and performance of storage dependent on size of repository. See the Sizing Guide for formulas.
Content Store	N/A	N/A	Mounted storage, local, SAN, or NAS.	Size and performance depends on size of repository. See the Sizing Guide for formulas.

The **Business** architecture has a duplicate system in place for High Availability (HA).

Enterprise

The following table specifies the hardware requirements for building the **Enterprise** reference architecture.

Server	CPU	RAM (GiB)	Storage	Notes
Alfresco / Share Server (2)	Quad core, Intel Xeon @2.6Ghz (exact model may vary) or better	16	500GB HDD	Local disk is used for application files and temporary files
Database Server	Sixteen core, Intel Xeon @2.6Ghz (exact model may vary) or better	128	Variable	Size and performance of storage dependent on size of repository. See the Sizing Guide for formulas.
Index Server (1)	Quad core, Intel Xeon @2.6Ghz (exact model may vary) or better	32	500GB SSD	Index servers require high disk performance. See the Sizing Guide for detailed index sizing formulas.
Content Store Server	N/A	N/A	Mounted storage, local, SAN, or NAS.	Size and performance depends on size of repository. See the Sizing Guide for formulas.

The **Enterprise** architecture has a duplicate system in place for High Availability (HA).

Enterprise - Scaled

The following table specifies the hardware requirements for building the **Enterprise - Scaled** reference architecture.

Server	CPU	RAM (GiB)	Storage	Notes
Alfresco / Share Server (4)	Quad core, Intel Xeon @2.6Ghz (exact model may vary) or better	16	500GB HDD	Local disk is used for application files and temporary files.
Database Server	Sixteen core, Intel Xeon @2.6Ghz (exact model may vary) or better	128	Variable	Size and performance of storage dependent on size of repository. See the Sizing Guide formulas.
Index Server (2)	Quad core, Intel Xeon @2.6Ghz (exact model may vary) or better	32	500GB SSD	Index servers require high disk performance. See the Sizing Guide for detailed index sizing formulas.
Content Store Server	N/A	N/A	Mounted storage, local, SAN, or NAS.	Size and performance depends on size of repository. See the Sizing Guide for formulas.

The **Enterprise x5** architecture has a duplicate system in place for High Availability (HA).

Software specifications

This section lists the specific software required to build a reference architecture.

Supported software stacks

The software stacks define several key pieces of system software used in a reference architecture. The reference architecture is not prescriptive about the vendors used to provide this software, only that it is listed as supported in Alfresco's Supported Platforms document.

1. Operating system
2. Database server
3. Application server

To allow for flexibility, a reference architecture allows you to choose from several supported stacks, depending on your requirements and preferences. For example, if you prefer using enterprise-grade commercial database servers and commercial Linux, there is a supported

stack. If you prefer to use Microsoft software there is also a supported stack you can use. If you prefer free open source software there is also a supported stack.

<https://www.alfresco.com/services/subscription/supported-platforms>

Configuration and validation

Ensure that your Alfresco Content Services installation is suitably configured for your deployment. See the following Alfresco documentation for details.

Configuration type	Documentation
Repository	5.1: http://docs.alfresco.com/5.1/concepts/intro-core.html 5.2: http://docs.alfresco.com/5.2/concepts/intro-core.html
Share	5.1: http://docs.alfresco.com/5.1/concepts/share-configuring-intro.html 5.2: http://docs.alfresco.com/5.2/concepts/share-configuring-intro.html
Database	5.1: http://docs.alfresco.com/5.1/concepts/intro-db-setup.html 5.2: http://docs.alfresco.com/5.2/concepts/intro-db-setup.html
Content store	5.1: http://docs.alfresco.com/5.1/concepts/manage-cs-home.html 5.2: http://docs.alfresco.com/5.2/concepts/manage-cs-home.html
Solr	5.1: http://docs.alfresco.com/5.1/concepts/solr-overview.html 5.2 (Solr 4): http://docs.alfresco.com/5.2/concepts/solr-home.html 5.2 (Solr 6): http://docs.alfresco.com/5.2/concepts/solr6-home.html
Authentication and security	5.1: http://docs.alfresco.com/5.1/concepts/auth-intro.html 5.2: http://docs.alfresco.com/5.2/concepts/auth-intro.html

Required Ports

Server Type	Default Port : Purpose
Alfresco Repository	21 - FTP (if FTP enabled) 137 UDP - NetBIOS Name (if CIFS enabled) 138 UDP - NetBIOS Datagram (if CIFS enabled) 139 - NetBIOS Session (if CIFS enabled) 445 - SMB (if CIFS enabled) 5701 - Hazelcast (for repository clustering) 8080 - HTTP 8100 - OpenOffice (local for transformations) 8443 - HTTPS 50500 - JMX
Alfresco Share	5901 - Hazelcast (for Share clustering) 8080 - HTTP 8443 - HTTPS 50500 - JMX
Index Server	8080 - HTTP (if search is configured to not use SSL) 8443 - HTTPS (if search is configured to use SSL) 8983 - Solr comms (Solr 6) 5701 - Hazelcast (if Alfresco repository installed locally)
Transformation Server	8080 - HTTP
Database	JDBC port is vendor dependent. See your database documentation for defaults or DBAs if defaults have been changed.
Content Store	Varies based on storage type and how it is attached.

Validation

Alfresco provides an environment validation tool (EVT) that helps ensure the environment into which Alfresco will be deployed is correctly configured, running a supported stack, and meets some minimum performance requirements. It is recommended that the EVT be run on a new environment before it begins accepting production workloads.

Benchmarking

Once you have installed your system, you should benchmark it in order to assess its performance. Once benchmarked, the system can be tuned, and scaled if required.

See the **Sizing Guide** and **Scalability Blueprint** documents for benchmark results.

Tuning

Alfresco Content Services is a Java application. Many of the performance tuning methods you can apply to Alfresco Content Services are as for other Java applications - for example, making sure enough RAM and stack is allocated to the JVM.

Some notes on tuning the JVM to suit Alfresco Content Services can be found in the following documentation:

JVM Tuning
5.1: http://docs.alfresco.com/5.1/concepts/jvm-tuning.html
5.2: http://docs.alfresco.com/5.2/concepts/jvm-tuning.html

Scaling

Once a reference architecture is deployed it will be necessary to benchmark, tune, and scale the deployment. This may occur soon after installation to meet your requirements, but may also happen over time as your requirements grow. Alfresco provides a **Scalability Blueprint** document to assist you with scaling your solution.

There are two possibilities for scaling:

- **Scale up** - upgrade machines (more CPU, more RAM, more cache, bigger HDD)
- **Scale out** - add more machines - more clustering - add dedicated transformation server

The scaling requirements will depend on the customer's specific needs, but would typically based on number of concurrent users, number and types of documents, required transformations, rate of ingestion of new documents and so on.

Scaling out is not linear and it is generally advisable to scale up your nodes initially, prior to scaling out, especially the repository nodes.

A typical scenario that would allow the reference architecture to scale out to an **Enterprise x5** deployment (see the **Deployment size definitions** section earlier in this document) would require increasing the Alfresco/Share nodes from two to four, and increasing the index server nodes from one to two, increasing the number of shards as required. The database server would also need to be scaled up (and out) accordingly.

For further information on scaling out, see the **Scalability Blueprint** document.

Sizing

The reference architecture provided in this document is a vetted starting point for an Alfresco deployment. The hardware recommendations provided will support the size of deployment indicated in each sizing option. Deployments that vary significantly from the reference architecture (such as higher concurrent / total user ratios, extremely large repositories, or metadata complexity beyond typical use cases) should refer to the **Alfresco Sizing Guide** for assistance in determining the required hardware.

Resources

Further information can be found from the following resources:

Resource	Link
Alfresco documentation	http://docs.alfresco.com
Alfresco consultancy services	https://www.alfresco.com/alfresco-consulting
Installing Alfresco	http://docs.alfresco.com/5.1/concepts/master-ch-install.html
AWS Quick Start	https://aws.amazon.com/quickstart/architecture/alfresco-content-services/
Alfresco Scalability Blueprint	https://www.alfresco.com/resources/whitepapers/alfresco-scalability-blueprint
Alfresco Sizing Guide	https://www.alfresco.com/resources/whitepapers/alfresco-one-51-sizing-guide
Clustering for redundancy	http://docs.alfresco.com/5.2/concepts/cluster-scenario-redundancy.html
Clustering for high throughput	http://docs.alfresco.com/5.2/concepts/cluster-scenario-throughput.html