Alfresco One 5.1 On-Premises
Reference Architecture
## Document History

<table>
<thead>
<tr>
<th>Document Version</th>
<th>Description of Change</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0</td>
<td>First version release.</td>
<td>19 January 2017</td>
</tr>
</tbody>
</table>
Table of Contents

EXECUTIVE SUMMARY ........................................................................................................ 1
DOCUMENT PURPOSE ......................................................................................................... 1
INTENDED AUDIENCE ..................................................................................................... 1
ASSUMPTIONS ................................................................................................................... 1
GLOSSARY OF TERMS ....................................................................................................... 2
ALFRESCO ONE OVERVIEW ............................................................................................ 3
ALFRESCO One architecture ............................................................................................... 3
ARCHITECTURE DIAGRAMS FOR THE STANDARD DEPLOYMENTS ......................... 4
DEPLOYMENT SIZE DEFINITIONS .................................................................................... 4
  Starter ............................................................................................................................ 5
  Business ....................................................................................................................... 6
  Enterprise ..................................................................................................................... 7
  Enterprise – Scaled ...................................................................................................... 8
KEY COMPONENTS ............................................................................................................ 9
  Repository and Share servers ...................................................................................... 9
  Database server ........................................................................................................... 9
  Content Store ............................................................................................................. 9
  Index server ............................................................................................................... 9
  Load Balancing ........................................................................................................ 10
HARDWARE SPECIFICATIONS ......................................................................................... 10
STARTER ......................................................................................................................... 10
BUSINESS ....................................................................................................................... 11
ENTERPRISE .................................................................................................................... 11
ENTERPRISE - Scaled .................................................................................................... 12
SOFTWARE SPECIFICATIONS ......................................................................................... 13
SUPPORTED SOFTWARE STACKS ................................................................................... 13
CONFIGURATION AND VALIDATION ............................................................................. 13
VALIDATION .................................................................................................................... 14
BENCHMARKING ............................................................................................................. 14
TUNING ........................................................................................................................... 14
SCALING .......................................................................................................................... 14
SIZING ............................................................................................................................... 15
RESOURCES .................................................................................................................... 16
CONTACT ......................................................................................................................... 16
Executive Summary

This document is a high-level best practices guide for deploying Alfresco One in an on-premises installation. It describes the recommendation for the Alfresco One deployment architectures.

Document purpose

The purpose of this document is to describe reference architectures for Alfresco One 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 proven deployment architecture
- 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 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.

Intended audience

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

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

Assumptions

This document assumes a general knowledge of Alfresco One.

Those implementing the chosen reference architecture will require an in-depth knowledge of Alfresco One 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 clustering is also highly recommended.

This Reference Architecture document details the deployment of Alfresco One 5.1.x and above.

For the specific supported platforms for Alfresco One 5.1.x, see the Supported Platforms page on the Alfresco website:

https://www.alfresco.com/services/subscription/supported-platforms
Executive Summary

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:

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

The main Alfresco glossary, with an extensive list of acronyms and terms, can be found here: http://docs.alfresco.com/5.1/glossary/glossary.html
Alfresco One overview

Alfresco One is an Enterprise Content Management (ECM) platform 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 One 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 synched to the cloud, avoiding the risks associated with uncontrolled consumer file sharing and enabling secure collaboration outside of the firewall.

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

Alfresco One architecture

At the core of the Alfresco One 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 One is a Java application that runs in an application server, such as Tomcat.

There are three main tiers:

- Repo-tier - repository
- Web-tier – Alfresco Share user interface
- Index tier - Solr indexing

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:

http://docs.alfresco.com/5.1/concepts/alfresco-components.html

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

Alfresco One supports 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:

http://docs.alfresco.com/5.1/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.

<table>
<thead>
<tr>
<th></th>
<th>Starter</th>
<th>Business</th>
<th>Enterprise</th>
<th>Enterprise - Scaled</th>
</tr>
</thead>
<tbody>
<tr>
<td>Users</td>
<td>300</td>
<td>500</td>
<td>1000</td>
<td>5000</td>
</tr>
</tbody>
</table>

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.
Starter

Starter Pack
Users: 100
Cores: 4

Clients/Browsers/APIs

Alfresco
Repo/Share/Search
4 cores

Database

Content Store
Architecture diagrams for the standard deployments

**Business**

Business Pack
- Users: 300
- Cores: 4
- HA: 4

Disaster Recovery Environment
- Alfresco CAS/Share/Search 4 cores
- Database
- Content Store(s)

DR Protocol
Enterprise

Architecture diagrams for the standard deployments

Alfresco Cluster

Load Balancer

Alfresco CAS/Share
4 cores

Alfresco CAS/Share
4 cores

Database

Content Store(s)

Index Server (Alfresco+SrH)
Scalable+Sharded

Disaster Recovery Environment

Alfresco CAS/Share

Alfresco CAS/Share

Database

Content Store(s)

IS
Architecture diagrams for the standard deployments

Enterprise – Scaled
Key components

Each reference architecture diagram comprises the key components of an Alfresco One 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 Alfresco repository nodes in the cluster, as well as any deployed index servers.

Index server

The index server (based on Alfresco plus Apache Solr) is responsible for indexing content, processing search requests, and returning search results.

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

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

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.

<table>
<thead>
<tr>
<th>Server</th>
<th>CPU</th>
<th>RAM (GiB)</th>
<th>Storage</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alfresco/Share/Index Server</td>
<td>Quad core, Intel Xeon E5-2676 v3</td>
<td>16</td>
<td>500GB HDD</td>
<td>Local disk is used for application files, temporary files, and index.</td>
</tr>
<tr>
<td>Database Server</td>
<td>Eight core, Intel Xeon E5-2670 v2</td>
<td>64</td>
<td>Variable</td>
<td>Size and performance of storage dependent on size of repository. See the Sizing Guide for formulas.</td>
</tr>
<tr>
<td>Content Store</td>
<td>N/A</td>
<td>N/A</td>
<td>Mounted storage, local, SAN, or NAS.</td>
<td>Size and performance depends on size of repository. See the Sizing Guide for formulas</td>
</tr>
</tbody>
</table>

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.

<table>
<thead>
<tr>
<th>Server</th>
<th>CPU</th>
<th>RAM (GiB)</th>
<th>Storage</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alfresco/Share/Index Server</td>
<td>Quad core, Intel Xeon E5-2676 v3</td>
<td>16</td>
<td>500GB HDD</td>
<td>Local disk is used for application files, temporary files, and index.</td>
</tr>
<tr>
<td>Database Server</td>
<td>Eight core, Intel Xeon E5-2670 v2</td>
<td>64</td>
<td>Variable</td>
<td>Size and performance of storage dependent on size of repository. See the Sizing Guide for formulas.</td>
</tr>
<tr>
<td>Content Store</td>
<td>N/A</td>
<td>N/A</td>
<td>Mounted storage, local, SAN, or NAS.</td>
<td>Size and performance depends on size of repository. See the Sizing Guide for formulas.</td>
</tr>
</tbody>
</table>

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.

<table>
<thead>
<tr>
<th>Server</th>
<th>CPU</th>
<th>RAM (GiB)</th>
<th>Storage</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alfresco/Share Server (2)</td>
<td>Quad core, Intel Xeon E5-2676 v3</td>
<td>16</td>
<td>500GB HDD</td>
<td>Local disk is used for application files and temporary files</td>
</tr>
<tr>
<td>Database Server</td>
<td>Sixteen core, Intel Xeon E5-2670 v2</td>
<td>128</td>
<td>Variable</td>
<td>Size and performance of storage dependent on size of repository. See the Sizing Guide for formulas.</td>
</tr>
<tr>
<td>Index Server (1)</td>
<td>Quad core, Intel Xeon E5-2670 v2</td>
<td>32</td>
<td>500GB SSD</td>
<td>Index servers require high disk performance. See the Sizing Guide for detailed index sizing formulas.</td>
</tr>
<tr>
<td>Content Store Server</td>
<td>N/A</td>
<td>N/A</td>
<td>Mounted storage, local, SAN, or NAS.</td>
<td>Size and performance depends on size of repository. See the Sizing Guide for formulas.</td>
</tr>
</tbody>
</table>

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.

<table>
<thead>
<tr>
<th>Server</th>
<th>CPU</th>
<th>RAM (GiB)</th>
<th>Storage</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alfresco/Share Server (4)</td>
<td>Quad core, Intel Xeon E5-2676 v3</td>
<td>16</td>
<td>500GB HDD</td>
<td>Local disk is used for application files and temporary files.</td>
</tr>
<tr>
<td>Database Server</td>
<td>Sixteen core, Intel Xeon E5-2670 v2</td>
<td>128</td>
<td>Variable</td>
<td>Size and performance of storage dependent on size of repository. See the Sizing Guide formulas.</td>
</tr>
<tr>
<td>Index Server (2)</td>
<td>Quad core, Intel Xeon E5-2670 v2</td>
<td>32</td>
<td>500GB SSD</td>
<td>Index servers require high disk performance. See the Sizing Guide for detailed index sizing formulas.</td>
</tr>
<tr>
<td>Content Store Server</td>
<td>N/A</td>
<td>N/A</td>
<td>Mounted storage, local, SAN, or NAS.</td>
<td>Size and performance depends on size of repository. See the Sizing Guide for formulas.</td>
</tr>
</tbody>
</table>

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 One installation is suitably configured for your deployment. See the following Alfresco documentation for details.

<table>
<thead>
<tr>
<th>Configuration type</th>
<th>Documentation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Repository</td>
<td><a href="http://docs.alfresco.com/5.1/concepts/intro-core.html">http://docs.alfresco.com/5.1/concepts/intro-core.html</a></td>
</tr>
<tr>
<td>Share</td>
<td><a href="http://docs.alfresco.com/5.1/concepts/share-configuring-intro.html">http://docs.alfresco.com/5.1/concepts/share-configuring-intro.html</a></td>
</tr>
<tr>
<td>Database</td>
<td><a href="http://docs.alfresco.com/5.1/concepts/intro-db-setup.html">http://docs.alfresco.com/5.1/concepts/intro-db-setup.html</a></td>
</tr>
<tr>
<td>Content store</td>
<td><a href="http://docs.alfresco.com/5.1/concepts/manage-cs-home.html">http://docs.alfresco.com/5.1/concepts/manage-cs-home.html</a></td>
</tr>
<tr>
<td>Solr</td>
<td><a href="http://docs.alfresco.com/5.1/concepts/solr-overview.html">http://docs.alfresco.com/5.1/concepts/solr-overview.html</a></td>
</tr>
<tr>
<td>Authentication and security</td>
<td><a href="http://docs.alfresco.com/5.1/concepts/auth-intro.html">http://docs.alfresco.com/5.1/concepts/auth-intro.html</a></td>
</tr>
</tbody>
</table>
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 will need to 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 One is a Java application. Many of the performance tuning methods you can apply to Alfresco One 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 One can be found in the following documentation:
http://docs.alfresco.com/5.1/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.

Briefly, 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 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.

See the Scalability Blueprint document (https://www.alfresco.com/resources/whitepapers/alfresco-scalability-blueprint) for further information on scaling out.
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:

<table>
<thead>
<tr>
<th>Resource</th>
<th>Link</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alfresco documentation</td>
<td><a href="http://docs.alfresco.com">http://docs.alfresco.com</a></td>
</tr>
<tr>
<td>Alfresco consultancy services</td>
<td><a href="https://www.alfresco.com/alfresco-consulting">https://www.alfresco.com/alfresco-consulting</a></td>
</tr>
<tr>
<td>Installing Alfresco</td>
<td><a href="http://docs.alfresco.com/5.1/concepts/master-ch-install.html">http://docs.alfresco.com/5.1/concepts/master-ch-install.html</a></td>
</tr>
<tr>
<td>AWS Reference Architecture</td>
<td></td>
</tr>
<tr>
<td>Alfresco Scalability Blueprint</td>
<td><a href="https://www.alfresco.com/resources/whitepapers/alfresco-scalability-blueprint">https://www.alfresco.com/resources/whitepapers/alfresco-scalability-blueprint</a></td>
</tr>
<tr>
<td>Clustering for redundancy</td>
<td><a href="http://docs.alfresco.com/5.1/concepts/cluster-scenario-redundancy.html">http://docs.alfresco.com/5.1/concepts/cluster-scenario-redundancy.html</a></td>
</tr>
<tr>
<td>Clustering for high throughput</td>
<td><a href="http://docs.alfresco.com/5.1/concepts/cluster-scenario-throughput.html">http://docs.alfresco.com/5.1/concepts/cluster-scenario-throughput.html</a></td>
</tr>
</tbody>
</table>

Contact

Details of Alfresco consultancy services can be found at the following page on the Alfresco website:

https://www.alfresco.com/alfresco-consulting