DEPARTMENT OF RURAL DEVELOPMENT & LAND REFORM – NATIONAL GEOMATICS MANAGEMENT SERVICE

CIS Revamp

CRVMP_DDP_0005 Solution Architecture

Authors:	Roger D'Souza / Bibhu Sahoo
Branch/Owner:	National Geomatics Management Service (NGMS)
Document ID:	CRVMP_DDP_0005 Solution Architecture
First Revision Date:	
Current Revision Date:	15/11/2018
Revision:	
Version:	1.0 Final

NOTICE

Please note that this document may contain information which constitutes confidential or proprietary information of the Department of Rural Development and Land Reform (DRDLR). Any such information shall only be used, by the respective parties, for the sole purposes of complying with their obligations within this project and may not be disclosed, copied, reproduced, divulged, published, distributed or circulated in any manner whatsoever or be exploited or used for any commercial purposes whatsoever.

APPROVAL

The signatories hereof, being duly authorised thereto, by their signatures hereto authorise the execution of the work detailed herein or confirm their acceptance of the contents hereof and authorise the implementation/adoption thereof, as the case may be, for and on behalf of the parties represented by them.

Acting Chief Surveyor-General : Ms. Rosalind Mdubeki	Date
Acting Chief Director: Cadastral Spatial Information - Mr. Sibusiso Dlamini	Date

CONTENTS

FIGURES

TABLES

Configuration Management

Document Audience and Distribution

Name	Office	Capacity	E-Mail	Category # Ref to note
Rosalind Mdubeki	NGMS	Acting Chief Surveyor-General	Rosalind.Mdubeki@drdlr.gov.za	AP
Nondwe Monyake	NGMS	Surveyor-General KwaZulu-Natal	Nondwe.Monyake@drdlr.gov.za	I
Julian Williamson	NGMS	Surveyor-General Western Cape	Julian.Williamson@drdlr.gov.za	I
Chris Williams-Wynn	NGMS	Surveyor-General Eastern Cape	Chris.WilliamsWynn@drdlr.gov.za	I
Victor Constable	NGMS	Surveyor-General Mpumalanga	Victor.Constable@drdlr.gov.za	I
Elbe Janse van Rensburg	NGMS	Surveyor-General Gauteng	Elbe.JansevanRensburg@drdlr.gov.za	I
Sibusiso Dlamini	NGMS	Acting Chief Director: CSI	Sibusiso.Dlamini@drdlr.gov.za	AP
Inban Moodley	OCIO	Chief Information Officer	Inban.Moodley@drdlr.gov.za	RV
Roy Naidoo	OCIO	Director: Service Management	Roy.Naidoo@drdlr.gov.za	I
Phuti Lekgau	OCIO	Director: Solution Development	Phuti.Lekgau@drdlr.gov.za	I
Boitumelo Sephoti	Internal Audit	Director: Internal Audit	Boitumelo.Sephoti@drdlr.gov.za	RV
Phindi Nkosi	NGMS	Deputy Director: CSI	Phindile.Nkosi@drdlr.gov.za	RV, RC,C
Gole Netshitshivhe	NGMS	Director Technical Services: SG-LMP	gole.netshishivhe@drdlr.gov.za	I
Piet Swart	NGMS	Director Technical Services: SG-WC	Pieter.Swart@drdlr.gov.za	RV
Elias Sithole	NGMS	Deputy Director: Quality Assurance	Elias.Sithole@drdlr.gov.za	RV
Babalwa Lujabe	OCIO	Project Manager	Babalwa.Lujabe@drdlr.gov.za	RV, C
Nonceba Jali	NGMS	Director Technical Services: SG-KZN	Nonceba.Jali@drdlr.gov.za	RV
Moagi Mohohlo	NGMS	Director Technical Services: SG-NW	Moagi.Mohohlo@drdlr.gov.za	I
Prince Mashele	NGMS	Spatial Data Manager	Prince.mashele@drdlr.gov.za	RV
Thembela Gazi	NGMS	Director Technical Services: SG-EC	Thembela.Gazi@drdlr.gov.za	RV
Bongani Mtshali	NGMS	Director – Technical Services SG-BFN	Bongani.Mtshali@drdlr.gov.za	I
Khaya Mkoko	NGMS	Control Survey Technician: BFN	Khaya.Mkoko@drdlr.gov.za	I
Nokuphila Mnyandu	NGMS	Director Technical Services: SG-MPU	Nokuphila.Mnyandu@drdlr.gov.za	I
Fani Motimone	NGMS	Control Survey Technician Gr B: MPU	Fani.motimone@drdlr.gov.za	RV
Stuart Mbokane	OCIO	Director Portfolio Management	Stuart.Mbokane@drdlr.gov.za	I
Wonder Modipa	NGMS	Director Technical Services: SG-GP	Wonder.Modipa@drdlr.gov.za	RV
Vijay Mandapati	Data World	Project Manager	vijay@dataworld.co.za	W, C
Parusha Malepfane	Data World	Project Administrator	parusha@dataworld.co.za	I
Phila Dlamini	Data World	Senior Business Analyst	Phila@dataworld.co.za	I
Rajesh Nooka	Data World	Business Development Manager	Rajesh@dataworld.co.za	RV
Hessie Molotsi	NGMS	Control Network Controller	Hessie.Molotsi@drdlr.gov.za	RV
Edith Mudugwe	NGMS	I T - Limpopo	Edith.Mudugwe@drdlr.gov.za	RV

Manuel Malapane	NGMS	Cadastral Officer	Manuel.Malapane@drdlr.gov.za	RV
Brandon Naguran	NGMS	IT - KZN	Brandon.Naguran@drdlr.gov.za	RV
EAPhillips	NGMS	STC Gr B	EAPhillips@drdlr.gov.za	RV
Miliswa.Kula	NGMS	IT -EC	Miliswa.Kula@drdlr.gov.za	RV
Coenraad.mouton	NGMS	Control Network Controller	coenraad.mouton@drdlr.gov.za	RV
NOMPUMELELO.DHLAMINI	NGMS	IT - MPU	NOMPUMELELO.DHLAMINI@drdlr.gov.za	RV
Nompumelelo.nkosi	NGMS	IT -MPU	Nompumelelo.nkosi@drdlr.gov.za	RV
Silungile Mthembu	NGMS	STC-GR A	Silungile.Mthembu@drdlr.gov.za	RV
Anastasia makgoba	NGMS	IT -Gauteng	anastasia.makgoba@drdlr.gov.za	RV
Faizel .khan	NGMS	Control Network	faizel.khan@drdlr.gov.za	RV

#NOTE

Categories of level of responsibility as determined below:

W	Document Author	Author of the document
RV	Review	Reviewer of the document (Could form one of the signatories)
C	Must be consulted	Contributes content
I	Must be informed	Person whom this document needs to be distributed to
RC	Recommender	Recommends approval of the document (One of the signatories)
AP	Approver	Approves document (One of the signatories)

Version History

Date	Author	Version	Changes made
13/11/2018	Roger Dsouza / Bibhu Sahoo / Amit	0.1	First Draft
14/11/2018	Vijay Mandapati	0.2	Implement review changes: Data Migration scope; Cosmetic changes and expatiating on training requirements and quality assurance.
18/11/2018	Roger D'Souza / Bibhu Sahoo	0.3	Added Network Architecture Overview Diagram section 6.3

Acronyms and Descriptions

Abbreviation	Description
CIS	Cadastral Information System
CSG	Chief Survey General
CSGO	Chief Surveyor-General Office
CSI	Cadastral Spatial Information
SDI	Spatial Data Infrastructure
OGC	**Open Geospatial Consortium**
DBA	Database Administrator
DDE	Due Diligence Exercise

Abbreviation	Description
DRDLR	Department of Rural Development and land Reform
FAQ	Frequently Asked Questions
ICT	Information Communication Technology
LDOM	Logical Domain
MB	Megabyte
NGI	National Geo-spatial Information
NGMS	National Geomatics Management Service
OCM	Organisational Change Management
OCR	Optical Character Recognition
OLA	Operation Level Agreement
PID	Project Initiation Document
PLS	Professional Land Surveyor
PMBOK	Project Management Body of Knowledge
POC	Proof of Concept
PQ	Participating Quota
QA	Quality Assurance
QC	Quality Control
SG	Surveyor-General
SGBFN	Surveyor-General Bloemfontein
SGEC	Surveyor-General Eastern Cape
SGGP	Surveyor-General Gauteng
SGKZN	Surveyor-General KwaZulu-Natal
SGLMP	Surveyor-General Limpopo
SGMPU	Surveyor-General Mpumalanga
SGNW	Surveyor-General North-West
SGWC	Surveyor-General Western-Cape
SLA	Service Level Agreement
SMS	Short Message Service
SSL	Secure Socket Layer
TBC	To be Confirmed
UAT	User Acceptance Testing
VM	Virtual Machine
WBS	Work Breakdown Structure

Abbreviation	Description
PMP	Project Management Professional
BPMN	Business Process Modelling Notation
CMIS	Content Management Interoperability Services
HTML	Hyper Text Markup Language
CSS	Cascading Style Sheets
CDN	Content Distribution Network
UI	User Interface
LDAP	Light Weight Directory Access Protocol
CMS	Content Management System
CAS	Central Authentication Service
SSO	Single Sign On
MVC	Model View Controller
ACL	Access Control List
WFS	Web Feature Service
WMS	Web Map Srevice
OGC	Open Geospatial Consortium
CGI	Common Gateway Interface
HTTP	HyperText Transfer Protocol
IoC	Inversion of Control
AOP	Aspect Oriented Programming
SPeL	Spring Expression Language
DAO	Data Access Object
JDBC	Java Database Connectivity
ORM	Object Relation Model
STOMP	Streaming Text Oriented Messaging Protocol
REST	Representational State Transfer
SAML	Security Assertion Markup Language
OAuth	Open Authentication
WSO2	Company that supplies Enterprise Middleware
RDBMS	Relational Database Management System
OEM	Original Equipment Manufacturer
POJO	Plain Old Java Objects
EMC	Company that provides SAN Storage Solutions

Abbreviation	Description
GB	Gigabytes
OS	Operating System
VM	Virtual Machine
RTO	Recovery Time Object
DMZ	Demilitarized Zone
CDN	Content Distribution Network
SITA	State Information Technology Agency, South Africa
IDS	Intrusion Detection System
NIDS	Network Intrusion Detection System
ETL	Extract Transform Load
GML	Geography Markup Language
JPA	Java Persisitence API
JVM	Java Virtual Machine

Terminology and Definitions

Term	Definition
Dirty Data	Data which is not fit for the new data structures and requires changing or formatting to comply with the planned data structures, these include duplicate records, incomplete or outdated data, and the improper parsing of record fields from disparate systems.
PMBOK	The Project Management Body of Knowledge is the entire collection of processes, best practices, terminologies, and guidelines that are accepted as standards within the project management industry.
TOGAF	The Open Group Architecture Framework (TOGAF) is a framework for enterprise architecture that provides an approach for designing, planning, implementing, and governing an enterprise information technology architecture. TOGAF is a high level approach to design. It relies heavily on modularization, standardization, and already existing, proven technologies and products.
PRINCE2	PRINCE2 is short for Projects in Controlled Environments, which is a Project Management Methodology to enable delivery of projects on time, within budget, scope and to the right quality.
Productionised	Implemented on the pre-production area, which has been configured the same as the Live Environment. This environment is also called the “pre-prod”.
Property Registration	
Backend	Backend is an application or program, not directly accessed by the user, which performs a specialized function on behalf of a main processor or software system. Backends in the context of this solutions typically implies a Database, Directory Server or other Web Service offered by the department
Back Office	Back office is the portion of a department made up of administration and support personnel who are not client-facing

Term	Definition
Frontend	
Demilitarized Zone	
Sharding	Sharding is a type of database partitioning that separates very large databases the into smaller, faster, more easily managed parts called data shards
Extranet	
Content Ingestion	Content Ingestion is the process of obtaining and importing data for immediate use or storage in a database.

References

Document Name	Document Description	Author	Version/ID	Status
CRVMP_DDP_0004_Business Requirement Specification	Business Requirement Specifications	Phila Dlamini	0.2	Draft
Training Strategy & Plan	CRVMP_DDP_0001_Training Strategy & Plan	Vijay Mandapati	0.2	Draft
Testing Strategy & Plan	CRVMP_DDP_0002_Testing Strategy & Plan	Vijay Mandapati	0.2	Draft
Knowledge Transfer Strategy & Plan	CRVMP_DDP_0003_Knowledge Transfer Strategy & Plan	Vijay Mandapati	0.1	Draft
Database Design & Model documentation	CRVMP_DDP_0006_Database Design & Model Documentation	Mahesh Ganeshan	0.2	Draft
Project Initiation Document (PID)	CRVMP_PLN_0003 Project Initiation Document-PID _v1.2 Final	Vijay Mandapati	1.2	Final

1. Introduction

The Solution Architecture in the CIS Revamp project serves to provide context and cohesiveness to Application components and controls, particularly for the Development phase. It therefore enables easy understanding and indexing of all the required architecture principles, Software / Hardware requirements applicable for the project.

All aspects crucial for the architecture of the application have been referenced within this document.

1.1 Document Purpose

This Solution Architecture document provides a comprehensive architectural design of the new CIS system, using different architectural views to depict different aspects of the system. It intends to capture and convey the significant architectural decisions which have been made on the system.

1.2 Intended Audience

The intended audience for this Solution Architecture document includes individuals that need to understand the solution architecture of the system, with specific focus on the Application and Technology architecture of the system.

- Developers, testers, analysts, part of the project's technical team.
- Key stakeholders and members of the DRDLR management team that require insight into the architectural approach adopted for the project.
- Staff from DRDLR, in charge of the technical aspects of the system and its future technical support and maintenance.
- Data world's and DRDLR's ICT team, as hosting center, in terms of compliance with their architecture and quality requirements.

2. Project Overview

The Department of Rural Development and Land Reform's (DRDLR) National Geomatics Management Service (NGMS) branch aims to provide efficient cadastral survey management service and to provide cadastral information services in support of land delivery and orderly development.

The NGMS branch makes use of the Cadastral Information System (CIS) to deliver its services. The legacy CIS application portfolio is out-dated and un-scalable as a result it became evident that there was a need to overhaul the legacy information system and provide a centralised, web-based, scalable and integrated platform which will enable the fulfilment of day-to-day Surveyor-General Offices' business processes.

3. Solution Overview

3.1 Solution Description

The CIS Revamp Solution aims to offer the following functionalities and benefits:

a) A web-based application which offers a single entry and has integrated the functionalities of the CSG Website and the SG Office backend systems. The application will be accessible and compatible with the latest smart phones and tablets. Electronic Lodgments will be available on the application and that service will be exposed to Professional Land Surveyors and Architects as a web service. The system will still accept Manual Lodgments which will be lodged via the same web interface. Technologies such as Mutual SSL Authentication, will be used to verify the Lodgments to validate the surveyor who is performing the electronic Lodgments;

b) Professional Land Surveyors (PLS) will be notified of their lodgment statuses at designated milestones e.g. on Lodgments, First Level etc.

c) The Document Number and Land Information searches will be exposed as a Web service. The functionality for the automated provision of images which are less than or equal to 5MB will be exposed as a web service.

d) There will be proper business process management tool for the SG offices. The system will have a workflow component which will be configurable by the offices. The solution will enable data driven user access and business rules. The solution will cater for the business value chain namely
1) Reservations, 2) Lodgements, 3) Examination, 4) Maintenance, 5) Archiving, and 6) Information Provisioning;

e) The solution will be a learning system which is able to keep track and suggest searches to the user, to allow the user to select a search criterion instead of having to retype the search criteria again.

f) The solution will have a spatial viewing component enabling viewing of a figure during capture it will also offer Spatial Data Management tools which enable the viewing, importing, exporting and manipulation of spatial data using a spatial mapping tool, data extraction tool(extract from legacy numeric table to spatial);

g) A comprehensive reporting utility will be developed and accessible at any relevant point by the PLS /Architect and the Back office as per the user profile and rights allocation.

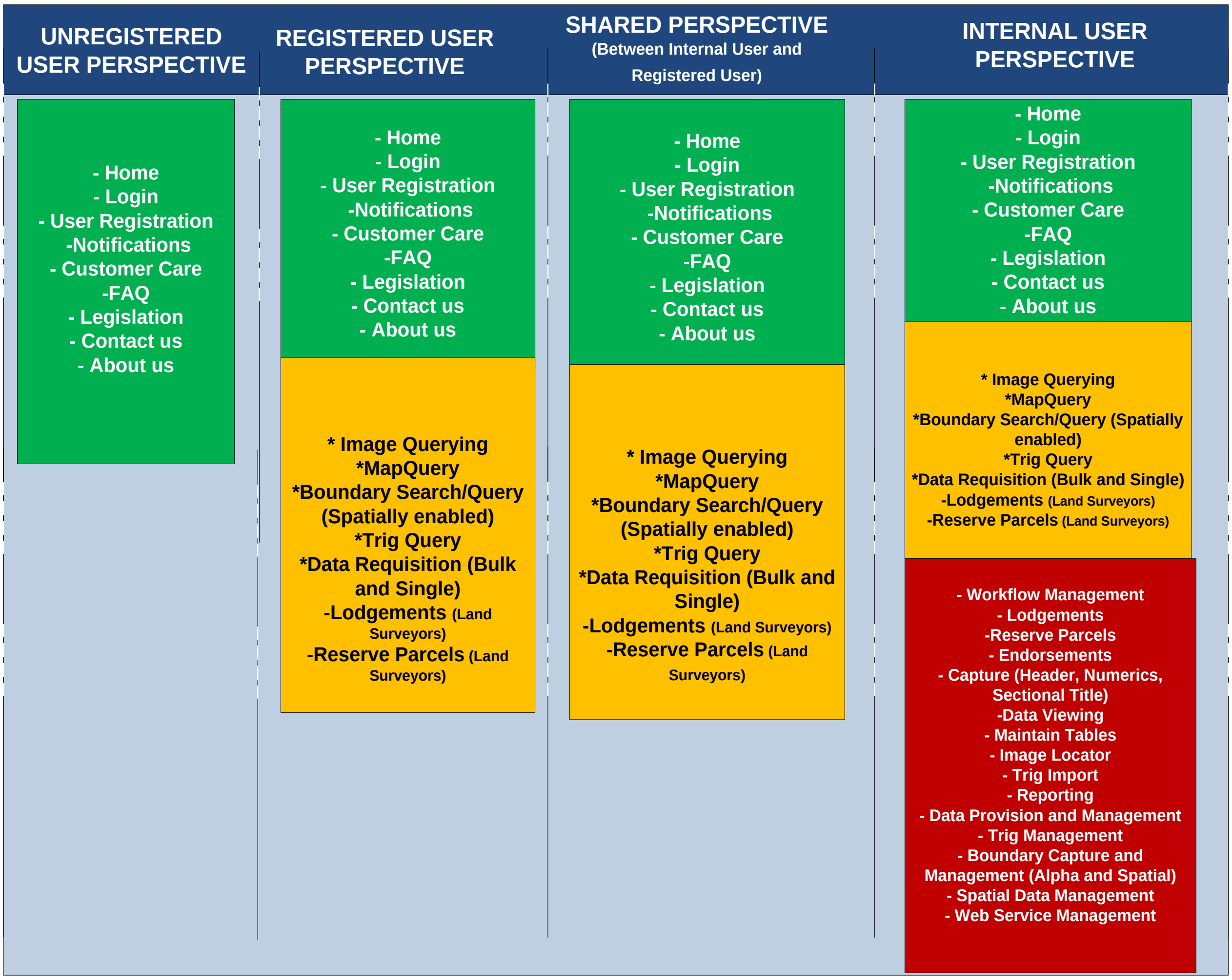
h) The system will have tools catering for area, distance and coordinate conversion and the calculation of Participating Quotas (PQs);

i) There will be an enterprise web portal whereby an internal and the registered user will be able to customize the portal web page according to their user roles;

j) The application will have a web portal which will enable content update by back office administrative staff.

3.2 Solution User Perspectives

The following is an illustration of the expected end-product, and the development thereof will be executed in incremental stages. The sequencing of these iterations will match the perspectives as outlined in the views below in Figure 1.



UNREGISTERED USER PERSPECTIVE

- Home
- Login
- User Registration
-Notifications
- Customer Care
-FAQ
- Legislation
- Contact us
- About us

REGISTERED USER PERSPECTIVE

- Home
- Login
- User Registration
-Notifications
- Customer Care
-FAQ
- Legislation
- Contact us
- About us

* Image Querying
*MapQuery
*Boundary Search/Query (Spatially enabled)
*Trig Query
*Data Requisition (Bulk and Single)
-Lodgements (Land Surveyors)
-Reserve Parcels (Land Surveyors)

SHARED PERSPECTIVE (Between Internal User and Registered User)

- Home
- Login
- User Registration
-Notifications
- Customer Care
-FAQ
- Legislation
- Contact us
- About us

* Image Querying
*MapQuery
*Boundary Search/Query (Spatially enabled)
*Trig Query
*Data Requisition (Bulk and Single)
-Lodgements (Land Surveyors)
-Reserve Parcels (Land Surveyors)

INTERNAL USER PERSPECTIVE

- Home
- Login
- User Registration
-Notifications
- Customer Care
-FAQ
- Legislation
- Contact us
- About us

* Image Querying
*MapQuery
*Boundary Search/Query (Spatially enabled)
*Trig Query
*Data Requisition (Bulk and Single)
-Lodgements (Land Surveyors)
-Reserve Parcels (Land Surveyors)

- Workflow Management
- Lodgements
-Reserve Parcels
- Endorsements
- Capture (Header, Numerics, Sectional Title)
-Data Viewing
- Maintain Tables
- Image Locator
- Trig Import
- Reporting
- Data Provision and Management
- Trig Management
- Boundary Capture and Management (Alpha and Spatial)
- Spatial Data Management
- Web Service Management

Figure 1: User Perspectives in relation to accessible functionality

3.3 Solution Features

The proposed CIS platform provides a full set of functionalities for all types of land administration tasks, such as cadastral operations, property registration, and physical planning. This platform is based solely on a stack of proven open source tools built with modern technologies. The main system features are as follows:

1. Web-based enterprise application based on the open source platform and technologies.
2. Simple user-friendly interface, easy to use, and customizable.
3. Fast configuration of prototype/pilot for client to get fully operational system solution for hands-on early testing and validation.
4. Unified, integrated database for records, textual information, and spatial data that ensures the data integrity and high reliability – in other words, both index and spatial data are stored in a single database.

5. Has a scalable architecture to address the increasing volume of operations in a variety of institutional and organizational structures of the land administration systems in different jurisdictions and countries.
6. Easily customizable data model, business processes, workflows, documents, and reports, with minimal need to recompile application.
7. Computer-guided and workflow-based cadastral and other procedures.
8. Strong security and extensive auditing capabilities that ensure high system stability, data integrity, and reliability of land administration services.
9. Efficient system monitoring, daily case management, and reporting.
10. Transparency of the procedures, controlled access to the data for different levels of customers, including business, owners, agencies, and management via corporate and public portals.
11. Flexible re-configuration of business rules to address potential changes in legislation and regulations in the country of operation.
12. Customization of workflows, rules and forms is available, allowing clients to perform client-side updates to quickly adapt the system to their needs and business environment and to the available technology, making this framework the opposite of vendor-locked “black boxes”.
13. Knowledge transfer from the development team to the client support team, with subsequent maintenance and development of the system with the client’s own development team. Refer document *CRVMP_DDP_0003_Knowledge Transfer Strategy & Plan*.

The automation of cadastral and registration operations, is the core feature, the system provides. This covers all aspects, starting from public areas, serving the clients to distribution of work for the back-office staff, management of the staff loads, provision of various reports on the efficiency of operations, working with maps, monitoring of all operations in the system, and building public and commercial portals. The system can also be positioned as an enterprise integration platform for Land Administration domain.

From a system design perspective, the solution presents a scalable architecture, object- oriented approach, compliance with industrial standards (OGC, BPMN, CMIS, HTML5/CSS3), configurability and customization for specific needs documents and reports, business processes, or organization structure of the land administration institution.

4. Solution Concept Diagram

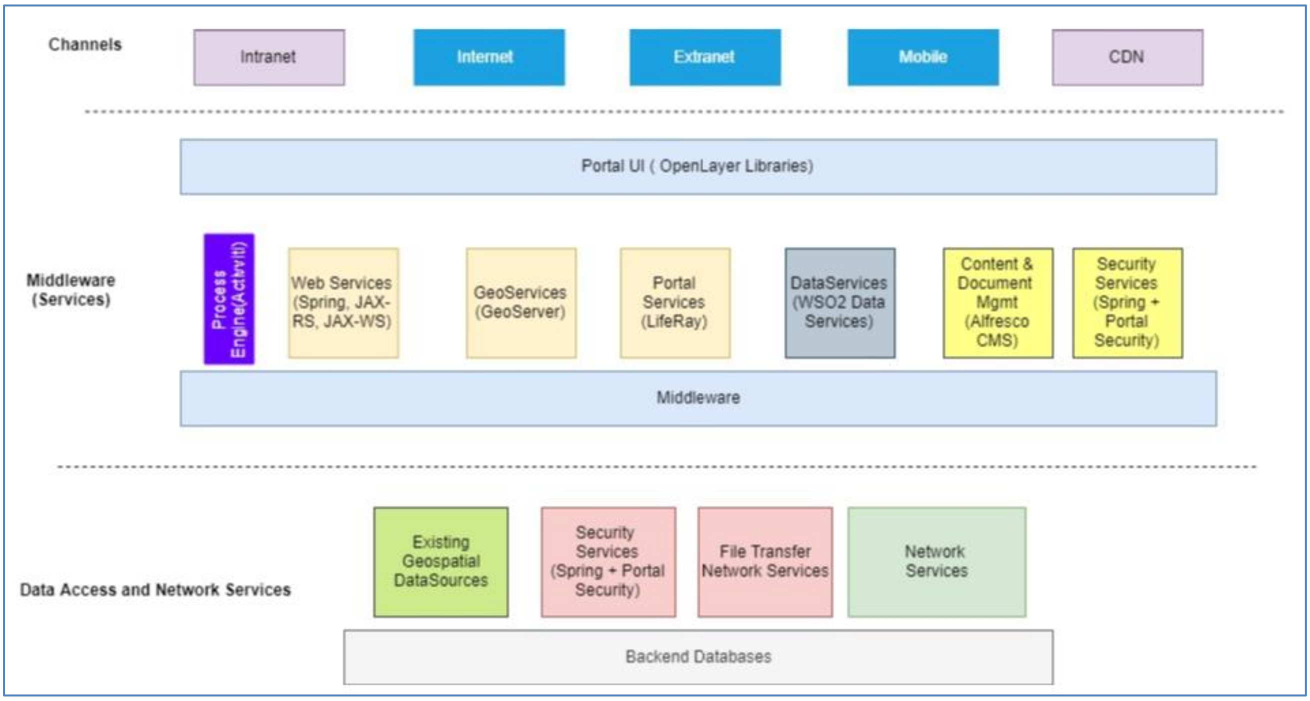

Channels
Intranet
Internet
Extranet
Mobile
CDN
Portal UI (OpenLayer Libraries)
Middleware (Services)
Process Engine(Activiti)
Web Services (Spring, JAX-RS, JAX-WS)
GeoServices (GeoServer)
Portal Services (LifeRay)
DataServices (WSO2 Data Services)
Content & Document Mgmt (Alfresco CMS)
Security Services (Spring + Portal Security)
Middleware
Data Access and Network Services
Existing Geospatial DataSources
Security Services (Spring + Portal Security)
File Transfer Network Services
Network Services
Backend Databases

Figure 2: Solution Concept Diagram

5. Solution Component Overview

5.1 Portal Services

Portal Service provides a user interface framework for quick and easy building of application portals and dashboards for the platform.

5.1.1 Web Portal Overview

The Web portal will provide a frontend for the application and interact with the rest of the CIS platform applications, so users can access specific application features using Portal Service. The portal will effectively provide a single glass pane view of the other application components and services offered by the solution.

The solution will enable non-technical users to make basic configuration changes to the UI using Editor and other components of the Portal Service. More advanced users can edit and extend portals, pages, and widgets. Expert users can use the Widget Editor tool to write scripts to power a portal, and even create rich web applications on the new CIS platform.

Liferay Portal has been chosen as part of solution platform to provide Portal Service. Liferay is an enterprise web platform. It is based on the Java platform and can be extended by adding new applications, customizing existing applications, modifying its behaviour, or and creating new themes. It is used as a headless platform to develop web apps with any technology of the departments choice. It is also used as a web integration layer, leveraging technologies such as portlets to allow several applications to coexist on the same web page.

5.1.2 Web Portal Component Details

Liferay provides a number of functions affecting several components of the entire CIS solution. The decomposition of the portal components are listed below

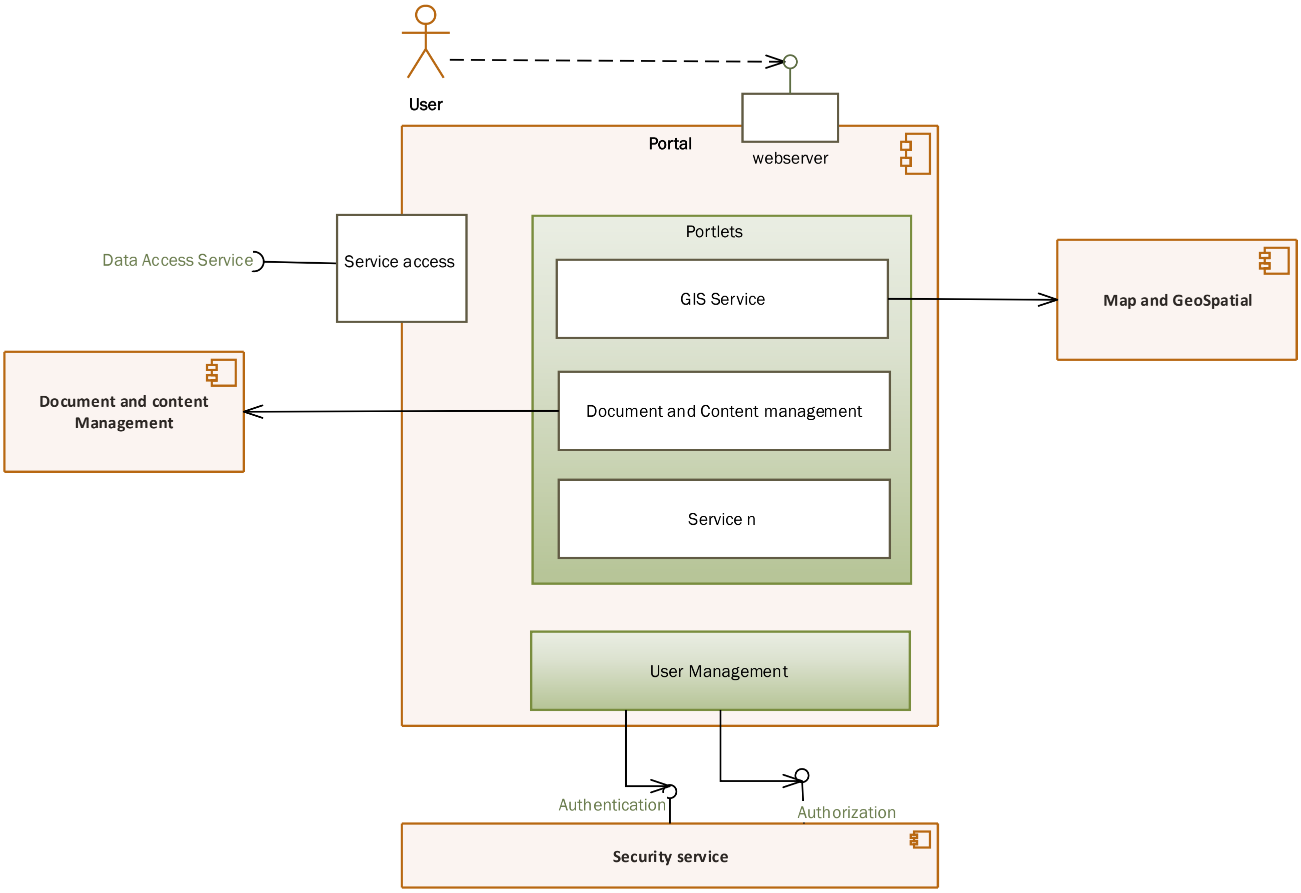


Figure 3: Web Portal Component Diagram

Portlets are pluggable user interface software components that are managed and displayed in a web portal. Portlets produce fragments of mark-up code that are aggregated into a portal page. A portal page is displayed as a collection of non-overlapping portlet windows, where each portlet window displays a portlet. Hence a portlet (or collection of portlets) resembles a web-based application that is hosted in a portal.

The portal provides user management functionality to control access to portal pages and components like portlets. However, the user account information including username, passwords, and others is stored in a separate component using the Lightweight Directory Access Protocol (LDAP) and databases. Thus, it is possible for all components of the solution to commonly access the users' identity information.

Integration:

Liferay Portal provides the ability to fully integrate with Alfresco content management system to take care of Users, Group, Roles and permission synchronization so that Users can see Alfresco as a Liferay CMS and use it through Liferay portlets. These portlets include Alfresco Client and Alfresco Content (More details on these to be covered in Alfresco CMS system).

Following consideration should be taken during Liferay and Alfresco integration

1. Use Alfresco document management as CMS repository of Liferay
2. Replace CMS in Liferay with Alfresco CMS in Liferay

For users to have seamless access across Alfresco and Liferay application Single sign-on access control will be setup. Solution like Central Authentication Service (CAS) are to be used for implementing SSO protocol in the CIS environment.

The following diagram shows full integration of Liferay, Alfresco, LDAP and SSO CAS.

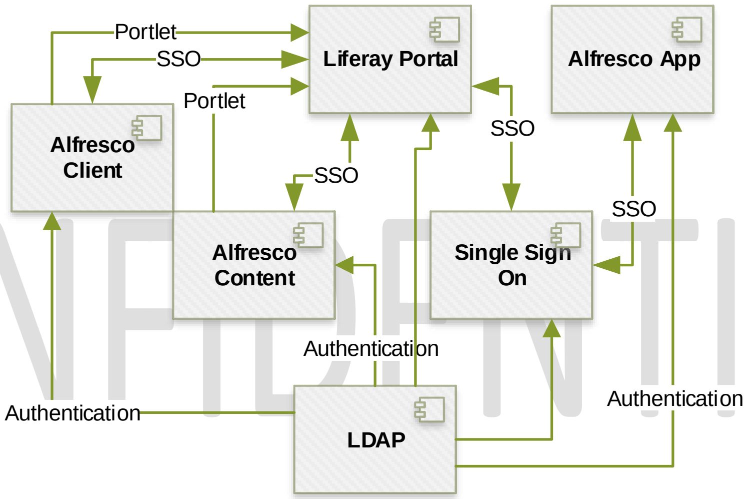


Figure 4: Integration of Liferay Components

For the creation of the user management service, we create a new Liferay module project using Liferay MVCPortlet framework (or other OOTB supported framework like Spring MVC) and service builders to manage Users, Roles and Permission.

The underlying portal architecture is a three-tier architecture-

Frontend

Liferay front-end is based on the implementation ofJSR-268, which is the specification for Java Portlets 2.0. Lifeway's UI uses templating in its views, HTML, CSS and an extension of the YUI JavaScript library, called AUI.

Middleware

The business logic for the portlet UI components will be housed here.

Liferay uses an ACL authorization system and all secured services will do a check against that authorization scheme. The non-secured services are used internally.

The most important operations that the services facilitate relate to:

- Management of portal assets (users, portal organizations, sites, pages, etc.)
- document management (the portal uses an abstraction of a document management system also comes with a default implementation based on Apache Jackrabbit)

Data Persistence Layer

- Liferay's persistence layer uses Hibernate ORM to store data managed by the portal (portal assets). Hibernate offers a lot of flexibility in that it makes Liferay totally independent from the database engine.

5.2 GIS Service

5.2.1 Overview

The GIS Service layer provides the first point of discovery of geospatial information in a map interface.

Below diagram details the service component in detail.

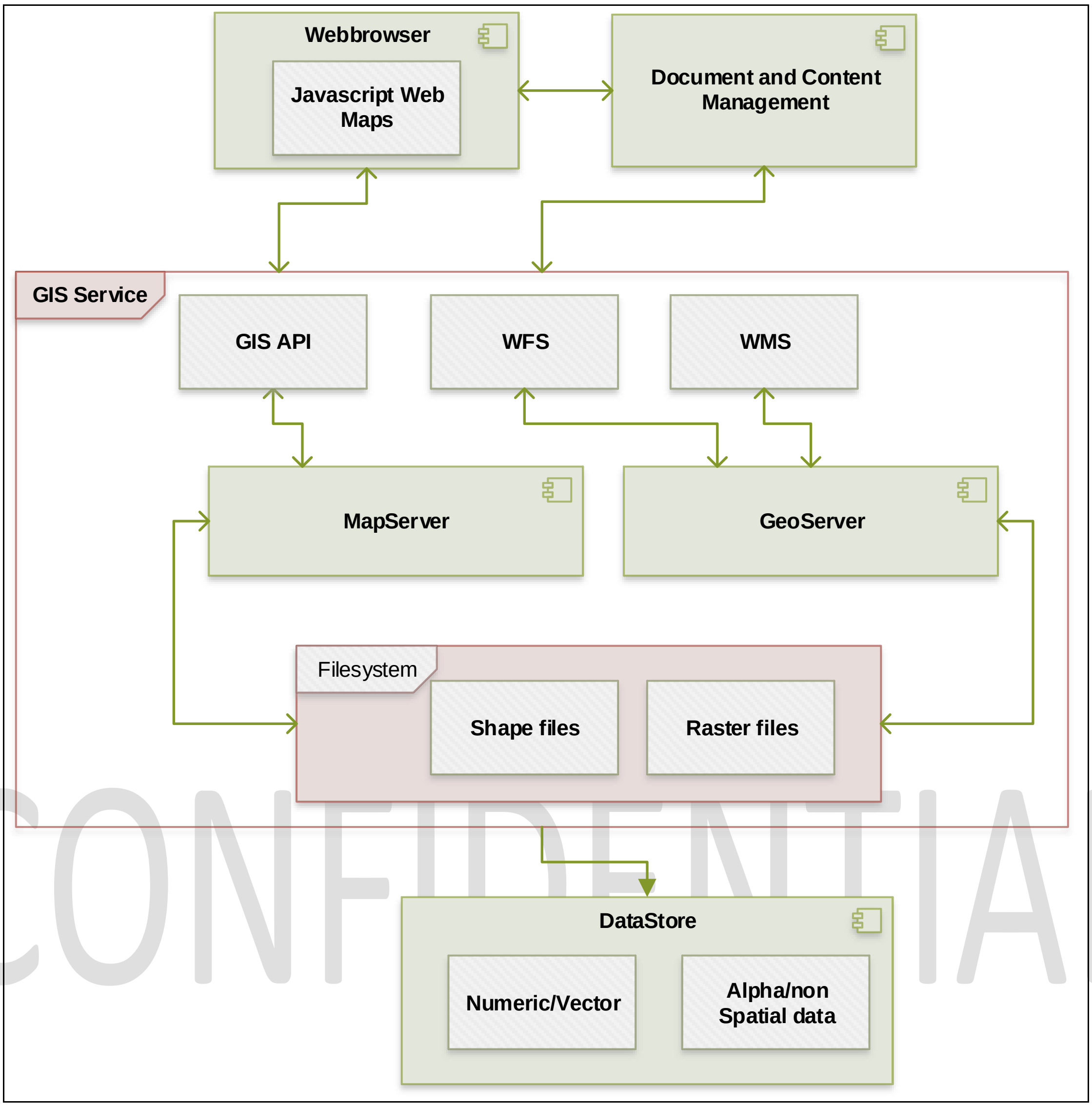


Figure 5: GeoServer Component Diagram

5.2.2 Geoserver

Geoserver is used to publish data from different major data sources using OGC standards and allowing the users to share their data. Unlike Mapserver, Geoserver has no mapping capabilities and is only used for publishing data in an interoperable and standardized way.

5.2.3 Map Server

MapServer is an open source geographic data rendering engine and development environment for building WebGIS applications and sharing data through OGC standards. It can run as a CGI program or via Mapscript which supports several programming languages.

5.2.4 Interface

Below two interfaces will be used profusely to access feature and map data.

WMS: The Web Map Service provides a simple HTTP interface for requesting geo-registered map images from one or more distributed geospatial databases. A WMS request defines the geographic layer(s) and area of interest to be processed. The response to the request is one or more geo-registered map images (returned as JPEG, PNG, etc.) that can be displayed in a browser application. The interface also supports the ability to specify whether the returned images should be transparent so that layers from multiple servers can be combined or not.

WFS: The Web Feature Service defines an interface that allows a client to retrieve and update features of georeferenced data encoded in Geography Markup Language (GML). The main difference between

WMS and WFS is that WFS gives direct access to the geometry and the attributes of a selected

geospatial data, meaning that a user can work with a dataset provided by WFS. In brief, the WFS is the specification to access vector datasets.

5.2.5 Datastore

Oracle RDBMS to be used for storing GIS service data. It stores Numeric/Vector data and Apha/NonSpatial data.

5.2.6 Web Mapping

OpenLayer is used for loading JavaScript maps in client browsers. It supports loading data from many sources: WMS, WFS etc. The JavaScript API allows full control over map from within JavaScript within web page.

5.3 Application Framework Services

The Spring Framework provides a comprehensive programming and configuration model for modern Java-based enterprise applications - on any kind of deployment platform.

The Spring Framework is an application framework and inversion of control container for the Java platform. The framework's core features can be used by any Java application, but there are extensions for building web applications on top of the Java EE (Enterprise Edition) platform.

A key element of Spring is infrastructural support at the application level: Spring focuses on the "plumbing" of enterprise applications so that teams can focus on application-level business logic, without unnecessary ties to specific deployment environments.

The solution will use Spring Framework and other industry standard modules wherever possible. The following Spring modules will be used with other added as deemed necessary to support the business requirements.

- Spring Core - IoC container, Events, Resources, i18n, Validation, Data Binding, Type Conversion, SpEL, AOP.
- Spring Data Access - Transactions, DAO support, JDBC, ORM, Marshalling XML.
- Spring MVC - Spring MVC, WebSocket, SockJS, STOMP messaging.

- Spring Security

Note: For integration purposes and implementation of Enterprise Integration Patterns, the Apache Camel framework will be used.

5.3.1 Spring Boot

Spring Boot makes it easy to create stand-alone, production-grade Spring based Applications. These applications embed Tomcat, Jetty or Undertow directly and provide opinionated 'starter' dependencies to simplify the build configuration. The applications provide production-ready features such as metrics, health checks and externalized configuration.

5.3.2 Web Service Framework

To facilitate integration between the different systems, Web services technologies such as JAX-RS or JAX-WS. Web Services enable loose coupling and facilitate integrations between systems. Web services application will be developed using the Apache CXF is an open source services framework. CXF will be used to build and develop services using frontend programming APIs, like JAX-WS and JAX-RS.

As a general principle, all new services that will be developed as part of the application implementation will use the JAX-RS for REST Services specification and will use the Open API specification (i.e. Swagger) specification.

5.4 Security Services

The security service covers the authentication and authorization part of the application service.

The CIS solution is designed to use an enterprise security framework based on industry standard Spring Security for authentication and authorization.

The portal provides user management functionality to control access to portal pages and components like portlets. However, the user account information including the username, password, and others is stored in a separate component using the Lightweight Directory Access Protocol (LDAP). And the client/customer information will be stored in databases. Thus, it is possible for all components of the solution to commonly access the users’ identity information.

For users to access seamlessly between systems and have an integrated Single Sign On System (SSO) used by the solution. Central Authentication Service (CAS) is a single sign-on protocol for the web. Its purpose is to permit a user to access multiple applications while providing their credentials (such as user_id and password) only once. It also allows web applications to authenticate users without gaining access to a user's security credentials, such as a password. The name CAS in this document also refers to a software package that implements this protocol.

Below diagram shows the architecture of the Security Service.

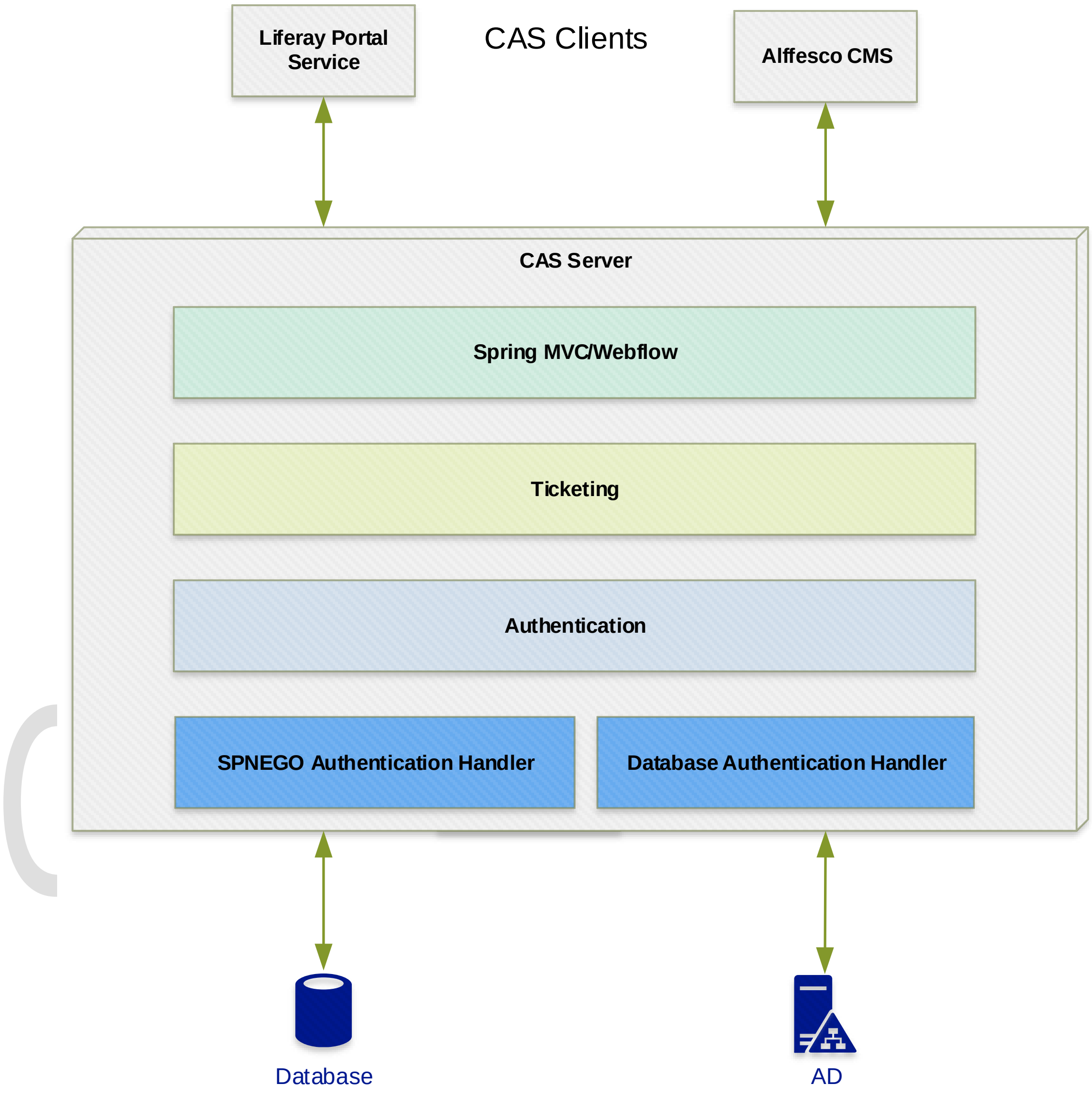


Figure 6: CAS Component Diagram

5.4.1 CAS Server

The CAS server is the core of the security server, whose primary responsibility is to authenticate users and grant access to CAS-enabled services, commonly called CAS clients, by issuing and validating tickets. An SSO session is created when the server issues a ticket-granting ticket (TGT) to the user upon successful login. A service ticket (ST) is issued to a service at the user's request via browser redirects using the TGT as a token. The ST is subsequently validated at the CAS server via back-channel communication. These interactions are described in great detail in the CAS Protocol document.

5.4.2 CAS Clients

A CAS client is any CAS-enabled application (in this case Alfresco and Liferay) that can communicate with the server via a supported protocol. The CAS client software package can also be integrated with spring embedded applications in order to communicate with the CAS server via some authentication protocol (e.g. CAS, SAML, and OAuth). CAS clients supporting a number of software platforms and products have been developed.

Underneath the CAS server is a three layered in subsystems.

- Web (Spring MVC/Spring Webflow)
- Ticketing
- Authentication

The Web tier is the endpoint for communication with all external systems including CAS clients. The Web tier delegates to the ticketing subsystem to generate tickets for CAS client access. The SSO session begins with the issuance of a ticket-granting ticket on successful authentication, thus the ticketing subsystem frequently delegates to the authentication subsystem.

Below diagram shows the flow for establishment of a Single sign on session by the ticketing service of CAS.

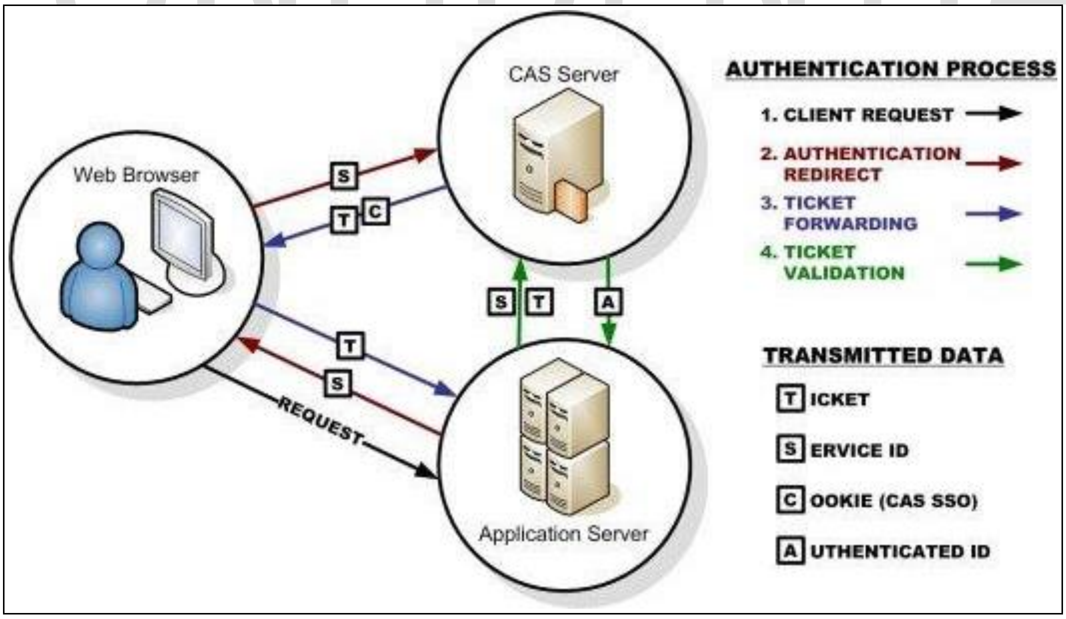


Figure 7: Single Sign on Service Flow

5.5 Data Access Services

5.5.1 Overview

The Data services layer is related to storing, retrieving or acting on data housed in the databases and repository. This service covers data access and integration between the data access layer and the business or service layer. The Data Services component will use the WSO2 Data Services application component as well as Spring Frameworks Data Access modules where JPA/JDBC data access is required via developed interfaces.

These provides three or four interfaces to client systems. These APIs include JAVA Persistence APIs, Hibernate native APIs, Webservice/Rest APIs.

5.5.2 Data Access Service Component Details

System constructed as part of the core solution are generally created as n-tier data systems with multiple APIs. An n-tier architecture means that a client system does not directly access the stored data (the data or persistence tier). Instead, clients use one or more interfaces to access domain objects that are provided by a middleware layer.

In the middleware layer, domain objects are translated into data access objects that are used to retrieve information from the data layer where the actual information is stored (or persisted). The advantage to such as system is that it shields clients from the implementation details of a software system. System owners can make changes to their data tier (change hardware or database type or design) without effecting a change in the API or interface used by clients.

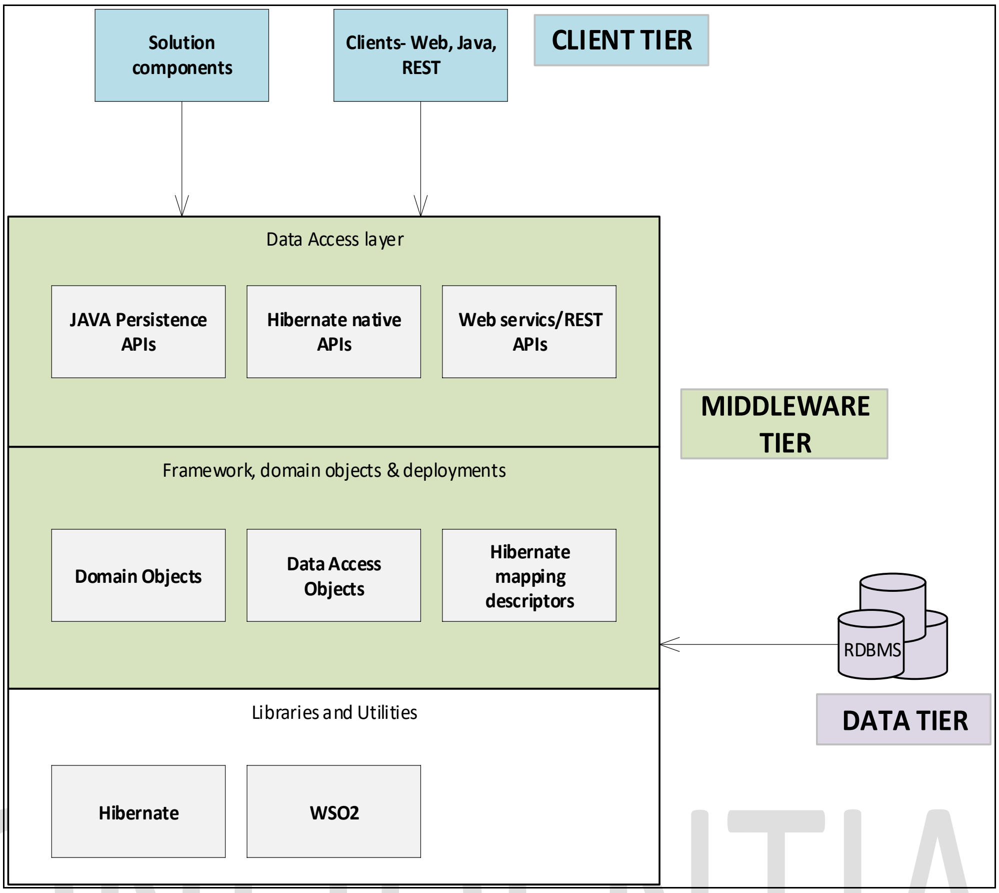


Figure 8: Data Access Service Architecture

Hibernate ORM framework

The implementation of the mapping between objects present in the middleware and API and the underlying data stored in RDMS is done by Hibernate Object/Relational Mapping(ORM) framework. The generation of codes that converts an object query into the Structured Query Language (SQL) used by the underlying RDBMS are also created using the ORM framework.

Based on how the mapping is done Hibernate native APIs or Java Persistence APIs can be chosen for the framework. For Persistent API, we can choose on the type of persistent code to be generated, either Static Methods, Factory Class, DAO or POJO. The decision to which API to select depends on the practice of coding.

5.6 Content Services and Document Management Services

The content services engine is the fundamental building blocks of the application and resides in the middleware tier. The Content Services is an Enterprise Content Management (ECM) platform that will enable the internal end users to manage any type of content: from text documents, to scanned images, photographs, and even large image files. The content services engine works with the process engine described in the section “Business Process Services” to manage the entire lifecycle of the document and related workflows.

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.

5.6.1 Content Services

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 or JBoss Wildfly.

There are three main tiers:

- Repository tier - Repository used to store content and metadata
- Web tier – Alfresco Share or another 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

In the architecture the two main components, repository and Alfresco Share, run on the same node, although there can be multiple clustered nodes.

Database server

- The Alfresco and Share nodes, and index servers, are configured to communicate with a clustered database server hosted on Supported version of Oracle. All metadata for repository content is stored in this database. The database server is clustered to prevent a single point of failure.

Alternatively, advanced features of the selected database such as replication can be used, as required. Configuration of advanced database features is vendor dependent and is out of the scope.

Content Store

- This is the File Content Store which is backed by a filesystem. 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 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) or the Alfresco Index Server (based on Alfresco Content Services plus Apache Solr) is responsible for indexing content, processing search requests, and returning search results.

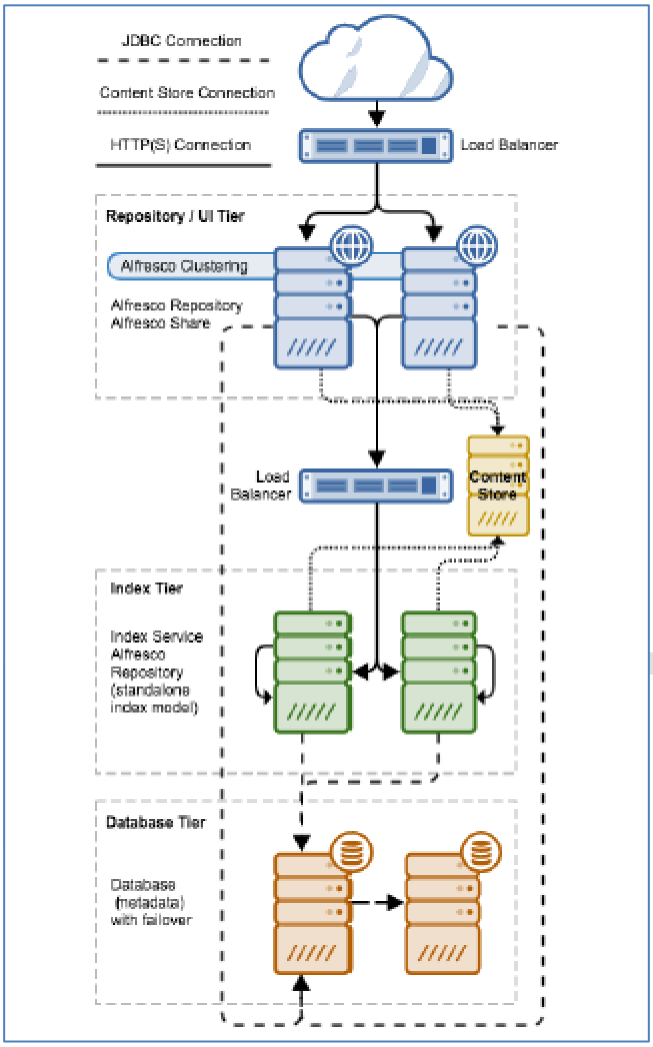


Figure 9: Clustered Alfresco Conent Services

5.6.2 Content Service Component Details

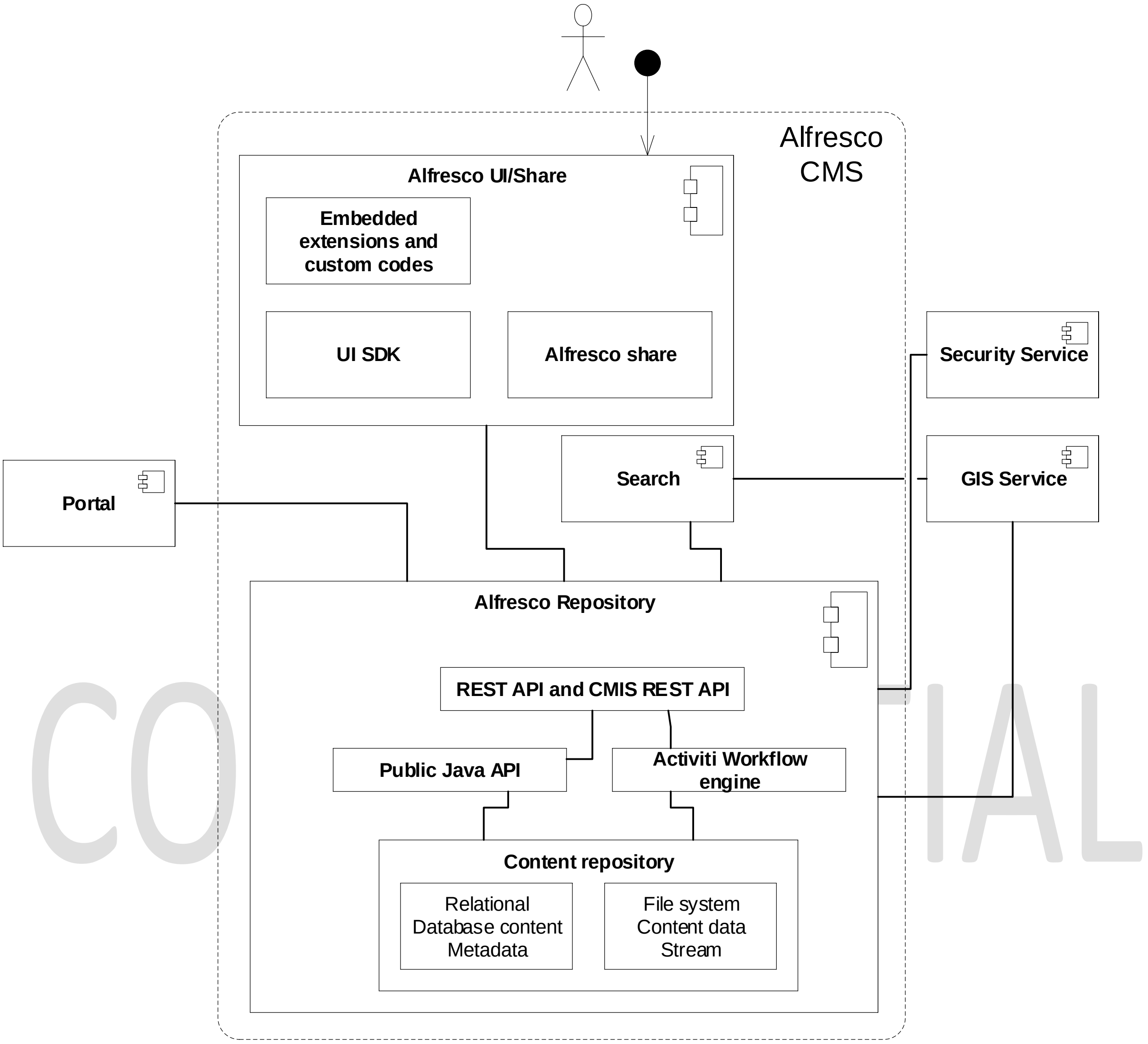


Figure 10: Content and Process Services Component Diagram

The above diagram illustrates the three main components, **Repository**, User Interface (**UI**), and **Search**. These components are implemented as separate web applications:

Repository:

The main component with all the core Enterprise Content Management functionality is called the **Repository**. It provides the repository where content is stored plus all the associated content services.

Note that content metadata is stored in a relational database system and the content itself is stored on the file system. Addition to basic content storage functionality, the repository platform provides a wide range of content-related services. These include core services such as the Node Service, and the Version Service. There are also higher-level services such as Thumbnail Service (for creating thumbnail images and renditions of documents), the Site Service used for creating and managing sites in the Share application, and the Tagging Service, which provides the ability to tag content with keywords

AlfrescoShare:

Alfresco**Share** provides a web client interface (i.e. UI) for the repository. Share will make it easy for users to manage their sites, documents, users and so on.

OCR Service

The CMS service will have **OCR** (Optical Character Recognition) service that allows to transform images (e.g., images scanned from paper documents) into editable text-based document. This service will be built using OCR Java SDK on the Alfresco share UI library.

Search:

The search functionality is implemented on top of Apache Solr and provides the indexing of all content, which enables powerful search functionality. Besides the web clients accessing the Repository via Share there are also mobile clients that will access the content via REST APIs provided by the platform.

Workflow engine:

Alfresco CMS is bundled with a process engine for workflow management based on the Activiti platform. The platform is usually also integrated with a Directory Server (LDAP) to be able to sync users and groups with Alfresco.

Content Model concept:

As described the Alfresco repository provides storage for documents and other content. The content metadata is stored in a relational database, while the content itself is stored directly on the file system. The relationships between content items, and their various properties (metadata) are defined in one or more content models.

Content models can be thought of as describing types of content and the relationships between pieces of content.

For the CIS platform storing of spatial content will be the fundamental function of the CMS system.   It will take the content management solution from being a glorified shared drive to be a system with classified content that can be searched, structured, and processed in many different ways. One of the main differences between a network drive and a Content Management System (CMS) is that the latter provides extra classification features.

Basically, the nodes are classified so it is easier to search for them, so we know how to display them in the user interface, what they can be used for, and so on. The properties that can be used to classify nodes, or if you prefer content, cannot be just any properties, as then the system would not know what each one of these properties represents. Because of this a content management system usually comes preconfigured with properties that can be used to classify content in the repository. These properties are usually organized into two types of groups called Types and Aspects. The main difference between types and aspects is that a node can only have one type applied but it can have multiple aspects applied.

1. Nodes do not live by themselves in the repository. They are related to one another in different ways. How the nodes are related to each other is defined with so-called Associations. A typical association that comes out-of-the-box with Alfresco is the one between a folder node and its child nodes, this type of association is a child-association. So when you start using Alfresco you can immediately organize folders and files into a hierarchy, like with a file system.
2. The Types, Aspects, Properties, and Associations are in turn organized into models that we call Content Models. Alfresco comes with several Content Models out of the box for different things like general folder and file content, workflow-related content, records content, web content, and so on.
3. As part of the CIS solution, the developers/ content administrators will create content models related to specific domains like Cadastral. This will enable the consumers of the system to define new types and aspects forming new custom content models related to CIS. The Meta model defines what syntax we can use when defining our content models.

5.6.2.1 Application Components Deployment Details

1. The Content Services Server and supporting components such as the Share, Search and indexing services will be deployed on a RedHat Linux Enterprise Server.
2. There should be minimum latency between the various application components with the application components being electronically close to the database servers.
3. Supported version of Oracle or OpenJDK will be installed on the server to run the Process engine services which is a Java application.
4. The Repository and Share will be hosted on two servers with solution offering the flexibility to add additional servers if required.
5. The Index services need high performance SSD disks for performing indexing operations with access to the content store. Appropriate sharding will be performed after a proper evaluation of the data storage requirements. Sharding is a type of database partitioning that separates very large databases the into smaller, faster, more easily managed parts called data shards
6. 32 GB Memory and CPU (8Core) requirements will be applied to all the repository and indexing servers for optimum performance. The Content service is a Java application and as such the memory will be allocated to the JVM heap for best performance.
7. The content service engine is a highly transactional server and requires very low database latency of about 2 ms. High DB latency can negatively impact the performance of the server.
8. A dedicated content ingestion server will also be setup to facilitate bulk import of large content into the alfresco repository.

5.7 Process Services (BPMN Process Engine)

5.7.1 Component Overview

Central to the implementation of the workflow engine is the process engine, which is an open source BPMN engine based on the Activiti project.

The Workflow Engine executes most of the symbols defined in the BPMN 2.0 standard. With BPMN, the solution will achieve reliable service orchestration, human task flows, event handling and enable end users to understand diagrams that are technically executable yet easy to understand.

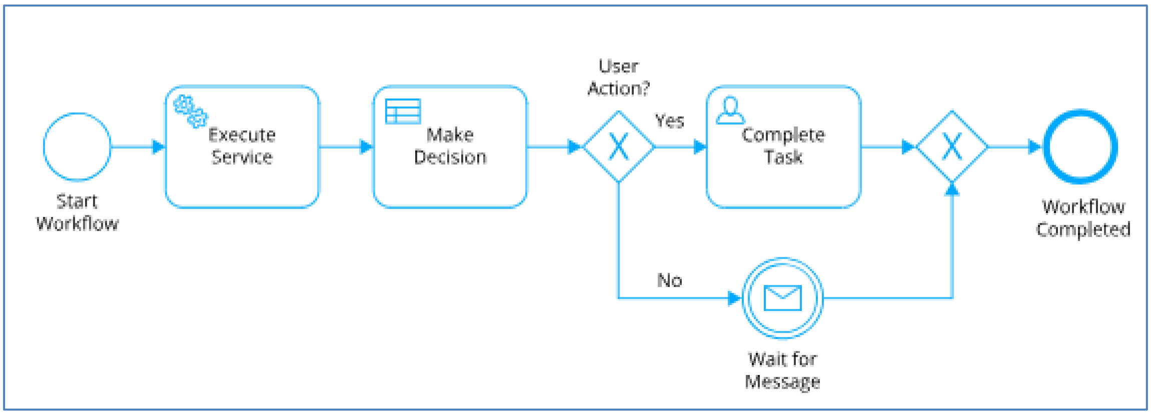


Figure 11: Sample Business Workflow

Activiti is a high performance open-source Java based business process engine with the flexibility and scalability to handle a wide variety of critical processes. The process engine will be used to orchestrate various business workflows in accordance with the BPMN standards. If required, Activiti extensions where necessary to extend the BPMN functionality.

The Process engine will run within the “business process” application server or as a spring boot application. They can run as independent engines or services or integrate to provide a rich suite for business process management. The engine integrates perfectly with spring, rich Java and REST APIs, they are the ideal base for orchestrating human or robotic automation processing. They enable micro service orchestration or software the micro services themselves.

If external integration is required, the business process engine will be modelled with “Service tasks” that invoke implementation classes to invoke external REST interfaces or leverage the Camel Service tasks to facilitate the integration.

5.7.2 Process Services component details

The following components provide details on the various components of the Activiti tool stack, which will be used for development and implementation of the workflow engine.

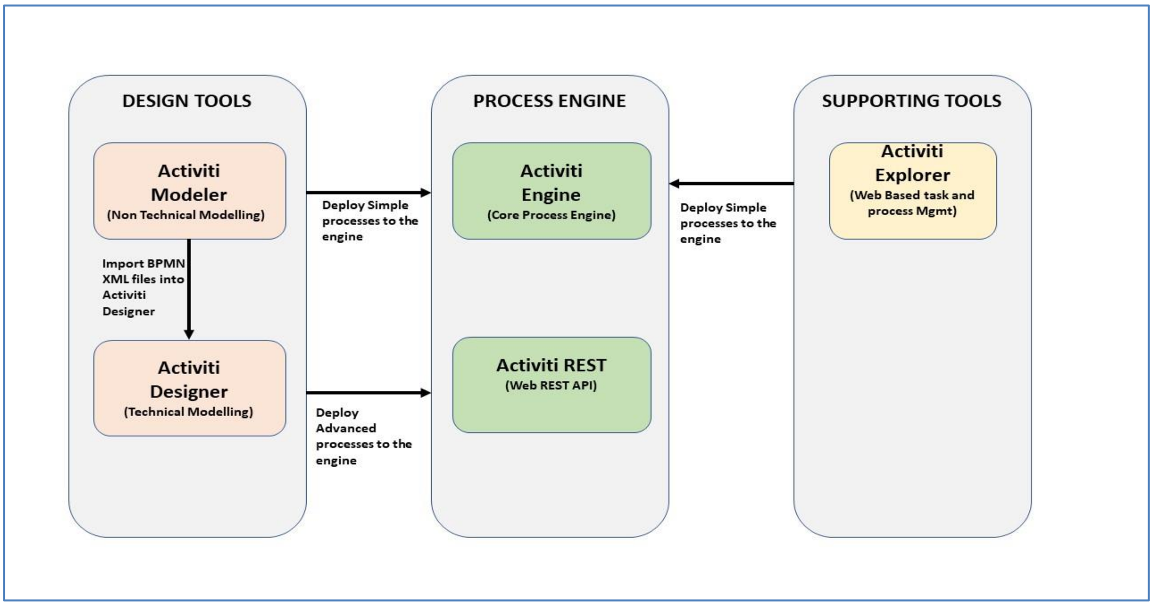

DESIGN TOOLS
Activiti Modeler
(Non Technical Modelling)
Import BPMN XML files into Activiti Designer
Activiti Designer
(Technical Modelling)
Deploy Simple processes to the engine
Deploy Advanced processes to the engine
PROCESS ENGINE
Activiti Engine
(Core Process Engine)
Activiti REST
(Web REST API)
Deploy Simple processes to the engine
SUPPORTING TOOLS
Activiti Explorer
(Web Based task and process Mgmt)

Figure 12: Workflow Engine component details and Process Engine

The following table provides details about the process engine components.

Table 1 : Process Engine Components

Category	Component	Description
Design Tools	Activiti Modeller (Non-Technical Design)	A web-based modelling environment for creating BPMN 2.0-compliant business process diagrams.
	Activiti Designer (Technical Design)	An Eclipse IDE plugin that can be used to design BPMN 2.0-compliant business processes with the addition of Activiti extensions, such as a Java service task and execution listeners.
Process Engine	Activiti Engine	The core component of the Activiti tool stack that performs the process engine functions, such as executing BPMN 2.0 business processes and creating workflow tasks.
	Activiti REST	A web application that provides a REST interface on top of the Activiti Engine. In the default installation the Activiti REST application is the entry point to the Activiti Engine.
Supporting Tools	Activiti Explorer	A web application that can be used for a wide range of functions in conjunction with the Activiti Engine. You can, for example, start new process instances and get a list of tasks assigned to you. In addition, you can perform simple process management tasks, like deploying new processes and retrieving the process instance status.

5.7.3 Process Services Component Architecture

1. The Business Process Server and process engine components will be deployed on a RedHat Linux Enterprise Server. To comply with the availability requirements of the solution the application will be setup on a cluster of servers consisting of a minimum of two servers.
2. The process engine is a Java application and requires a supported version of Oracle or OpenJDK /JRE will be installed on the server to run the Process engine services which is a Java application.

3. The User Interfaces for the process engine will be deployed onto a java application server such as Apache Tomcat or Wildfly.
4. The Activiti process engine is stateless and will be running on two or more Linux nodes and will point to the same database to access process definitions or runtime information housed in the database repository. Enterprise standard load balancer will be setup to route traffic between the two or more nodes in the cluster.
5. The Activiti Process engine requires a database engine to maintain process metadata, process design and running instances of the process definitions in a backend database.
6. To align with the business requirements, it is proposed that the backend for the Activiti database that is deployed on the Oracle database, requires a database engine to maintain metadata and workflow information.
7. Relevant libraries such as Spring, Apache Camel, and JDBC Drivers etc. will be deployed or added to the class path of the application server to facilitate integration or support for the deployed BPMN business processes.
8. The Activiti process engine offers the flexibility to source User, group and role information from the enterprise Active directory to assign tasks to users or groups.
9. When the workflow is initiated, the users will have the ability to view their task queues and “claim/take” tasks from the queue to perform the necessary manual actions. The task queue will be queried using the Task API offered by the Activiti engine. A user (for example, an information manager) can take a task from the task queue and assign the task to an information officer within the business unit/region.

5.7.4 Process Service Integration / Interfaces

5.7.4.1 Web Services Integration

The Activiti process engine exposes a REST interface to start up runtime instances of BPMN processes and pass parameters to the processes. Form parameters on lodgement and request forms can be used to start process workflows. The solution design incorporates the use of this feature of the process engine and offers the flexibility to choose internal forms or external form interfaces such as Angular / Liferay forms to pass data to initiate the workflow processes.

5.7.4.2 User and Role management for Task Assignment (Process Services)

The Process engine offers the flexibility to assign tasks to users and groups that are not explicitly defined in the engine. The proposed solution adopts this approach of looking up the users or groups from the external store and assigning the tasks to the user or group for completion.

Access to the tasks and the assignment of the tasks are managed via Spring Security services module, implemented as part of the application security infrastructure. The users, groups and roles will be defined as part of the data modelling for the User Access Management module.

The Activiti Workflow Engine offers a REST Services interface that can be accessed to start process instances, complete tasks and much more. With the external tasks pattern, the solution will be developed to operate as microservices, completely decoupled from the Workflow Engine, and let them pull the work via REST whenever it suits them.

The workflow engine will be designed to adopt an event driven processing pattern. The event producer will generate events that will trigger the creation of runtime instances of the processes.

The following sample walk-through will illustrate how an event is processed by the system to trigger start-up of workflows.

1. A registered end-user will login to the application and submit a request (i.e. information provisioning request or lodgement request or other types of requests) through the various user interfaces such as the Web UI request management interface.
2. This request will be recorded in the various application tables and initiate call to the Workflow Process engine to initiate the workflow tasks. The completed form attributes will be provided as part of the REST call or alternatively a reference to persisted record will be provided for future lookups.
3. The workflow will evaluate the move through the various automated service tasks such as looking up referential data, notifying user of successful lodgement etc.
4. If a task is assigned to a group, the relevant stakeholders will be notified of the new tasks depending on the notification policy. The task will also show up in the task queue for the user.
5. The user will “take/claim” the task from the queue and complete or assign the task. For example, the information manager can take the task from the queue and assign the task to one of the information officers within their business unit.
6. The workflow process will iterate through all the workflow steps until completion. The workflow process will also provide important analytics data and audit data. This audit data will be stored in the process engine history /audit tables. Alternatively, event listeners will be configured to ship this data to the analytics engine for reporting purposes.

6. Core Solution Views

6.1 Application Communication Diagram

This section covers the interactions between the CIS platform, external application and potential end-users. The diagram represents an integrated system of software, users and web services capable to effectively support CIS activities. Overall, the system can be described by three components:

1. *CIS GeoService Platform:*

The geo platform is an integrated web-based application. Data is managed in a SDI over the entire workflow, from collaborative upload in a web portal, to creation of meta data, the choice of appropriate visualization, the composition of maps and the elaboration through specific modules producing final maps and information.

2. *Web UI Interface and other services:*

Spatial layers and maps managed in the Geo Services Platform can be downloaded in different formats directly from the web interface or can be reused in desktop GIS applications for further investigation by connecting to standard web services. With available plugins like Geo Server explorer, upload of visual data and layers are easy to handle, allowing collaborators and domain experts to better contribute to the update and maintenance of Geo Services platform content.

3. *Third party data repositories*

CIS related workflows need relevant and updated data to be analysed and processed from third party databases or repositories. Using inter-operable OGC complaint web services to exchange spatial information.

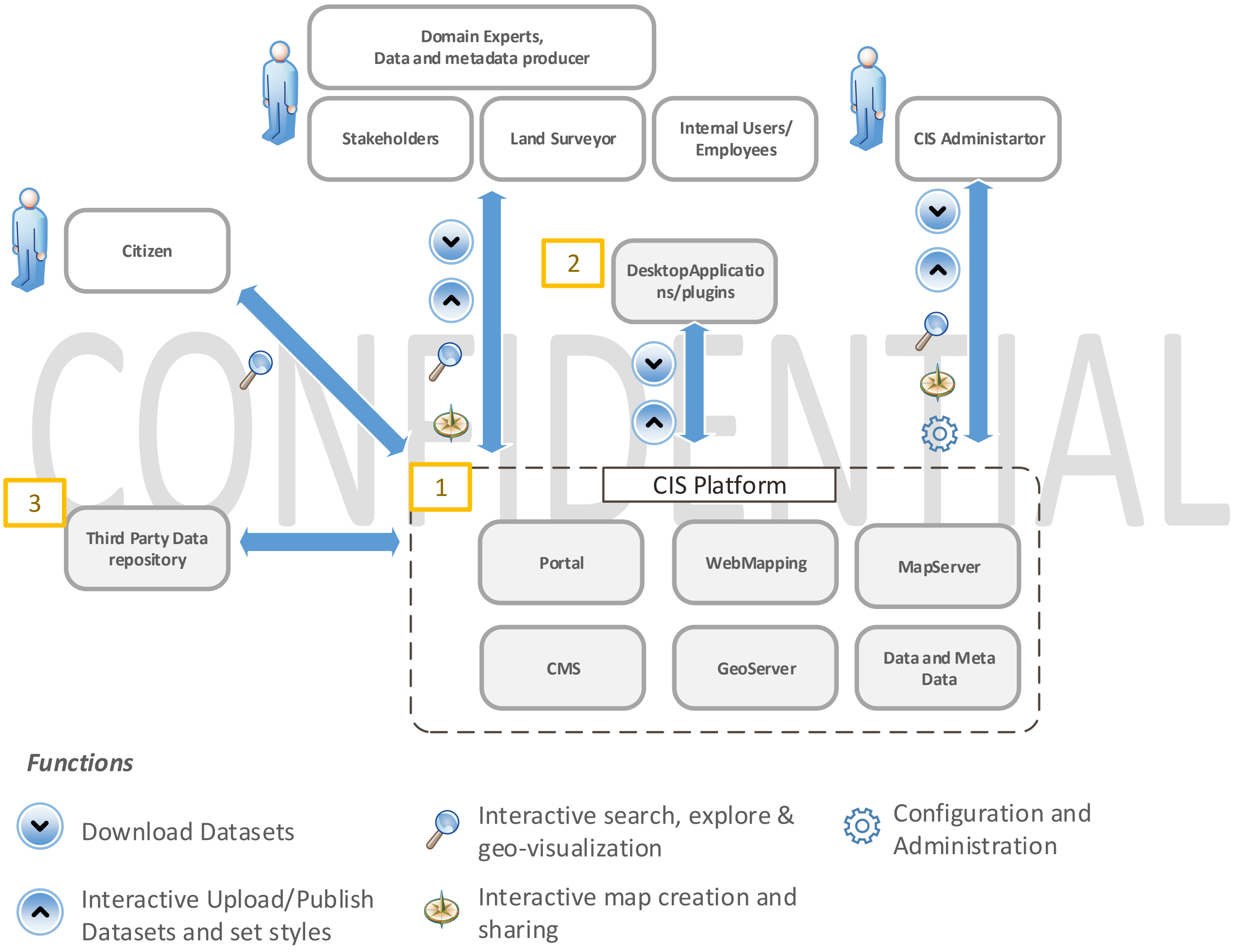

Domain Experts,
Data and metadata producer
Stakeholders
Land Surveyor
Internal Users/ Employees
CIS Administartor
Citizen
2
DesktopApplications/plugins
3
Third Party Data repository
1
CIS Platform
Portal
WebMapping
MapServer
CMS
GeoServer
Data and Meta Data
Functions
Download Datasets
Interactive Upload/Publish Datasets and set styles
Interactive search, explore & geo-visualization
Interactive map creation and sharing
Configuration and Administration

Figure 13: Data Repositories

6.2 Deployment Views

This section covers the solution components and how they are distributed to the physical or virtual resources available for the system. Among others the following aspects have been considered while creating the deployment model:

Security- Components with high security requirements are isolated against other, less sensitive components and thus be run on a separate system node.

Performance- All components are distributed in a way that availability of sufficient system resources is ensured.

Maintainability- To ensure that the possibly complex distributed system can be managed efficiently, components are grouped together in a sensible way.

The UML diagram below covers all the components and how they are distributed to physical resources available for the system. Each nodes may represents cluster of two or more server providing application server clustering at each component level. To enhance security, we recommend a network zoning model comprised of an internet zone, a demilitarized zone (DMZ) and an internal zone. This deployment model also reflects the different levels of control that can be executed by the owners of the data.

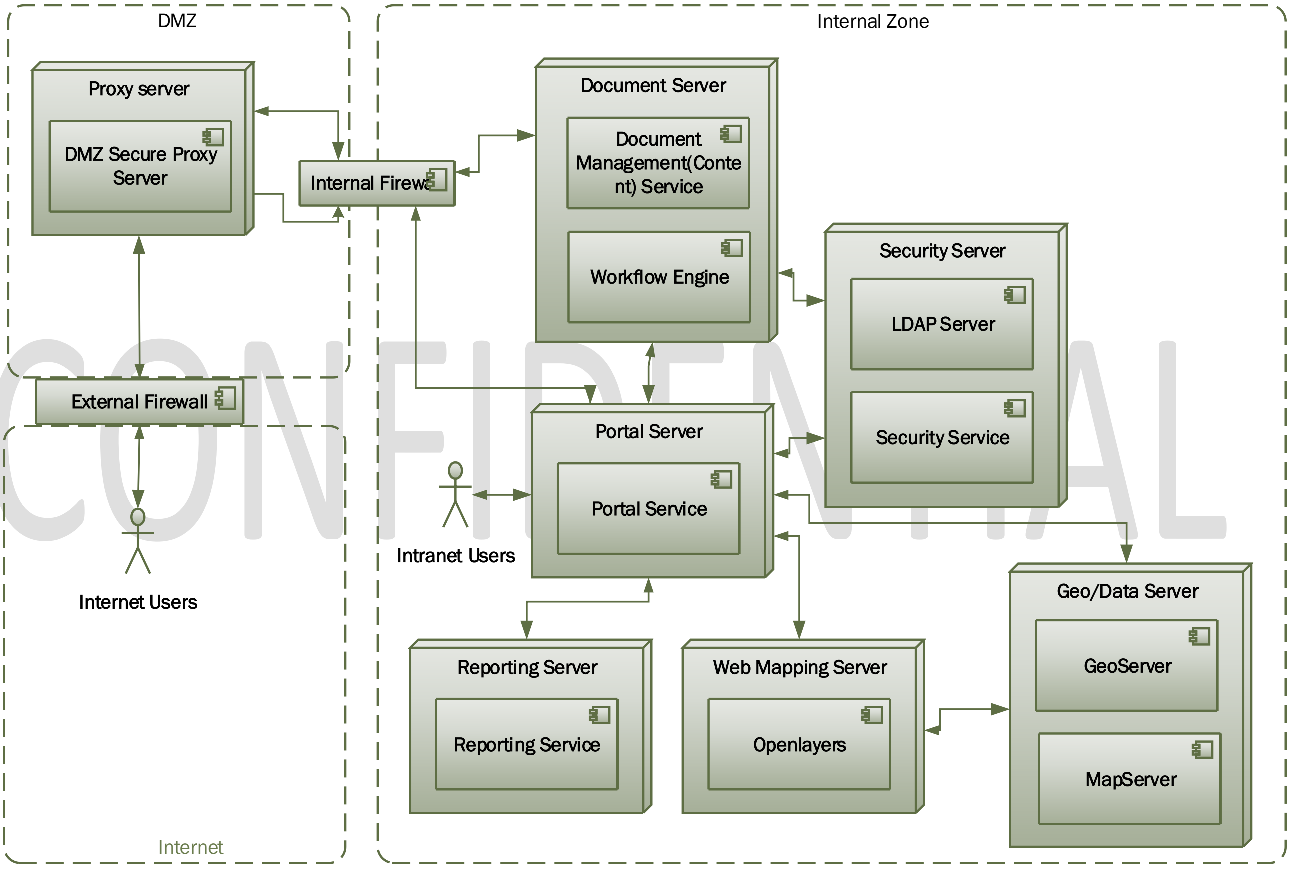


Figure 14: Clustering

The configuration shown in figure 14 is a single DMZ configuration. The DMZ proxy server deployed in DMZ, routes all the requests to external customers. Firewall restrictions make it impossible for users from internet to directly access the services in the secure internal network.

6.3 Network Architecture Overview

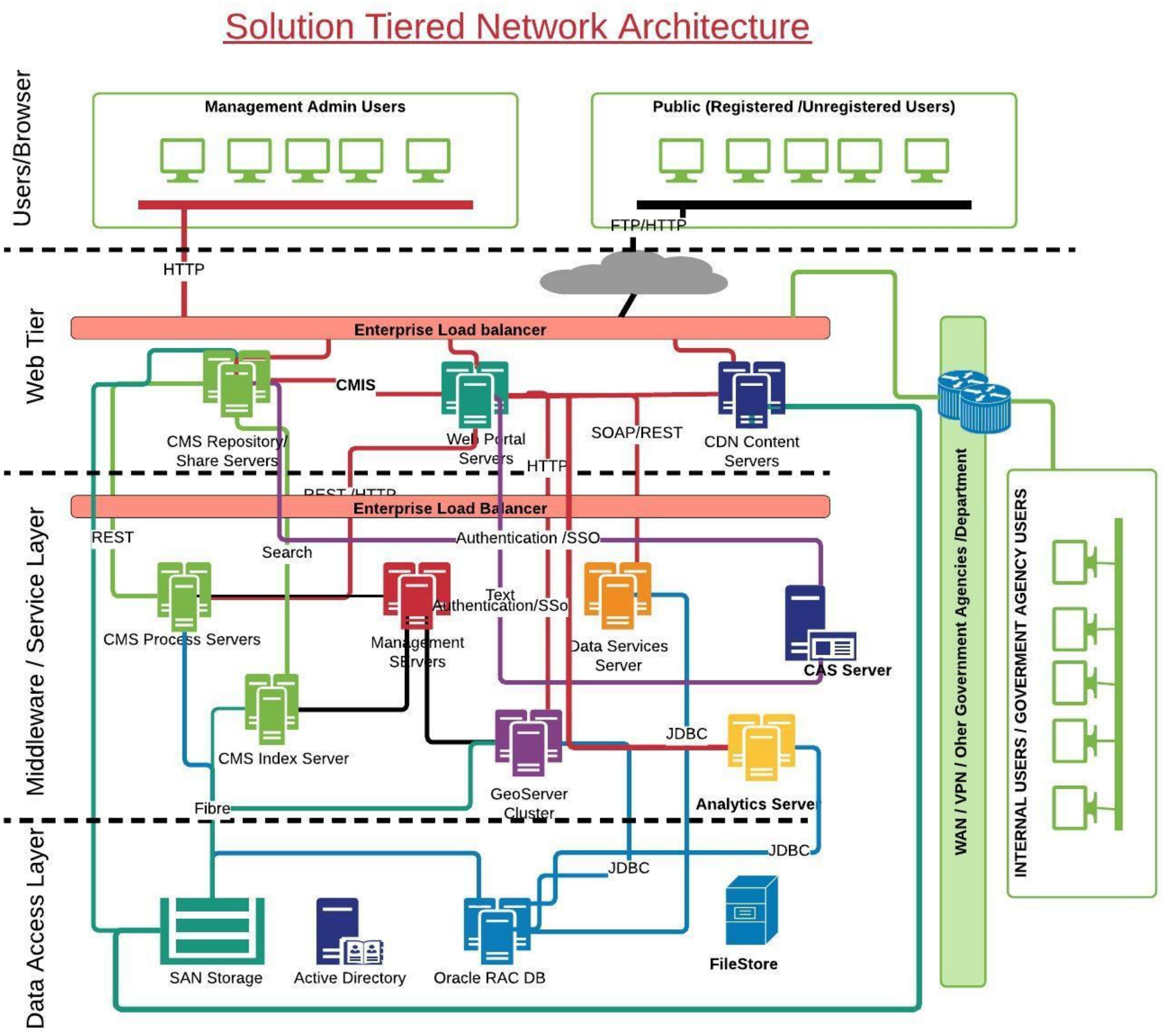


Figure 15: Solution Network Architecture & Integration Specification

The above diagram reflects the network architecture overview of the solution. The hardware specifications for the solution will be added as part of the annexure of system specifications.

1. All components in the solution are clustered with a minimum of two or more servers per application to provide redundancy.
2. Web content and site information will be served by Liferay portal servers. Content designers will design and publish content for the sites using the portal interfaces. The portal server will communicate with Alfresco content services using the CMIS protocol. Backend services developed on the portal server will using the Service Builder Ide.
3. Content such as image files, Shp and other files will be delivered to consumers using the CDN servers which will deliver content using HTTP and FTP.

4. The solution will provide authentication for internal users using the active directory services via LDAP Authentication. All internal users that are consumers of the application will be authenticated using LDAP.
5. There will be 3 instances of Data Services servers that will be linked to legacy data sources via JDBC content and will make the content available via REST/SOAP services.
6. Initially, 2 servers are proposed to run the process services which will drive the workflow process for the solution. A network load balancer will be placed in front of the process services to facilitate interaction between the front end (i.e. web tier) servers and the process engine via REST Services.
7. Alfresco content services integrated with the Process engine services and will enable users to manage the entire lifecycle of the document published into the content server.
8. The index servers are based on Apache Solr and Lucene indexing engine. The servers will index the content that is stored in the content repository to facilitate faster access and offer search capabilities.
9. All events of interest will be forwarded to a cluster of analytics server that will log all the content into the analytics database housed on the Oracle RAC. Reports will be generated against content stored in the analytics server.
10. All the content servers will be connected to the SAN using fibre channel connectivity to enable faster access to the underlying datastores.

6.4 Environment and Location

Table 2: Environment Details

Environment	Description	Notes
DEV	Development Environment	There is no proposed development environment for the project as all the development work will be done offsite at Dataworld premises
TRN	Training Environment	The application will be deployed in the Training environment for users to do hands-on with the application functionality and gain knowledge.
UAT	User acceptance Testing Environment	The application will be deployed in the UAT environment for users to test the application and solution features and functionality. This environment will be available to the various application testing teams to validate if the solution functionality is aligned with the business and functional/non-functional requirements.
PROD	Production Environment	The production environment will be available to the consumers of the system and will house production data
DR	Disaster Recovery environment	To comply with the requirements set forth for the implementation a DR Environment has also been proposed for the solution implementation.

Below diagram shows the reference deployment consideration for production and non-production environment in datacentres.

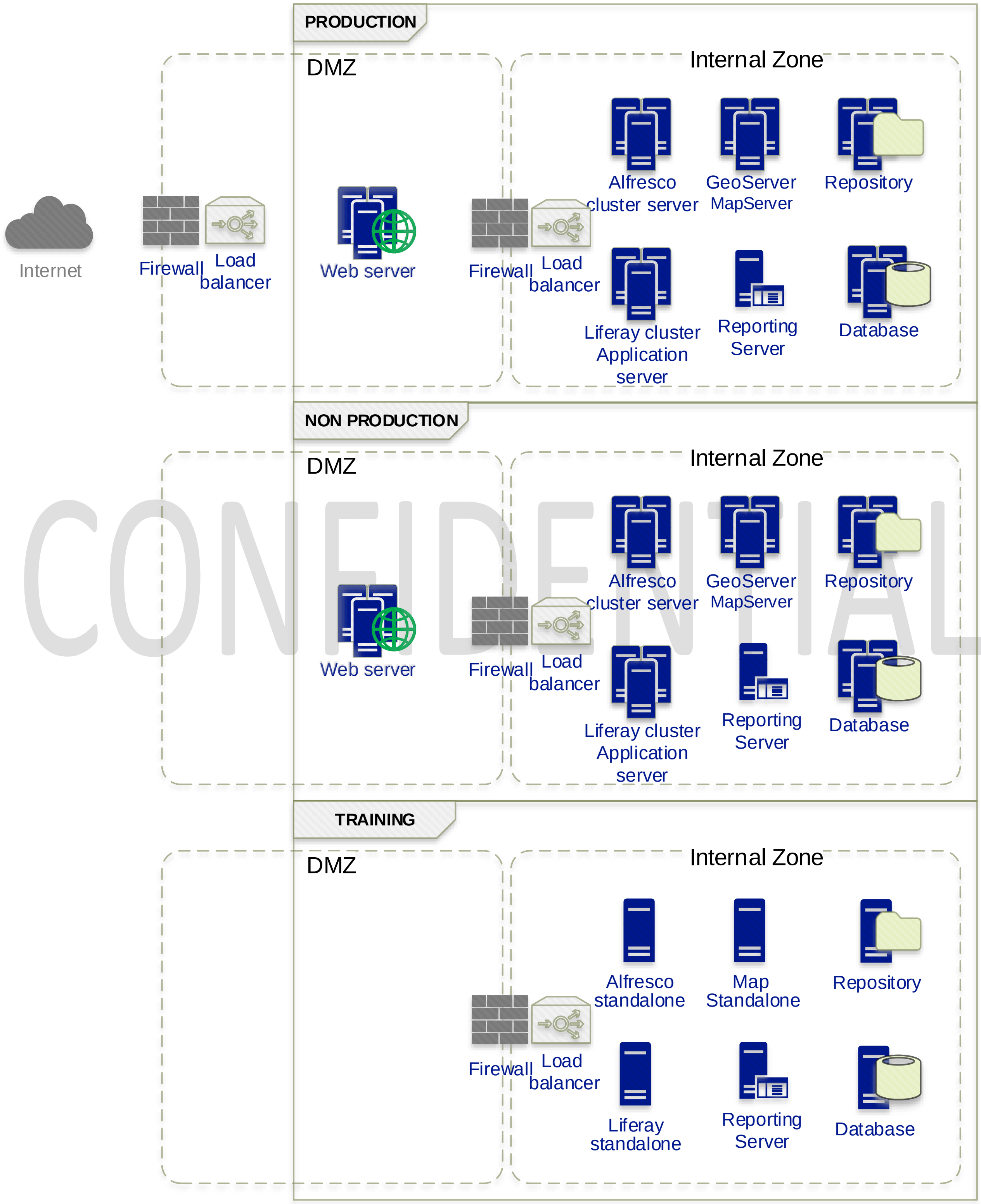


Figure 16 :Server Environment

7. Quality Attributes

7.1 Reliability

To ensure optimal reliability and supportability, the CIS system architecture relies on standard COTS product integrations & configurations wherever possible. External integrations are implemented by using standard APIs/Interfaces designed for that purpose. Customization will be used only where it cannot be avoided to achieve mandatory project goals.

This section summarizes the performance, throughput and scalability characteristic of the Solution. Scaling can be characterized as:

- **Horizontal Scaling –** the ability to add servers (virtual or physical) to the architecture to meet scalability requirements.
- **Vertical Scaling –** the ability to implement multiple instances of a solution architecture component on a single server (virtual or physical).

Table 3: Reliability Attributes

Quality Attribute Identifier	Reliability Design Decisions
All CIS component node	The initial CIS cluster installation comprises of two nodes on active – active mode. Horizontal scaling can be used for adding additional servers to the solution if required.
Scalability – User	The solution is designed to handle (concurrent users): 200 operational users 50 administrative users 4000 External Users(approximate)

Additional notes on concurrent user's calculation.

The concurrent user load is calculated based on the below formula:

Average Concurrent Users = Peak Visits per hour / (60 min/hour / average visit)

Peak Visit per hour is the number of unique visitors to the portal in its busiest hour. For operator we have taken 1200 visits per hour (putting together all operators)

Average visit length is the number of time, on average, each user spends on the CIS Portal. For operator we assumed 10 mins.

So, the average concurrent users for operators is = 1200 visits per hour x (10/60 average visit) = 200

Similarly, the assumption for administrative user's peak visits per hour is 300 and average visit length of 10 minutes. And for external users the assumption for user's peak visit hour is 25000 and average visit length is 10 minutes.

Dataworld cannot guarantee any specific sizing limits and response times for the solution, as these are largely dependent on the hardware configuration, and/or interfaces with third party/customer-developed software products, and the load in the production environment.

7.2 Performance and Throughput

To increase overall performance of the solution, network components (like switches, routers and firewalls) should be inspected. Increase the throughput of these devices to minimize eventual communication delay between solution components. To increase performance and throughput of the core components increase the available system resources in terms of CPU and Memory.

The solution is designed to be horizontally scalable to increase performance and throughput. Additional server nodes can be added to the cluster to scale out the application.

7.3 Availability

It is required that the solution meets DRDLR quality related standards expected of an IT System in a Production environment. As the primary role of CIS solution is providing the platform for the Cadastral information, is imperative that the solution design ensures, within the bounds of reasonable cost and complexity, 24x7 solution availability supported by as short a recovery time objective (RTO) as possible and associated Disaster Recovery attributes.

Three types of availability characteristics have been identified that are required from the solution:

- *Continuously Available* – this characteristic implies that no downtime is acceptable for the aspects of the solution to which it applies. The areas of the solution that DRDLR require to be continuously available must be able to run in a group, or in parallel with either system, in this configuration being able to service a specific request.
- *Highly Available* – this characteristic implies near-continuous availability of the aspects of the solution to which it applies. In other words, the areas of the solution that DRDLR require to be highly available cannot be subject to any prolonged outages during business-as-usual operations; downtime must be kept to the lowest possible level should a fault occur.
- *Rapidly Recoverable* – this characteristic implies that a period of unavailability is acceptable for the aspects of the Solution to which it applies. In other words, the areas of the Solution that DRDLR requires to be rapidly recoverable can be unavailable for a short period during business-as-usual operations.

This section describes how the CIS Solution addresses the uptime attributes defined in the Solution Requirements document.

Table 4: Availability Characteristics

Solution Component	Availability Design Decisions
All CIS components	Highly available

7.1 Disaster Recovery

This section covers the Disaster Recovery design considerations and approach. The DR provisioning is planned by creating a redundant site at another geographic location.

The DR / failover site should be the “mirror” of the live site in terms of data to avoid losing as much possible when switching to failover site (DR at another geographic location) is required.

A hot site recovery approach has been considered for the DR with the below-

Failover: Separate instance of all the solution component part of Active site to be deployed in the DR Datacentre. Active site is the “live” site exposed to end-users. DR site is used only for failover(only in crash of primary site).

Clustering: We simulate the cluster nodes for each application components, to have a hot standby system rather than a real cluster. Here the standby instance is mainly used for backup purpose (and of course failover). The failover site will be in read only mode and it will be synchronized in real time.

Data Replication: Replication and synchronization of file, storage and databases to be done using tools to have real-time(near to real-time) data.

The DR planned is a variation of high availability in an active-passive topology, both the active site and the DR site have a full deployment. The deployment includes synchronized database and filesystem. However, only the active site will actively handle the network requests from the users. The DR site becomes active only when the Active site experiences a service disruption. In that case, all new network requests route to the DR site.

The failover from Active to DR Site is much faster because the services are already deployed. This topology provides a very low RTO(Recovery Time Objective, is the duration of time and a service level within which a business process must be restored after a disaster in order to avoid unacceptable consequences associated with the disruption). The DR failover site must be ready to go immediately after failure of the primary(Active) site.

The following diagram shows the model of CIS Active and DR site.

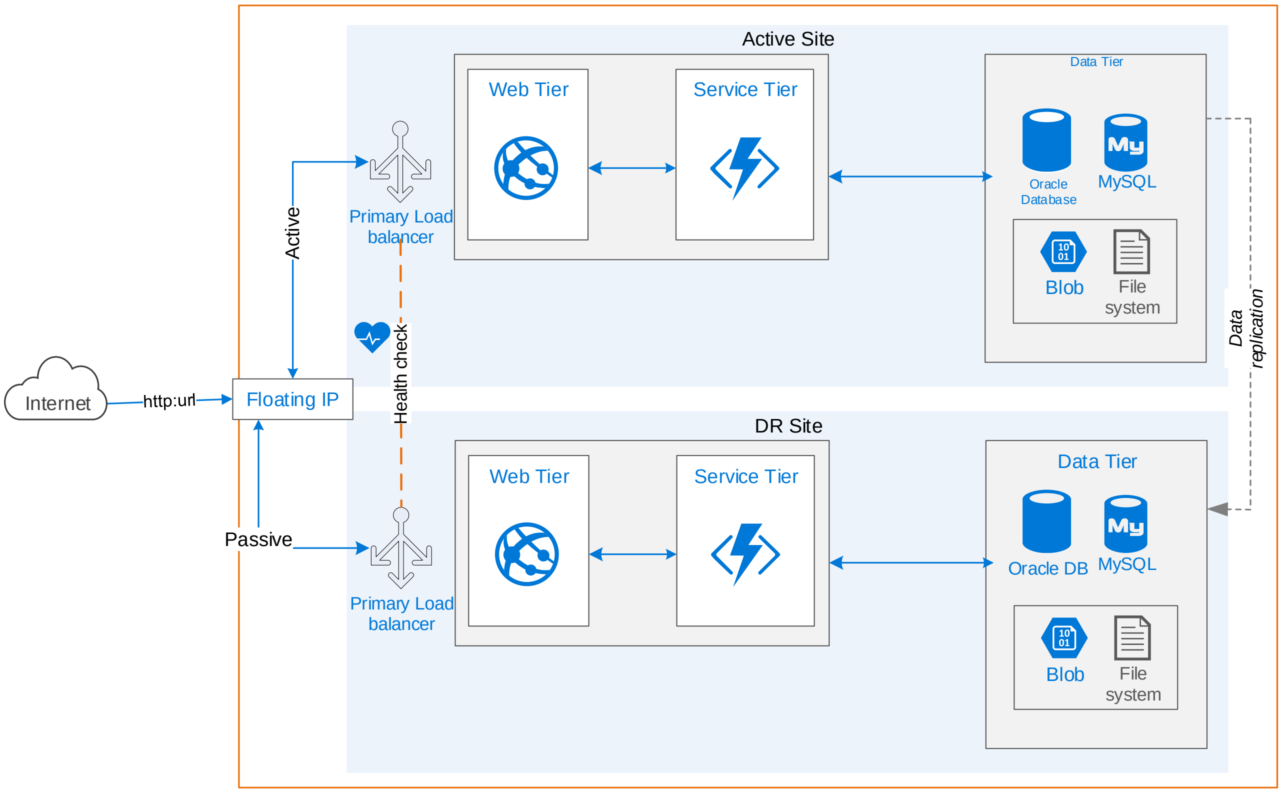


Figure 177: Active and DR model

The DR side of the diagram represents servers in a process 'cold' state (no solution processes are actively running), but the server OS can be active and running.

The synchronization of the server configuration (profiles, directives, XML files, etc.) can be propagated to the target DR servers when

- Changes to the active instance is changed
- Periodically, based on a specified time interval

And below assumption-

- Can be accomplished through existing synchronization, or replication tools for storage, filesystem and databases.
- Dataworld does not have a formal recommendation about the best tool to use, and will rely on existing standards of DRDLR
- Industry wide load balancers to be used

The load balancers will have keepalive daemons to check health status and to automatically failover to stand by if problem occurs. A floating IP address is required that can be moved between the primary and secondary load balancers. These will each be configured to split traffic between backend web servers. If the primary load balancer goes down, the floating IP will be moved to the second load balancer automatically, allowing service to resume.

7.2 Serviceability

7.2.1 Promotion Model

Design decisions regarding environment promotion are important items to document as Dataworld and DRDLR jointly develop the solution. Dataworld recommends a multi-tiered environment to ensure that the solution is tested appropriately before going to production.

Content, process and/or configuration related aspects and data of the Solution will need to be developed and adequately tested before being introduced into production. This is done through the establishment of non-production environments and a lifecycle approach taken to managing changes or updates to the Solution. Adequate testing of said data should be performed in each environment before considering promoting to the next.

The environments that are considered as required for the CIS solution are now as follows:

Table 5: Serviceability

Quality Attribute Identifier	Deployment Design Decisions
Dev/Non-Prod/ Prod	In the CIS Project, 3 environments are defined: • Dev: Dataworld build and test • Non-Prod (UAT): Test for Service owner for running test with CIS platform • Production: target environment for CIS platform
Software Distribution	Dataworld will define the steps needed to install and configure solution components manually.
Move to central repository and UAT-Production	After the services have been created in the UAT environment and tested on a functionality basis, these will be moved to a central repository, whose management will be outside of our scope. This external environment will also support revisions.
Move to production	After the service test plans have been fully validated on the service in the non-prod environment, they will be eligible for promotion to the Production environment. Similarly, as to what happened for the move to non-Production, the services will be promoted into production from the central repository. Any changes that must be done to the Production environment service should be done on the Development environment, as they will be mirror copies of each other. There should be no changes to the services done in the Production Environment.

7.3 Maintenance

Maintenance periods for the CIS solution components should be aligned to the server maintenance periods. The following maintenance schedule for server was identified:

- Needed and important updates: delivered by Dataworld as per maintenance agreement
- Critical updates: immediately, delivered by Dataworld

The downtime for each solution component due to maintenance tasks, are directly related to the availability requirements and constraints. This section summarizes how the maintenance model requirements defined in the Solution Requirements document are addressed by the Solution.

Table 6: Maintenance Strategy

Quality Attribute Identifier	Maintenance Design Decisions
Downtime during CIS Appliance/OS Upgrades	The Solution cannot be down during upgrades while in production. Portions may be taken offline to reduce capacity and enable the upgrade process, but the overall system may not be taken down for routine system patches. Utilization of Load Balancers during this task is critical.
Downtime during Application Patching	This can be first done on a separate VM and the changes can be tested. Once the patches are found to be there acceptable, a seamless switch of the load balancer connectivity to the new Portal will ensure near to no downtime.

7.4 Administrative Functions

This section summarizes how the administrative functions are addressed:

Table 7: Administrative Function

Quality Attribute Identifier	Administrative Function Design Decision	Actor

Quality Attribute Identifier	Administrative Function Design Decision	Actor
Master Image /code Maintenance	DRDLR will be required to maintain the master images/code for environments that make up the solution and store this in a suitable, centralized location. The Routine Backup Strategy will be referenced in the below section 5.1.2.4. • For the upgrade process, Dataworld recommends: ○ backing up the current system, ○ downloading the patches ○ ~~and~~ applying these patches to the master images to the Development environment first before starting non-Production and Production deployment. • Updates should be completed to the master VM images and fully tested in Development and non-Prod environments prior to rolling out the updates to the Production environment via rollout processes.	CIS Administrator
Routine Appliance/Server Health Checks	DRDLR will be required to perform routine system health checks on a periodic basis across all environments, but with priority on the Production environment. Dataworld recommends two major methods: • Utilize the network monitoring service to monitor the appliance for disk space, memory usage, and processes currently running and archive the findings. • Manually access the appliances and monitor the disk, space, memory usage, and processes currently running and archive the findings.	CIS Administrator / Enterprise Tools / Network Management Team

Quality Attribute Identifier	Administrative Function Design Decision	Actor
Routine Audit and Log Monitoring for Gateways	DRDLR will be required to perform routine Audit and Log Monitoring, periodically, of the systems across all environments, but with priority on the Production environment. Dataworld recommends to utilize applications that can parse and help read Audit database entries and Log files.	CIS Administrator

7.5 Backup Strategy

Dataworld is not responsible for configuring backup services for the Solution, but those requirements are documented here for reference. Optionally the Solution Run Book documents will address what needs to be backed up and how. Additionally, if there are special backup scripts/programs which are needed to back up a Solution component via means other than file system backup; those will be designed and configured as necessary to meet these requirements. From a physical architecture perspective DRDLR has a requirement for the Solution, where appropriate and applicable to do so, be instantiated on virtual platforms. As such it is intended that the capabilities of the underlying virtualization technology will be used as one mechanism for the backup and/or restore of components of the solution. This section describes the platform specific backup methods in use:

Table 8: Customer Backup Strategy

Server	Customer Backup Strategy Design Decision
Appliance Backup -All CIS nodes	Method: VMWare/ OracleVM Snapshot Occurrence: Weekly: Full (or Consolidation) Daily: Incremental Before Upgrade: Full (or Consolidation) Environmental Difference? None
Service-level Version Control	Repository: [Git, File server, Local appliance]

Server	Customer Backup Strategy Design Decision
Production Remote Audit Database	The customer has defined policies and procedures for performing backups of database servers. At a high level they perform full offline backups once a week and incremental backups daily.

7.6 Recovery Strategy

The following table summarizes the customer's recovery strategy:

Table 9: Recovery Strategy Considerations

Component	Customer Recovery Strategy Design Decision
All CIS nodes Appliance	Method: [Appliance-level restore VMWare/OracleVM Snapshot] Occurrence: [When issue arises] Environmental Difference? No
Service-level	Method: [Revision History Comments] Occurrence: [Major version change testing issues, Environment promotion issues]
Production Remote Audit Database	The customer has defined policies and procedures for performing restores of database servers. At a high level, they perform a base OS and DBMS install before a restore of the full offline database backup plus and incremental backups taken since the last full backup.

8. Supporting Network Services

8.1 Network Architecture Overview

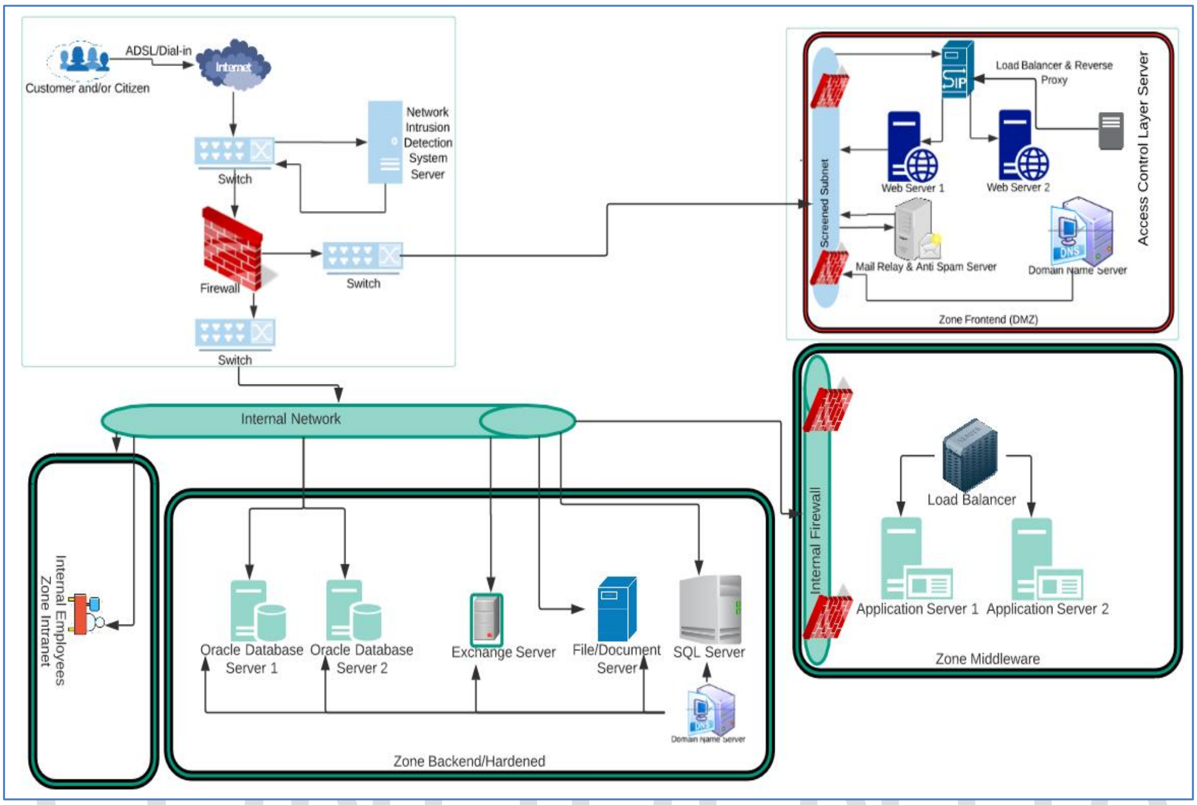


Figure 18: Network Architecture

8.2 Internet

The CIS solution will have web portal services that are internet facing to enable external users (i.e. registered and unregistered), other government agencies. The Web tier servers will be hosted in the DMZ behind the SITA standard load balancer and reverse proxy servers. As communication will be via SSL, all certificates installed on these servers will need to be issued by a public certificate authority.

The content distribution network (CDN) server will also have to be internet facing and will host large files and Geo spatial data, images etc that can be consumed by the requestors of the information. The content will be stored for a finite amount of time in accordance with the storage provisioning policies adopted by the department.

8.3 Network Switching

A switch is a device that filters and forwards data packets between network segments. Switches operate at the data link layer and sometimes the network layer, therefore switch support any data packet protocol.

The proposed solution recommends using Network switches with at least 1 Gigabit interfaces for routing network traffic. All processing nodes “virtual machines/physical” will be added to a subnet with security policies and firewall rules controlling access to the subnet.

8.4 Network Intrusion Detection System Server

The CIS Solution requires using the department enterprise standard NIDS and other network security controls to monitor and control access to the web portal and other content management services. The department can use the enterprise standard security devices to control access to the application.

An intrusion detection system (IDS) is a device that monitors a network or systems for malicious activity or policy violations.

8.5 Firewall

A firewall is a network security system that monitors and controls incoming and outgoing network traffic based on predetermined security rules. A firewall typically establishes a barrier between a trusted internal network and untrusted external network, such as the Internet.

8.6 Screened Subnet

The recommended solution for firewall implementation for the implementation of the CIS platform solution is the Screened subnet-based implementation where the web portal servers are hosted in the DMZ and other middleware servers are hosted on the internal network in a separate subnet.

A screened subnet is a network architecture that uses a single firewall with three network interfaces.

- Interface 1 is the public interface and connects to the Internet.
- Interface 2 connects to a DMZ (demilitarized zone) to which hosted public services are attached.
- Interface 3 connects to an intranet for access to and from internal networks.

8.7 DMZ Zone

The CIS solution is to add the load balancer, reverse proxy and the web portal web tier to the DMZ. Other application components will be housed in the internal networking

- A DMZ or demilitarized zone (sometimes referred to as a perimeter network) is a physical or logical subnetwork that contains and exposes an organization's external-facing services to an untrusted network, usually a larger network such as the Internet.
- The purpose of a DMZ is to add an additional layer of security to an organization's local area network (LAN): an external network node can access only what is exposed in the DMZ, while the rest of the organization's network is firewalled.
- The DMZ functions as a small, isolated network positioned between the Internet and the private network.
- This network feature allows the organization extra time to detect and address breaches before they would further penetrate into the internal networks.

8.8 Load Balancer and Reverse Proxy Server

Load balancer distributes incoming client requests among a group of servers, in each case returning the response from the selected server to the appropriate client.

A reverse proxy is a type of proxy server that retrieves resources or accepts request on behalf of a client from one or more servers. These resources are then returned to the client as if they originated from the Web server itself.

8.9 Mail Relay and Antispam Server

In order not to be considered "open", an e-mail relay should be secure and configured to accept and forward only the following messages (details will vary from system to system - in particular, further restrictions may well apply)

- Messages from local IP addresses to local mailboxes
- Messages from local IP addresses to non-local mailboxes
- Messages from non-local IP addresses to local mailboxes
- Messages from clients that are authenticated and authorized

In particular, secured SMTP mail relay will not accept and forward arbitrary e-mails from non-local IP addresses to non-local mailboxes by an unauthenticated or unauthorized user.

The relay is still effective as it is easy to forge e-mail header and envelope information, however it is considerably harder to successfully forge an IP address in a TCP/IP transaction because of the three-way handshake that occurs as a connection is started

9. Notes on Data Migration, Consolidation and Cleansing

DRDLR is responsible for performing the data migration to the centralised production environment including its consolidation and cleansing thereof. The DRDLR Data Migration will be planned to coincide with the project milestones and will be done in incremental stages.

The DRDLR will be responsible for the development of the Data Migration Strategy and Plan document. Data World will support the DRDLR team during this process through necessary input.

The High-level scope for Data Migration is as follows:

a) Scoping sessions to determine requirements;

b) Analysis of requirements;

c) Creation of the Data Migration Strategy and Plan;

d) Creation of migration scripts;

e) Creation of the new databases; in a centralised environment; utilising the configurations as specified by the solution requirements;

f) Migrating Alphanumeric and Spatial Data (from all 8 regional SG offices) to the centralised databases i.e. Production, Testing and Pre-Production/Staging area through ETL process;

g) Consolidating SG data at schema level for all 8 regional offices;

h) Identification of “dirty” data and data cleansing;

i) Testing of the migration and exception handling;

j) Migration Report generation;

k) Capturing of crucial data by the SG offices necessary for the new application;

l) Migration of the Map guide server to the centralised environment;

m) Migration of Image files from 8 regional offices to a centralised environment;

n) Archive data which will not be used by the new system.

o) Pilot Migrated Data for each SG office.

9.1 Quality Assurance

The Test Strategy will be developed during the Project Design and Documentation phase.

Table 10: CIS Revamp Project Quality Assurance Priorities

Dimension	Driver/Objective	Constraint/limits	Degree of Freedom/ Latitude
Location and Facilities	Each SG office will provide the Service Provider with the facility to setup for and to conduct the UAT testing.	Each UAT (User Acceptance Testing) for each milestone will be done at all regional SG offices.	Travelling of internal staff will be the responsibility of each SG office. Setting up of the UAT testing environment at the SG offices will be the responsibility of the Service Provider. The SG office will provide the necessary assistance to the Service Provider for a successful UAT environment setup. The infrastructure access and support must be provided by DRDLR. Prior to the release of milestone 1, the environment for UAT at the regional SG offices must have been readied by the Service Provider with the assistance of DRDLR.
Environment	Training and Users Acceptance Testing will be run from a centralised Pre-production area.	The Development, Testing, Pre-Production, and Production environments must be created during the Design and Documentation Phase of the project.	

10. Security Design Considerations

The security is defined based on best industry standards and noted in the Related Documents table. This has to be reviewed by the customer to meet their standards.

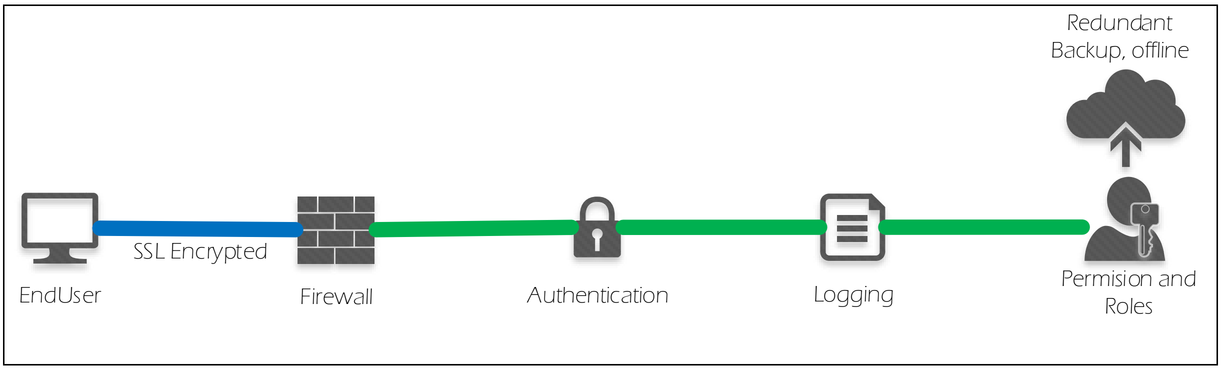


Figure 19: Security Design Consideration

The solution provides an enterprise security and compliance all the way to your end user, while still maintaining the simplicity and ease of access of the CIS services.

10.1 User Roles, Functions and Scope

This section describes the User Role and Functions constraints of the products and common components that are part of the CIS solution.

These roles can be stripped down to basically 3 different role types: Read-Only, Administrators and Super Administrator roles. By using “Group” like objects (groups, models, policies …) the solution will still support to create different grouping objects for the needed purposes.

The following table lists the actors that the solution will support:

Table 11: User Roles, Function and Scope

Quality Attribute Identifier	Business Responsibility	Description	Permission
Admin	None	Administrator of CIS Platform and all environment	• Customize Portal, • Administer configurations, users, etc. • Full access: read, write, assign Portal owner, etc. • Ability to stop and start services as well as manage the portal application.

Quality Attribute Identifier	Business Responsibility	Description	Permission
User	Consumer registered and unregistered	The *user* role will be given to ALL users as it is the base role for all CIS Portal users. This role is usually the one and only role of users and enables them to consume CIS Services or generate content that can be uploaded and managed by the content management system	• Modify and manage (develop, document, version, etc.) relevant to the role and context that applies to the user.
Geospatial Administrator	Geospatial Content Designers / Administrators	This user / role will have specialist knowledge of developing and managing applications in the Geoserver and open layers space	• Creating and customizing styles and other content data/metadata relevant to Geo Spatial content housed in the application. • Will be responsible for management, generation and facilitating the provisioning of Geospatial data
Internal User	Staff member/ Internal User	User will have appropriate privileges to generate, publish and consume content on the portal. These users will be presented with task interfaces relevant to their job function, description and access privileges.	• Generate, search, modify and publish content using the various interfaces provided by the solution
Process designer	Business Analyst / Process Developers	Business Analyst / Non-Technical users These users are the portlet developers that will have privileges to create BPMN models and workflows. Process Developers: • Responsible for developing the actual business logic for linking the business process to the concrete implementation and automation interface offered by the application	• Design business process models. • Develop, deploy, execute and test that the designed models align with business requirements.

Quality Attribute Identifier	Business Responsibility	Description	Permission
Content publishers	Produce/ design, validate and publish content that will be consumed by other users of the portal	T	Limited- API user permissions to read published internet APIs User is stored in a separate directory which needs to be connected first
Web admin	Responsible for the general upkeep and management of the portal and other content management and GeoServices applications	Will be responsible for provisioning and managing user access and assigning appropriate roles to the users.	• Edit the look and feel on the Portal. • Looking after the templates and brandings. • Look after the default page messages and default email alerts.

10.2 Authentication / Authorization / Entitlement

The following table outlines the authentication, authorization and entitlement design decisions as defined by the CIS solution. It also includes additional information relating to specific actors within the solution that is considered important to understand as part of these requirements:

Table 12: Authentication, Authorization and Entitlement

Quality Attribute Identifier	Actor/Actor Synonym	Description of Requirement
Authentication	All Actors and Actor Synonyms	Authentication to access any aspect of the CIS solution will be achieved by integration to existing SSO infrastructure, integrated to the specific DRDLR Directory or Database Store infrastructure. The users that are provided access to the CIS solution need to have a common attribute, like group membership or other Directory attribute that can be evaluated by SSO - Directory synchronization mechanism.
Authorization	All Actors and Actor Synonyms	Authorization to access any aspect of the CIS solution will be determined by membership to specific SSO or solution component groups (pre-determined) but can also be derived from Directory or Database Store if required.
Entitlement	All Actors and Actor Synonyms	Entitlement is achieved based on the security model of the respective CIS solution component. This may include • User membership in specific LDAP group • Using local solution component groups • Using local solution component security model

Quality Attribute Identifier	Actor/Actor Synonym	Description of Requirement
Provisioning	All Actors and Actor Synonyms	User provisioning is achieved by either one or more of the following • Adding the user to a specific Directory group • Adding the user to a specific solution component group • Creating a user security model in the solution component

Table – Authentication / Authorization / Entitlement

10.3 Firewall

The CIS is to be protected by multiple levels of firewall. This provides a solid basis for network security.

10.4 Data and Communications Security

Based on compliance, regulatory, security and audit requirements provided, the following communications security requirements have been identified:

Table 13: Data Encryption and Security

Quality Attribute Identifier	Description of Requirement
Internet & Intranet Communications	Initially the solution will not be exposed to the Internet (only to be done in later Phase). Within the DRDLR Intranet the solution will be setup with the capability to use secure communications wherever possible (access using HTTPS through SSL/TLS).
Data encryption	No data encryption will be provided beyond the out of the box encryption levels.

10.5 Audit Tracking and Access Logs

Access to the service can be audited by setting up security infrastructure to help identify security breaches if they occur. For current phases basic auditing of activity is also provided, in the logs of each solution component. Based on compliance, regulatory, security and audit requirements discussion with DRDLR, the solution may require review of the security infrastructure.

10.6 Data Retention

This section summarizes how the data retention requirements, are addressed by the Solution.

Because of the variety of solution components there are different requirements for data retention. E.g. information should be aggregated after some days and still provide information after 1-2 years whereas active alerts should not be purged at all.

The maximum retention time is basically limited by cost, disk storage and application performance.

DRDLR needs to outline data retention policies in more granular level. In further discussion with DRDLR, Dataworld can add controls in the solution component to have data retained in the solution environment and cater to forward archived data to external storage system.

10.7 Security Hardening Considerations

The following security hardening considerations/requirements have been considered in the CIS solution:

The DRDLR’s team of information security professionals should handles information security incidents, like:

- new security vulnerabilities in applications or operating systems
- viruses and Trojans
- Hacker attacks against web applications.
- All CIS servers to be scanned on a regular base for vulnerabilities and detailed reports are distributed to the Team with their findings and recommendations.