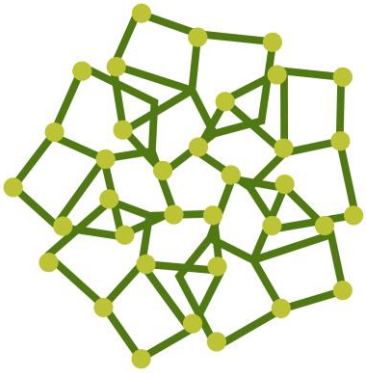


CRVSNOW

CRVSNOW

Enterprise Architecture Integration

Version: 1.0.0

Release State: RELEASED

Released: 04/04/2019

Confidential

Object CONSULTING

Table of contents

Revision history

Issue	Date	Description	Name
1.0.0	04 Apr 2019	Released	Brett McDowall

1 Introduction

1.1 Purpose of Document

This document provides an overview of how CRVSNOW fits into the Enterprise Architecture of a Government Department running a Births, Deaths and Marriages Registry. While CRVSNOW runs in the cloud and is complete within itself, the Registry needs to connect into it at a number of different levels. This document describes how this is done and what options there are to support different needs.

1.2 Background

CRVSNOW is a modern Java web application, using Angular for the core user interface, Spring components for infrastructure and Hibernate for persistence in association with an enterprise grade relational database and a NoSQL database. It uses several open source products to provide specific services and capability.

See the CRVSNOW Technical Solution Architecture document for more details about the internal architecture of the solution.

1.3 Scope

The scope of this document is the integration aspects of CRVSNOW which are needed to deploy this system for a Government Department.

1.4 Terms and Abbreviations

BDM	Births Deaths and Marriages or Civil Registration of Vital Events
Channel	Different ways of interacting with the registry, e.g. Online, B2B
CRVS	Civil Registration and Vital Statistics
CRVSNOW	CRVS System from Object Consulting available as SAAS. CRVSNOW is based on Connected BDM which is the data centre version of CRVSNOW. Connected BDM is based on the Enterprise Registry Management platform (ERM)
Core	Entry point for Registry officers to CRVSNOW
EFTPOS	Electronic Funds Transfer at Point of Sale
ePublic	Entry point for the public to CRVSNOW
eRegistry	Entry point for external stakeholders to CRVSNOW
ODS	Operational Data Store
POS	Point of Sale
RDD	Register Document Database
RDW	Register Data Warehouse

2 Positioning

A Government Department manages its own enterprise architecture and needs to ensure that any solution it uses fits into its guidelines and policies while also being best practice. The enterprise architecture defines how users should be managed, how security is provided, how systems are deployed and managed, how data is managed, how user devices are managed and much more. A solution such as CRVSNOW needs to fit into this enterprise architecture in a way that works for the department.

CRVSNOW has a relatively small number of integration points with the department's systems. The diagram below shows a conceptual view of the main system elements and connections. All users, including staff, connect to the system using a modern browser. Some Staff PCs, such as those on the counters, may need to support integration with specific hardware such as EFTPOS terminals and local scanners, some of which need specific software to be installed. Managed networked printers also need to be provided near the counters for printing of certificates. The Print Room and Mail Room need to be able print and scan certificates in bulk.

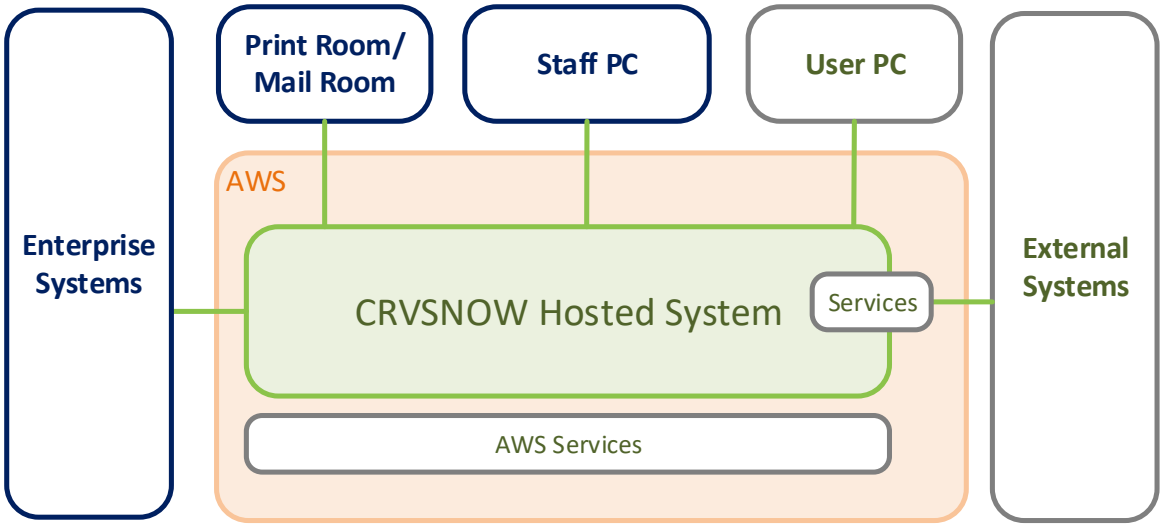


Figure 1. Conceptual View of Enterprise Integration Points

Enterprise systems can include:

- directory services
- content management systems
- document management systems
- finance systems

Note CRVSNOW is perfectly capable of working without these systems, however, integration is available for one or all as required.

External Systems provide several important services and use the services provided by CRVSNOW. This include:

- payment services through Westpac
- document verification through DVS (which is two way)

- accreditation verification through AHPRA
- proof of identify through IDMatrix
- address verification using QAS or Harmony
- an interface to a printing vendor to print commemorative certificates.
- stakeholder systems to submit notification documents directly into CRVSNOW for processing e.g., Hospitals

2.1 Interfaces

The main interfaces between the different system elements are shown in Figure 2. The separation between different physical sites is shown. Because CRVSNOW runs in the AWS cloud it can be used from anywhere with an Internet connection. It is also possible to use direct connect gateways to control staff access.

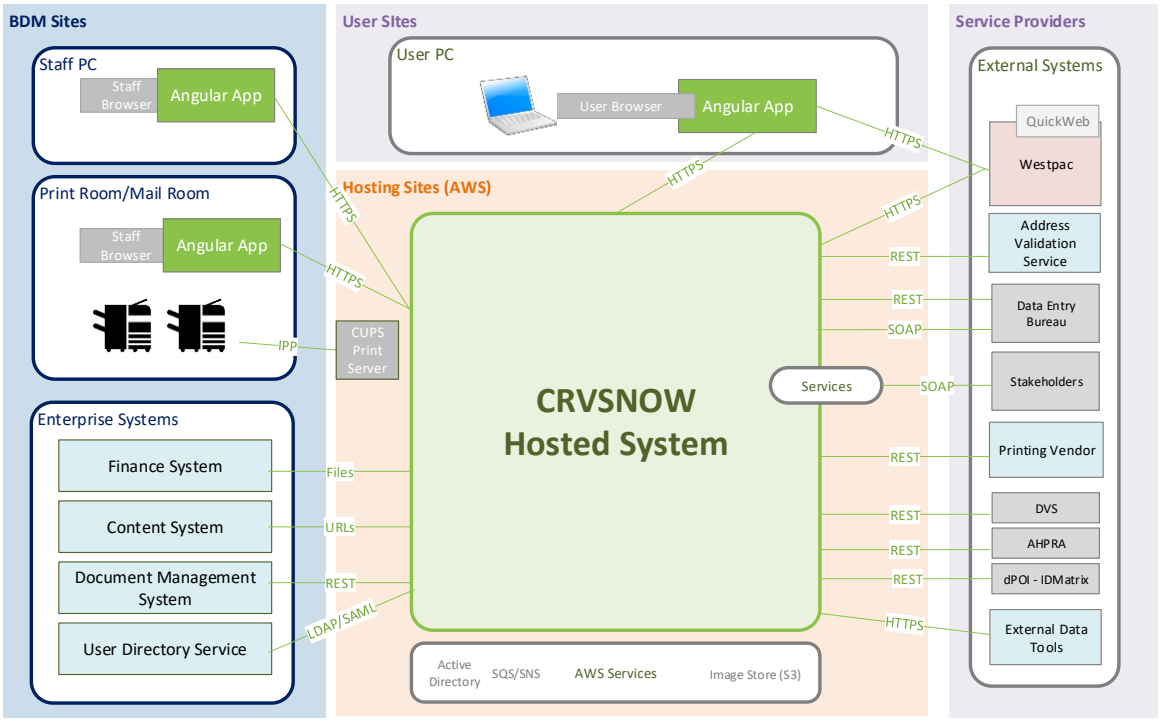


Figure 2. Physical sites and their integration points

eRegistry and ePublic users (User Sites) connect to CRVSNOW via the Internet using a browser, so there should be no integration issues. Staff who don't need POS access can also connect via the Internet using a browser. Staff PC's in departmental sites will need to have appropriate security settings to allow access to CRVSNOW as well as ensuring that access to the system is not blocked by network firewalls because CRVSNOW is in AWS. Government IT may setup a private VPN for staff access. The benefit of this is that access to the Core user interface (which is the most powerful) and its logon page, can then be further controlled.

The services to external parties are fully managed by CRVSNOW so they don't go through any departmental servers. This leaves the following interfaces that need to be supported by the department:

- Managed printing needs to directly access specific trays on designated printers. These printers need to be integrated with the CUPS Server running in CRVSNOW.
- CRVSNOW produces reports for consumption by the finance systems which are normally interested in information at the level of the general ledger. CRVSNOW can assist with reconciliations, but inevitably there is a banking element with time variations that need to be considered. There is no system level integration for finance currently.
- For ePublic and eRegistry users, the department's web site can link into CRVSNOW to allow complete control over the user experience.
- Help screens are URLs which can be hosted in an external content management system.
- CRVSNOW stores all documents and images in in AWS S3 buckets, so there is no need for a department document management system. However, CRVSNOW has been integrated with Vignette previously.
- CRVSNOW has its own directory – the AWS Directory Service which uses Microsoft Active Directory. It is used to authenticate staff and stakeholder's users. The CRVSNOW User & Stakeholder Management subsystem handles the interface. There is no need to directly integrate the departments directory service into the CRVSNOW directory as single sign-on is not a desirable feature for a BDM Registry. User access needs to be carefully and individually managed. Bulk upload facilities are provided to make user account creation simple. CRVSNOW does support shared sign-on between user applications.

3 Components & Services

The main software components and services are shown in Figure 3 below. It is a more detailed version of the diagrams above and shows the five app servers (running on Tomcat) and the five adjunct servers. The databases and search are shown along with the hardware and software needed for a staff PC (typically located at a counter or service area for the public).

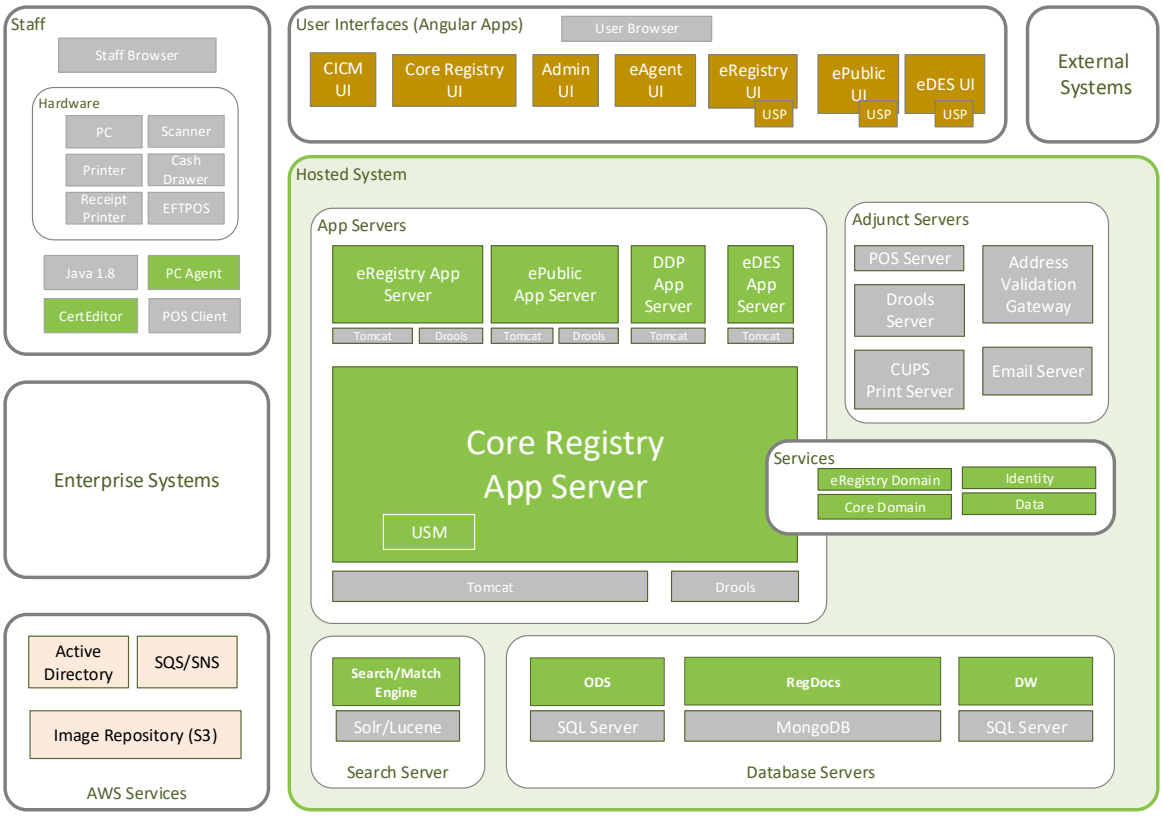


Figure 3. Software components and systems

3.1 App Servers

There are five app servers:

1. Core Registry - the main application server is known as “Core”. It supports all the registry functions including processing data, creating documents, and generating certificates. It supports the CCIM, Core, Admin and eAgent user interface applications.
2. eRegistry - the app server for the eRegistry functions supports the eRegistry user interface and integration into the Core Register app server.
3. ePublic - the app server for the ePublic functions supports the ePublic user interface and integrates into the Core Register app server.
4. DDP – the Data Delivery Portal allows users to collect data (in sets of files) in a secure way to avoid sending data via email. There is also support for SFTP – both inbound and outbound.
5. eDES – the Data Extract System allows registered users to extract data from the system under contracts which can require payment. eDES integrates into the Core Register app server to

manage the process but the data extracts work directly off the data warehouse. Includes integrated data delivery support using DDP.

3.2 Databases

There are three distinct databases:

1. Operational Database, also called the Operational Datastore (ODS)
2. Registry Documents Database (RDD)
3. Registry Data Warehouse (RDW)

The Operational Datastore consists of the core transactional database for managing the system including user roles and privileges, products and prices, reference tables etc.

The Registry Document Database (RDD) is a NoSQL database which is well suited to storing many document types with complex data structures.

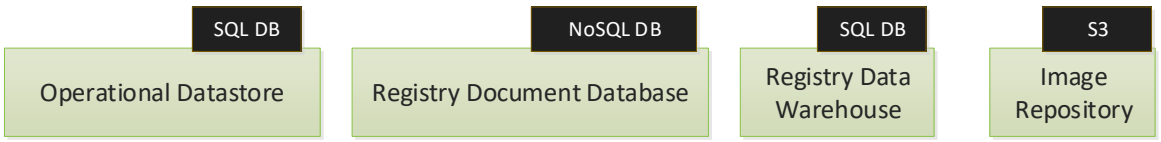


Figure 4. Data Tier Architecture for CRVSNOW

The Registry Data Warehouse (RDW) uses its own instance of SQL Server as it has different operating characteristics. It is updated by the core server as updates are made to the registry documents in RDD. There is no direct connection between any of the databases – just references that can be resolved by the application servers.

At present the ODS and RDW use Microsoft SQL Server 2014. Customers can choose to use MySQL to reduce operating costs.

The images (files that are documents and images) are stored in AWS S3 buckets with the ODS managing the metadata and the links to the files. The Image Repository is implemented as a standalone subsystem with a web service interface to the core system.

3.3 Searching & Matching

The search facility uses Solr/Lucene to provide high speed searches of all the data in the system. In some cases, Solr/Lucene has enough data in its indexes to fulfil the entire search request, i.e. it doesn't need to lookup data in the RDD which makes these searches extremely fast. Solr/Lucene also provides the underlying capability for matching. The search indexes are kept up to date by the Core Register application which is responsible for all changes to the Core Register Databases.

Matching is trying to find other documents that have the same key criteria as the source document. The principles of thresholds, weightage and total score are used to bring up the most relevant results. The number of records returned is controlled by field weightings and Partial / Full Threshold configuration. Matches can be

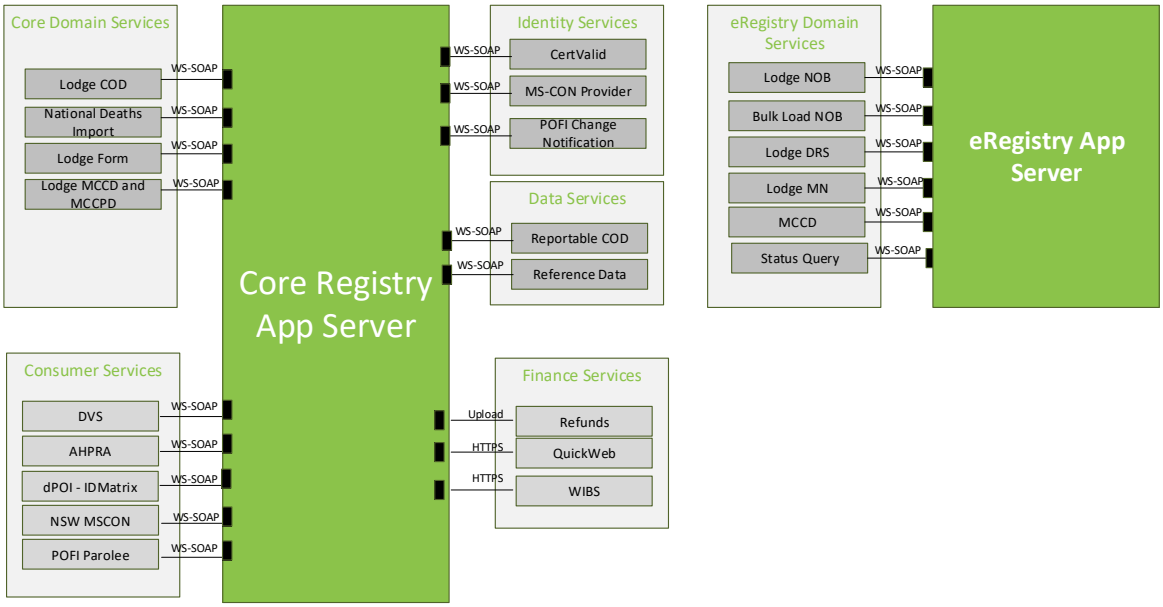
- perfect (one match only on all fields)

- perfect and partial (less than 100% match)
- full (many records with 100%)
- full and partial
- partial and no match.

In some cases, the results of the matching engine (which uses Solr) are re-rated by checking specific details of the records.

3.4 Domain Services

CRVSNOW is both a provider and consumer of services. The diagram below shows these services grouped into categories. SOAP is used for many of the interfaces since by their nature they are contractual, i.e. the submission of a document (as in data) needs to meet strict requirements. SOAP makes it easy to define the contract and specify the data (as against RESTful interfaces which are better for flexible interactions, such as for the user interfaces).



At present there is no API Management software used as it isn't currently necessary. There are a small number of well-defined and mature interfaces that are used to by a tightly managed group of stakeholders. API Management could be added to support versioning, testing, monitor service levels, and manage access.

3.5 User & Stakeholder Management

The users and stakeholders need to be registered and approved, and their permissions set appropriately. Stakeholders request registration, providing details of their key admin users. Once approved they manage their own users. Users can have different types of accounts which use different authentication methods.

3.6 Payments

Westpac Bank services have been integrated credit card payments eFamilyHistory, eRegistry and ePublic.

Westpac's QuickStream QuickWeb product provide the UI after ePublic directs the user to the Westpac hosted payment page (with CRVSNOW branding). Once payment is accepted or declined, ePublic is given appropriate status information and resumes control. There is no ability to refund a transaction via QuickWeb.

BPAY is supported through integration with Westpac's WIBS. This allows CRVSNOW to automatically download payment information and match them with transactions awaiting payment by BPAY.

Counter transactions (pay in person) is handled by the integrated POS and Westpac EFTPOS pin pads. The POS controls the pin pads and the pin pads communicate directly with the bank.

3.7 Finance

CRVSNOW provides operational accounting, which means that it is handling the transactions and payments and recording what happens. It supports a range of operational reports and can allocate received payments to general ledger accounts. However, the true accounting which involves reconciliation and recognition of close off period, accrual, refunds, journal entries etc. is not supported by CRVSNOW. It is expected that the departments accounts department will use the information provided by CRVSNOW to feed into an accounting system.

3.8 Printing

Printing is supported via the CUPS Print Server which connects to printers at Registry offices. CUPS is a standards-based, open source printing system. It is hosted on a standalone EC2 Linux instance. The Internet Printing Protocol (IPP) is the native protocol of CUPS, though several other printing protocols, such as LPR/LPD, are supported for printing to local and network printers.

There are two main types of printing:

1. Internal Legal Certificate printing
2. External Commemorative printing

Internal Legal Certificate Printing

CRVSNOW supports managed printing for secure paper. This means that the printed pages need to go to a specific print queue which is connected to a specific tray on the printer which contains a specific batch of secure paper. There are no user selectable print queues. When the certificate printed there are three barcodes on the certificate which must be scanned to allow the paper to be tracked and linked to the certificate and the registry item.

For bulk printing there is also a bulk scanning facility which links all the printed certificates to their paper sheet via the scanning process which reads the barcodes (in bulk).

Straight-through-printing (STP) is used in the Mail Room. It is a process whereby CRVSNOW spools output to a defined print queue, once all criteria for processing an application are compliant (which could happen automatically). A key attribute of STP is the static assignment of paper types, as specified in the templates for each document or certificate type, and the physical printer trays in which the paper is loaded. CRVSNOW will automatically print compliant or completed requests to the CUPS queue and printer tray for the designated paper type. No user intervention is required or possible.

External Commemorative Certificate Printing

A web service is provided for external 3rd party vendors who provide commercial printing services. It is possible to support multiple print vendors. The SOAP web service is called by the vendor to get a list of orders which contain the information needed to populate the templates to be printed. The vendor sends back a message when each one is printed and delivered.

3.9 Scanning

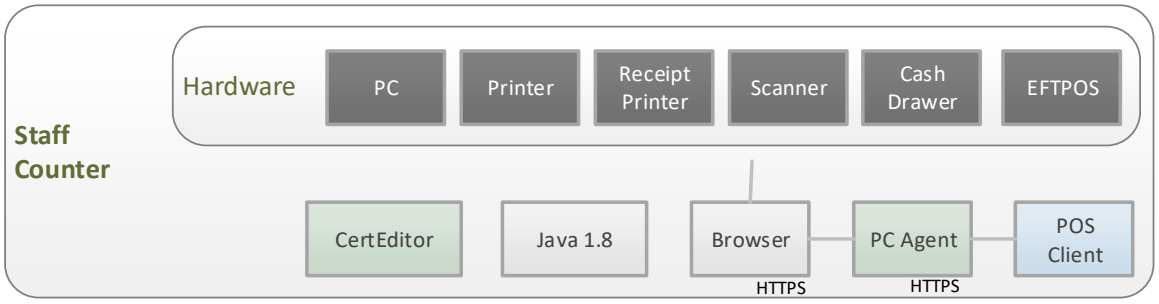
For scanning on staff PC's, the scanner writes its images onto a designated directory on the PC's hard drive. The PC Agent watches this directory and transfers any files to CRVSNOW, so the user doesn't have to explicitly deal with these files. The user clicks two buttons: one for the scanner which scans and puts and image into the folder, and, one in the user interface to select the file.

3.10 Email

An SMTP mail server for outbound email and a MAPI mail server for inbound email. A hosted email service is provided using DoveCot for IMAP and Postfix for SMTP. The mail server uses the elastic file system to share and maintain currency of mail data across availability zones. AWS's email services, SES, is not bound to a region so it is unsuitable for CRVSNOW, however, SES can be used for testing.

3.11 Staff Counter PC

The staff counter has several hardware peripherals which need to be integrated into CRVSNOW – via the PC. This allows staff to take payments, open/close a cash drawer, print receipts, print documents, and scan documents. Every staff PC is configured (as its SOE) to be a counter PC but only some will have all peripherals connected.



Due the security constraints of modern browsers a PC Agent is required to interface to hardware and other applications running on the PC such as the POS Client.

4 Data Migration

Data migration is the process of converting any existing electronic data into a format that can be directly loaded into CRVSNOW. This involves mapping record types and data fields and transforming data at both the record and field level. One of the challenges is that the data models used will invariably be quite different (unless the existing system is an earlier version of CRVSNOW in which case we have made the data model upwards compatible). Older systems have ad hoc data models where data can be kept in different fields using a range of formats and some new data records are interpreted by the code from old records or by combining records. This leads to the need for data cleansing to fix errors in the data and find code-related data. The most complex data migration mappings require one-to-many or many-to-many mapping of data records, e.g. a single birth record may hold adoption and/or change of name information which may require the creation of several related records in CRVSNOW.

There are three approaches to data migration (in order of least time, effort and cost first):

1. Simple migration of records with the focus on being able to retrieve them so they can be fixed manually when needed. Only minimal cleansing is performed to allow data to load and be indexed.
2. More complex mapping of records to make the data more like it has been entered directly into CRVSNOW but with limitations and many rejections. More complete cleansing is needed here, and rejections need analysis (or can be referenced in the old system until needed).
3. Full mapping of all records so that the data is exactly as if it was just entered – including history. Obviously, this can be a challenge. All records are rigorously cleansed, so the result is no errors or rejections.

Cleansing can be performed by the data migration processing when there is a well-defined mapping/algorithm and where rejection/error reports can be used to edit the data in the source system.

Jurisdictions need to provide the schemas for the data and representative samples or obfuscated datasets for test processing. The schemas should be mapped against the standard UN data model.

If the existing system is an earlier version of CRVSNOW then there is no need for a full data migration since the data is already in the correct model. It is simply a matter of transferring the data between the systems. Object will fully manage the process.

Specific EC2 instances are created to perform the data migration processing to reduce elapsed time. This is also very useful during testing and it allows more tests to be run in a given period.

5 Deployment Models

CRVSNOW is designed to run in the cloud – currently AWS. However, it is can run in a government data center of some other hosting facility if needed. Given the trends for all business and government to move their computing systems to the public cloud, there seems little benefit in not using the cloud hosted version.

There are two distinct deployment models which use AWS:

1. Managed Hosting – CRVSNOW is installed in cloud infrastructure on AWS using accounts owned by the department. Object installs and manages software in partnership with the department. The exact delineation of support responsibilities can be defined in the contracts. This approach means the department is buying the software license and taking full responsibility for the system, including paying the AWS fees.
2. SAAS – CRVSNOW is provided by Object as a shared service. There are yearly fees but no upfront license fees. This approach means the department is outsourcing the responsibility of running and maintaining CRVSNOW to Object Consulting and its support partners.

In each case there are different support models available ranging from best efforts to premium, which can be tailored to specific needs.

5.1 AWS Hosting

CRVSNOW is designed to run within a single AWS region to meet data sovereignty requirements. It uses the multiple Availability Zones which AWS provides within each region to ensure high availability. Availability Zones are connected to each other through private fiber optic networking, so that applications can automatically fail-over between zones without any interruption.

CRVSNOW uses the following AWS services and resources:

- Elastic Compute Cloud, or EC2 is an IaaS service that allows subscribers to run application programs in the computing environment. The EC2 can serve as a practically unlimited set of instances (or virtual machines), created from an Amazon Machine Image (AMI). The AMI are like templates that are configured with an operating system and other software, which determine the user's operating environment.
- EC2 Auto Scaling permits the dynamic instantiation of additional EC2 instances, according to user defined policies that trigger when monitored metrics breach a threshold.
- Amazon Simple Queue Service (SQS) is a scalable, pay-per-use web service for storing messages in transit between systems, applications or micro services. It moves data between distributed applications with decoupled components, without the overhead of creating and maintaining message queues. SQS supports tasks that process asynchronously.
- Amazon Simple Notification Service (SNS) is a service for coordinating the delivery of push messages from software applications to subscribing endpoints and clients. SNS uses the "publish-subscribe" pattern, whereby notifications are delivered to clients using a "push" mechanism that eliminates the need to periodically check or "poll" for information. SNS has multiple notification channels available. The Short Message Service (text messaging) delivery channel is used in CRVSNOW.
- S3 is an IaaS service that provides persistent storage accessible directly through RESTful web requests, rather than by block mode transfers via host adapters, typical of local/SAN disk. S3 is a standalone, highly scalable storage service, with 99.999999999% data durability. It is

particularly useful for storing backups and large volumes of infrequently accessed static data (e.g., images).

- EBS volumes stores data persistently as blocks of the same size, organized hierarchically like a traditional file system. They can be used only in conjunction with Amazon EC2, or else kept in a standby mode. EBS stores data in volumes of a provisioned size, attached to an EC2 instance. It can't be easily scaled, and if more storage is required, new volumes of a larger capacity must be configured. EBS volumes are designed to be highly available and reliable. Amazon EBS volume data is replicated across redundant infrastructure in an Availability Zone to prevent the loss of data from the failure of any single component.
- Amazon Route 53 is a scalable and highly available Domain Name System (DNS) service. Route 53 connects user requests to AWS infrastructure services, including the public facing Elastic Load Balancing load balancers (see below), web servers and S3 buckets. It also provides health check integration to CloudWatch.
- Amazon CloudFront is the content delivery network service at AWS. CloudFront is a globally distributed network of proxy servers deployed in multiple regions. It optimizes the delivery of static assets on a website, such as HTML, CSS or JavaScript files, serving them from edge cache locations, which are closer to the consumer of the content. CloudFront offers a single point of public ingress into the CRVSNOW AWS environments, and an additional layer of network protection (e.g. from DDOS attacks) in association with AWS Shield.
- AWS elastic load balancing (ELB) is an IaaS service that automatically distributes incoming application traffic across multiple EC2 instances, enabling application fault tolerance. It dynamically scales its request-handling capacity in response to incoming application traffic volumes. ELB integrates with EC2 Auto Scaling to ensure that enough back-end capacity is available to meet demand levels without requiring manual intervention. ELB optimizes fault tolerance by adjusting capacity according to incoming application traffic. ELB can be enabled within a single availability zone or across multiple availability zones to maintain consistent application performance. ELB will route traffic based on either network connection load or application load. Network level loading (concurrent connections) is enough. ELB can detect unhealthy instances and route traffic away from them to other instances in other AZs.
- AWS security groups control the access that traffic has to AWS resources. The two key controls used in CRVSNOW are:
    - VPC security groups, which control access to EC2 instances and the database (DB) instances inside a VPC. VPC security groups are implemented by defining security group rules, which permit each rule to allow/disallow a specific source to access a resource (e.g. DB instance) associated with that security group
    - EC2 security groups, which control access to a specified EC2 instance. An EC2 security group filters all traffic entering or exiting a specific server and can protect a server from attacks that originate from both within the same subnet and other subnets.
- Amazon Relational Database Service (RDS) is a fully-managed SQL database service. It provides cost-efficient and resizable capacity, while managing database administration tasks such as migration, backup, recovery and patching. The CRVSNOW relational database uses Microsoft SQL Server 2014 Standard Edition.
- AWS Identity and Access Management (IAM) securely controls individual and group access to AWS resources for users and applications. IAM features are used to securely give applications that run on EC2 instances the credentials that they need to access other AWS resources, like S3 buckets and SQS.
- Amazon CloudFormation (CF) provides a mechanism for creating and updating a collection of related AWS resources. CF uses templates in JSON format to create a "stack" of AWS resources using the Management Console, APIs or AWS command line.

- CloudWatch is a monitoring service for AWS IaaS resources. It can monitor resources such as EC2 and RDS instances, as well as any application generated log files. CloudWatch collects and tracks metrics, monitors log files, set alarms, and automatically react to changes in AWS resources.
- Amazon QuickSight is a business intelligence (BI) service that makes it easy for users to analyze and visualize their data to gain business insights. QuickSight accesses data from AWS storage services, collates and formats it and injects it into an in-memory computation engine (SPICE) for high performance analysis and presentation. A dashboard creates visualizations, tables and other data displays. QuickSight will read data from multiple AWS data stores, including RDS and S3, and external sources such as data in a local file or uploaded to an S3 bucket.

5.2 Availability

All EC2 instances run active/active in auto scaling groups across two of the three availability zones. This configuration provides high availability and performance, while minimizing costs (AWS services and software licenses). This configuration has been deemed to be enough given the availability requirements which do not require strict HA. If it is determined that availability/recovery requirements approach a near-zero downtime, then EC2 instances running (or available) in three AZs will be considered. In general, workload processing will be distributed evenly across the AZs to optimize performance. A SAAS version can have increased availability.

To reflect the criticality of the data stored in the Register, the MongoDB database is deployed across three EC2 instances, one per availability zone. A MongoDB Replica Set of two nodes and an arbiter provide improved high availability of Register data should one node in the cluster fail.

The distribution of the functional components across these AWS services not only enhances the security framework of the solution but reflects the prevailing cloud computing model of optimizing spend and service continuity through smaller and lower cost servers. Lightweight instances, such as the T2 series, utilize CPU cycles better and more consistently. Load balancing and autoscaling instances improves availability and performance, while lowering the impact risk of a server outage by limiting its scope to only the hosted application.

6 Enterprise Options

CRVSNOW is a very powerful and complete system, however, there is still room for customizations and extensions to suit each customer.

CRVSNOW can be extended to integrate with virtually any other system. In some cases, there are ready made interfaces, but if necessary new interfaces and end points can be added. This could include integrations with identity systems or health systems.

The way that information is exchanged with other parties, such as other jurisdictions can be automated, controlled and managed in several ways. CRVSNOW already has facilities to automatically deliver reports, however, there could be a need for very secure and controlled B2B integration between the CRVS systems of different jurisdictions.

ePublic can be independently developed so it is a seamless part of existing government web sites. However, the compliance rules for the data still need to be enforced so there is integration work required. The RESTful APIs that are used by the ePublic user interface could be used by an independently developed application if desired.