Solution Design

<Project Name>

Author: TBC

Version: 0.1

Status: draft

Issue Date: 02/01/2018

Protective Marking: Official

Contents

[0 Document Information 3](#_Toc503878616)

[0.1 Document History 3](#_Toc503878617)

[0.2 Lifecycle History 3](#_Toc503878618)

[0.3 Referenced Documents 3](#_Toc503878619)

[1 Purpose 4](#_Toc503878620)

[2 Background 4](#_Toc503878621)

[2.1 Objective 4](#_Toc503878622)

[2.2 Key Drivers for Change 4](#_Toc503878623)

[2.3 Business Outcomes 4](#_Toc503878624)

[2.4 Solution Overview 5](#_Toc503878625)

[2.4.1 PDI Overview: 5](#_Toc503878626)

[2.4.2 Process Flow 5](#_Toc503878627)

[2.4.3 HLBRs 5](#_Toc503878628)

[3 Solution Requirements 6](#_Toc503878629)

[3.1 User Stories 6](#_Toc503878630)

[3.2 Scope: 6](#_Toc503878631)

[3.3 Physical Component View 6](#_Toc503878632)

[3.4 Data Ingestion 6](#_Toc503878633)

[3.5 Data Exploitation 6](#_Toc503878634)

[4 Data Architecture 6](#_Toc503878635)

[4.1 Data Flow Diagram 6](#_Toc503878636)

[4.2 DM Data Model 6](#_Toc503878637)

[4.3 Data Dictionary 6](#_Toc503878638)

[4.4 Data Access and Security 6](#_Toc503878639)

[4.5 Data Retention 6](#_Toc503878640)

[5 Key Assumptions Risks, Issues & Dependencies 7](#_Toc503878641)

[5.1 Assumptions 7](#_Toc503878642)

[5.2 Risks 7](#_Toc503878643)

[5.3 Issues 7](#_Toc503878644)

[5.4 Dependencies 7](#_Toc503878645)

[6 Architecture Decisions 8](#_Toc503878646)

[Appendix A: Glossary 9](#_Toc503878647)

# Document Information

## Document History

| Version | Date | Author / Contributor | Reviewers | Reason for Issue |
| --- | --- | --- | --- | --- |
| 0.1 | dd/mm/yyyy | TBC |  | Initial draft for review |

## Lifecycle History

| Lifecycle Stage | ARB Approval Date |
| --- | --- |
|  |  |
|  |  |
|  |  |

## Referenced Documents

| Ref | Document | Current Version | Location |
| --- | --- | --- | --- |
|  |  |  |  |

# Purpose

This Solution Design has been produced to give a functional and technical overview of the <Project Name>. The document will provide details of:

1. An overview of the solution
2. Key requirements for the project
3. Any dependencies on architectural changes needed to support the migration requirements above, specifically where any impact to the Platform Component Model or Solutions Design are identified
4. Data Architecture
5. Process Flow

**This document forms part of the main governance document set and is subject to review by the Architecture Review Board.**

# Background

A brief background about the project stating its objective, key drivers and business outcomes.

## Objective

## Key Drivers for Change

## Business Outcomes

## Solution Overview

### PDI Overview:

PDI is the Pentaho Data Integration tool that is used to orchestrate and process data within CbC. It processes definition files in xml format called Jobs (extension .kjb) and Transformations (extension .ktr). These files are commonly referred as kettle jobs and transformations and can be read and interpreted by either a client tool (kitchen, pan, spoon), by the Pentaho server or by Carte slaves. These tools then execute the instructions (called steps) as specified in the Jobs and Transformations using the kettle engine to achieve the goal intended.

Steps in Jobs are executed sequentially; Jobs are used essentially for orchestration and file management. Steps in Transformations are executed in parallel; they are used for data transfer, transformation and generation.

In CbC, we make use of two Pentaho servers, aided by two Carte slaves. Spoon, a visual development tool for Jobs and Transformations) is used for development. Carte is a standalone web server that runs the kettle engine. Jobs and transformations stored in the Pentaho server can be forwarded to Carte for remote execution. In this way we can parallelize file processing.

*Note:*

*Prior to version 7.0, the BA (business analytics) and the DI (data integration) servers were different products. After version 7.0, inclusively, these two products were combined into one, now named the Pentaho Server. For that reason, we will refer to the Penatho Server in OfDS but to the DI and BA Servers in the EDH.*

### Process Flow

Description about each component in the project.

### HLBRs

# Solution Requirements

## User Stories

For detailed view on the user stories, refer to the following links:

User Stories can be found <here>.

## Scope:

Detailed scope of the project

## Physical Component View

A brief explanation of each component in terms of its physical characteristics i.e. the kit it is running on, compute capacity etc.

## Data Ingestion

## Data Exploitation

HLBRs for CBC Analytics

# Data Architecture

## Data Flow Diagram

TBC

## DM Data Model

TBC

## Data Dictionary

TBC

## Data Access and Security

TBC

## Data Retention

TBC

# Key Assumptions Risks, Issues & Dependencies

## Assumptions

## Risks

## Issues

## Dependencies

# Architecture Decisions

Any key architectural decision worth noting.

# Appendix A: Glossary

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| ID | Term | Description | Classification  (Acronym or Business Term) | Status  (Proposed/  Approved) |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |