Design (QATProm06)

Introduction

The Enterprise Address Validation Service (EAVS) provides address standardization, verification as an Enterprise Shared Service (ESS) across the HHS 2020 systems. A third-party tool, Experian CorrectAddress, enables the EAVS. This design addresses the requirements identified in this Jira filter.

Enterprise Address Validation Service (EAVS)

This document details the design to address the use cases identified for the Address Standardization and Validation (ASV) QAT promotion. The solution design follows the guidance provided in the System Design Document (SIPLT1) and the NM HSD Reference Technical Architecture (RTA). This document provides the logical, physical, and communication diagrams of this component of the SI Platform.

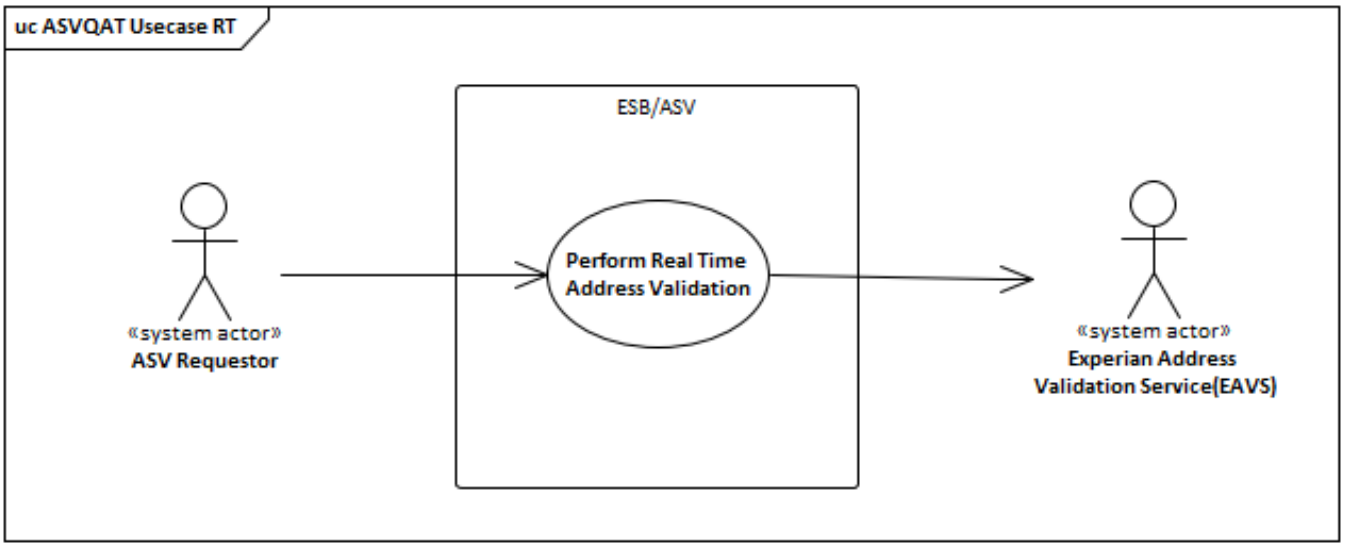
Use Cases

As identified in the JAR sessions for the ASV workstream, the use cases for EAVS are the real-time, and batch invocations. The design of the ASV as an enterprise shared service will primarily address these use cases.

Real-Time Use Case

This use case describes functionality to request a real time validation from a service by performing bi-directional format transformation using canonical schema.

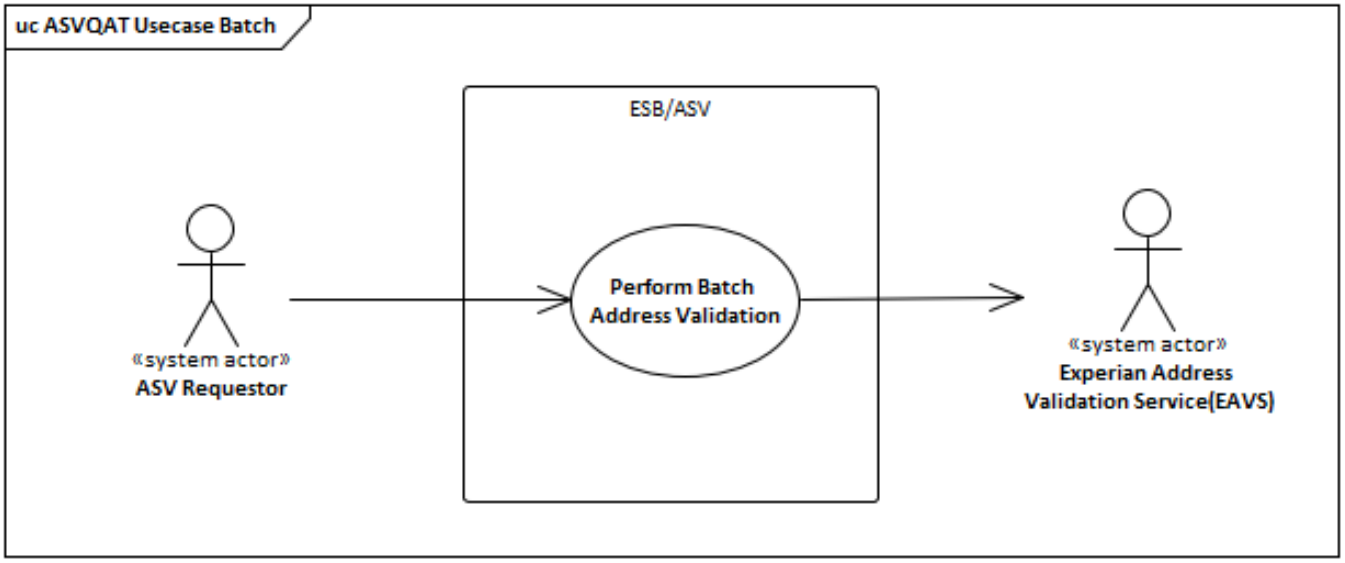
Figure 1: EAVS – Real-Time Use Case



Batch Use Case

This use case describes functionality to request batch validation from a service by performing bi-directional format transformation using canonical schema.

Figure 2: EAVS – Batch Use Case



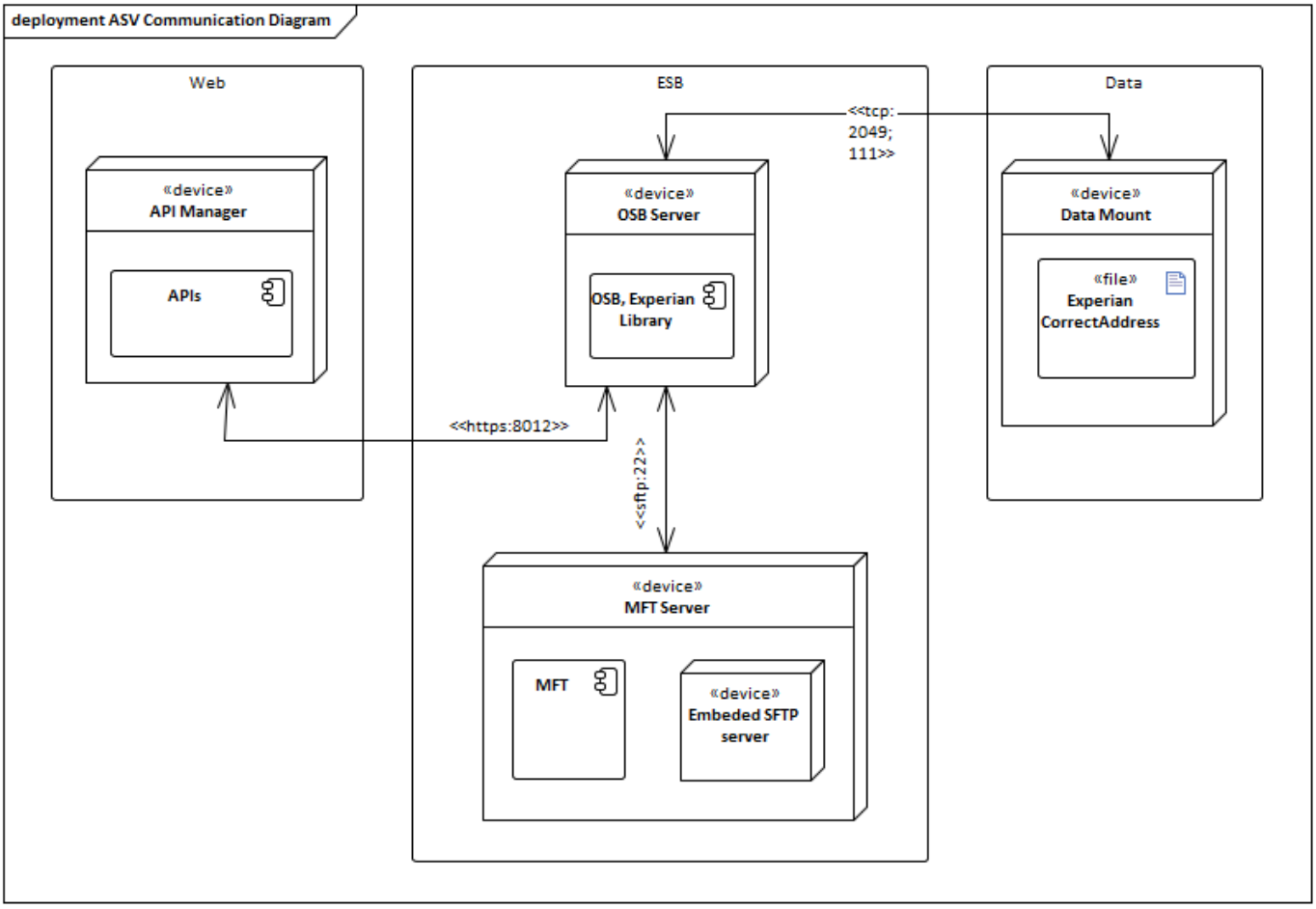
Communication Diagram

The following figure shows a communication diagram for the EAVS platform. The components are divided into three subnets:

- Web Trusted – The Web Trusted zone hosts the API Manager component of the ESB, and handles all incoming Hyper Text Transfer Protocol (HTTP)/ Secure HTTP (HTTPS) traffic.
- ESB subnet – The ESB zone consists of several components, primarily Oracle Service Bus, and Oracle Managed File Transfer. These two components handle the real-time and batch use cases respectively.
- Data subnet – The Data zone holds the Experian libraries and US and Canada address data.

The ASV in ESB design follows the N tier architecture with clear segregation of each layer with its dedicated components and responsibilities.

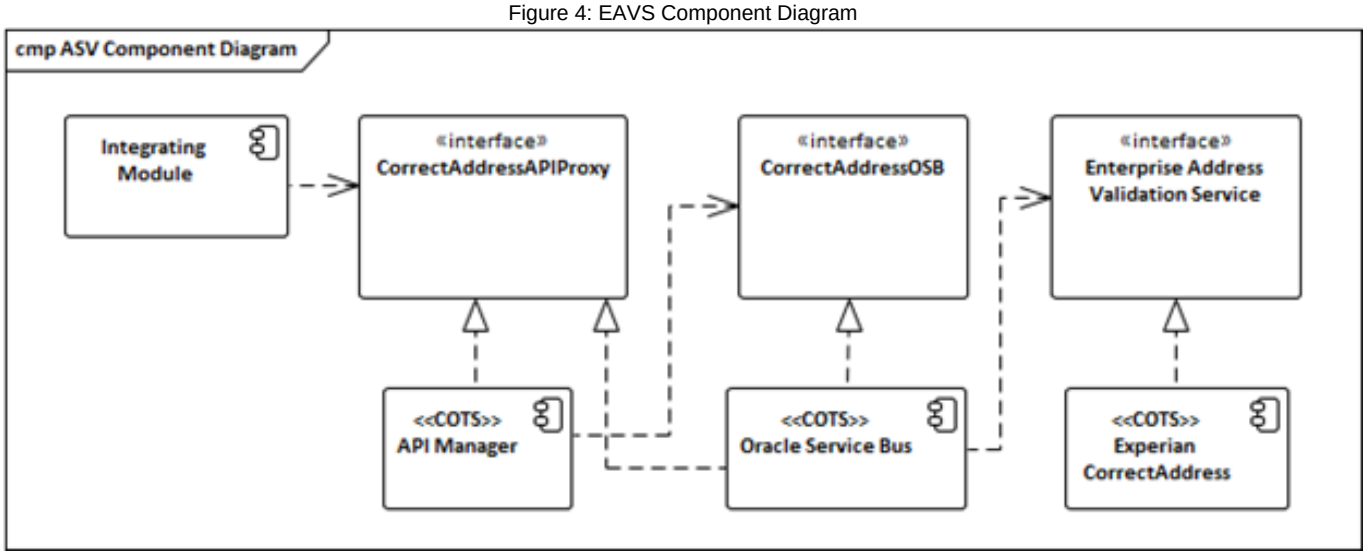
Figure 3: EAVS Communication Diagram



Component Diagram

The following figure shows the component diagram for EAVS. The diagram depicts how an integrating module invokes the process and subsequent dependencies.

Figure 4: EAVS Component Diagram



API Proxy Service

The CorrectAddressAPIProxy is an Oracle Service Bus (OSB) instance published as proxies to enable service discovery, versioning, and other capabilities of the Oracle API manager. The proxy delegates the call to the virtualization layer.

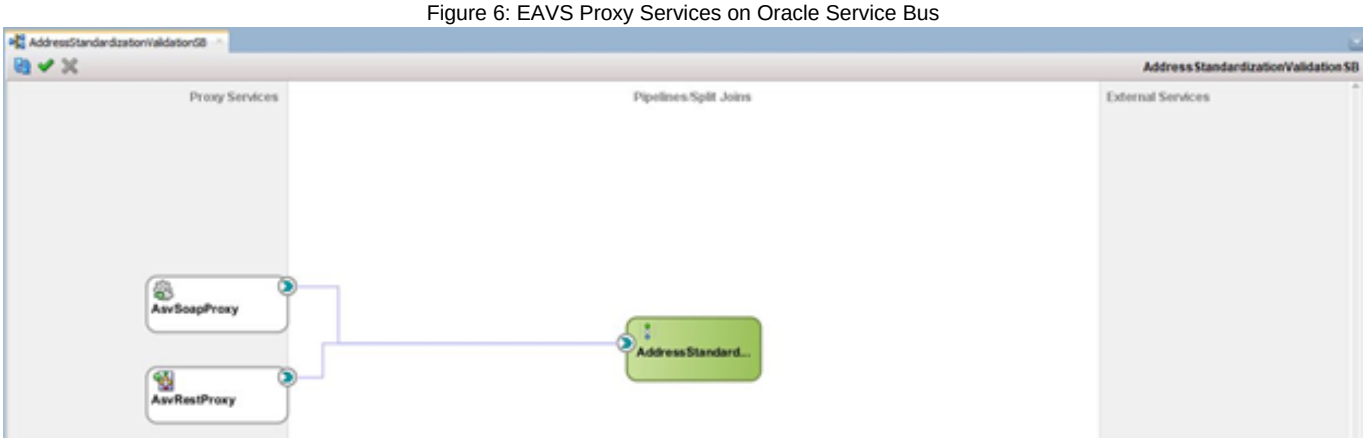
Figure 5: EAVS on API Manager

REDACTED DUE TO SECURITY CONCERNS

OSB Inbound/Outbound Instances

The CorrectAddressInboundOSB is an OSB instance used to virtualize the inbound call to the Enterprise Address Validation API. The service takes the request from CorrectAddressAPIProxy and applies validations and transformations to the request as needed. The OSB instance provides response back to the clients with validated CorrectAddress details. The OSB instance transforms the response message back to the canonical model.

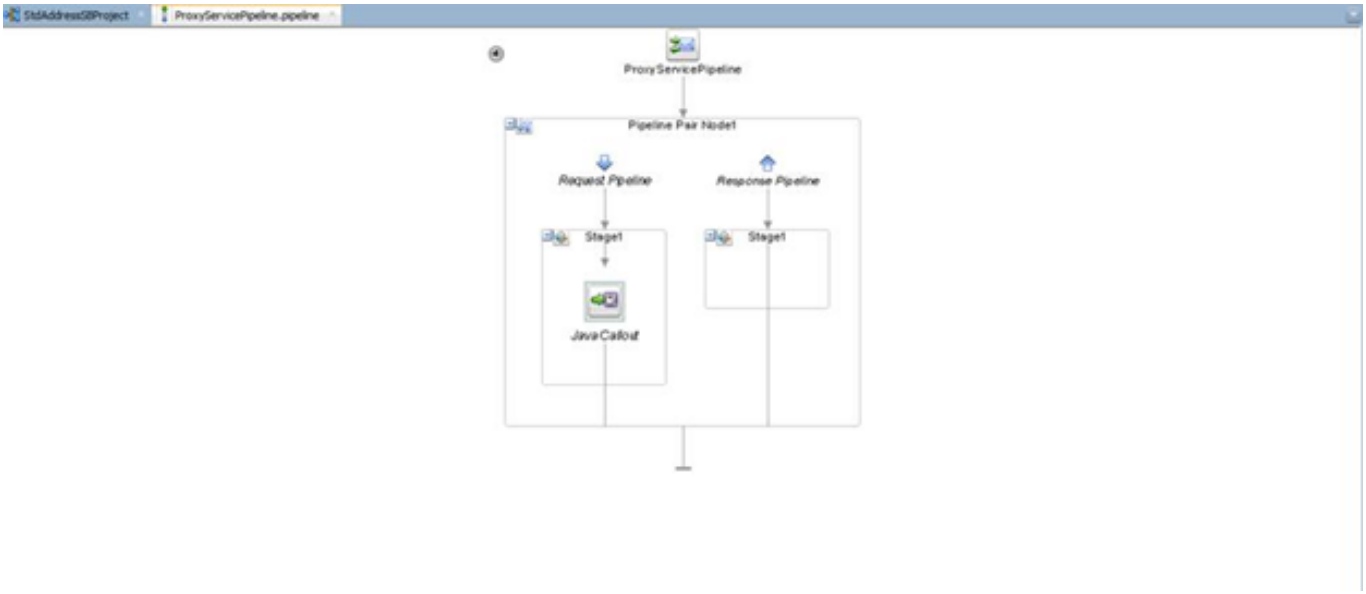
Figure 6: EAVS Proxy Services on Oracle Service Bus



CorrectAddress Java Callout

The application uses Experian CorrectAddress as the address standardization and validation solution. The Service Bus application uses the Java callout action to invoke the Java library from within the pipeline. The Java callout action takes input parameters required to invoke the Java methods and creates an object to pass to the library. The response object from the library in native format is translated into a canonical object by the Service Bus.

Figure 7: EAVS Java Callout in Pipeline



Experian CorrectAddress – Technical Features

The underlying component, Experian CorrectAddress provides the following features to effectively handle address validation and standardization across the HHS 2020 platform. It includes:

- Address verification (USA and Canada)
- Advanced name and address parsing
- Address standardization
- Recognition of addresses of major corporations
- ZIP® correction
- ZIP+4® Code appending
- DPVTM (Delivery Point Validation) for USA and PoC (Point of Call) verification for Canada
- LOT (Line of Travel) coding
- LACSLink™(Locatable Address Conversion System)
- SuiteLink™ coding
- Delivery point bar coding with Postnet barcode font included
- Carrier route codes
- Support for FIPS codes
- Geocoding Add-on
- Merge-Purge and Deduplication Add-on

Functions

The following primary functions are invoked as part of the underlying Java callout process.

CorrectAWorld: Validates and CASS-standardizes input address with Delivery Point Validation (DPV) and LACSLink™ processing. Verifies and corrects Canadian addresses. All near matches are returned when applicable. This function is applicable for the real-time use case.

RunCABatch: Accepts configuration file as input, runs a batch of addresses through CorrectAWorld function according to specifications in the configuration file. This function is applicable for the batch use case.

The complete list of functions provided by Experian CorrectAddress are listed here on Confluence: Experian CorrectAddress - Technical Capabilities

Return and Error Codes

Some of the important return codes return by Experian library have been listed here:

Table 1: Experian CorrectAddress Return Codes

Return Code	Description
1	Match found; four-digit ZIP add-on assigned.
> 1	Multiple possible results; no exact match made. Number of results is the value of return code.
• 99	No match found; the original input has been returned

Some of the important error codes return by Experian library have been listed here:

Table 2: Experian CorrectAddress Error Codes

Error Code	Description
07	Address non-deliverable; no add-on assigned. Delivery Point Validation (DPV) check failed.
11	No match; failed CASS multi-component rule; number of results is absolute value of return code. Several problems were found in the input address. Unable to match using CASS logic. Number of near matches returned is equal to the absolute value of the return code.
58	Address not standardized; invalid ZIP Code. Input ZIP code was incorrect. Unable to standardize address.

The complete list of return and error codes of the Experian CorrectAddress are listed on Confluence here: CorrectAddress Return Codes and Error Codes

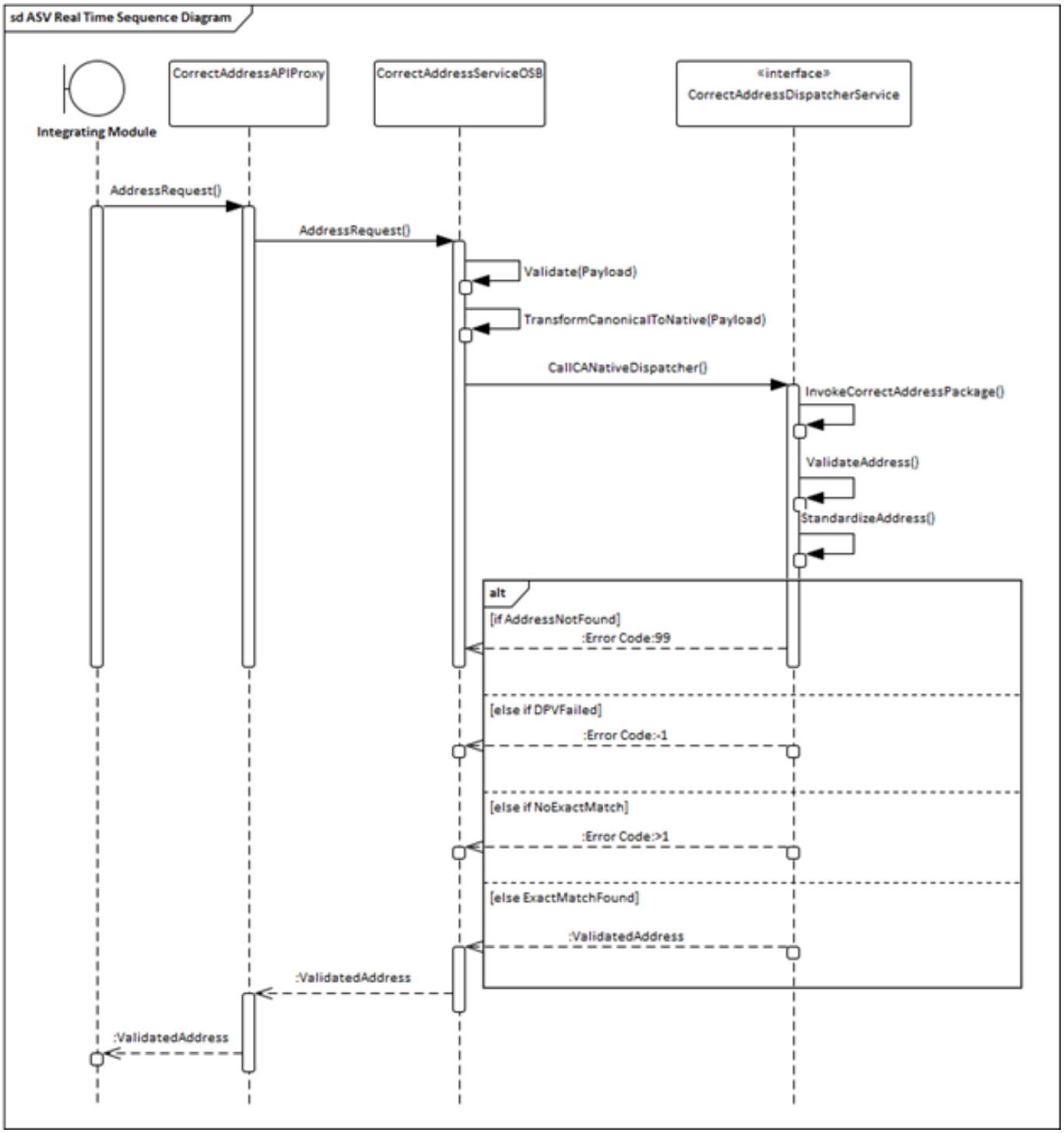
EAVS – Real-Time Validation Process

The real-time process exposes the CorrectAWorld function that validates and CASS-standardizes input address with Delivery Point Validation (DPV) and LACSLink™ processing information. As described in the component diagram, the service enablement of EAVS is achieved through the following Oracle Fusion Middleware components: Oracle API Manager, Oracle Service Bus. The benefits of this approach includes:

- Ability to expose the EAVS web service in both SOAP (Simple Object Access Protocol) and REST (Representational State Transfer) formats.
- Integrate with existing frameworks like Common Process Logging Framework developed for the enterprise.
- Create atomic and composite services to expose various Experian CorrectAddress functions as enterprise shared service.

The following figure shows the sequence of events with a few negative and positive alternate flows.

Figure 8: EAVS Real-Time Sequence Diagram



Sequence of Events

- The integrating module (Client) makes an address validation request via the API Proxy
- API Manager delegates the request to the OSB.
- OSB handles the validation and transformation both to and from CorrectAddressDispatcherService
- The CorrectAddressDispatcherService realizes the operations involving standardizing and validating the address.
- The sequence diagram shows a handful of alternate flows, few negative flows and a positive flow.
- The ValidatedAddress object is returned to the integrating module in canonical schema.

SOAP and REST Requests

The following figures show the real-time invocation of the EAVS using SOAP and REST protocols via the test automation tool, SOAPUI. The request payload in either cases use the canonical schema.

Figure 9: EAVS SOAP Request via SOAPUI

REDACTED DUE TO SECURITY CONCERNS

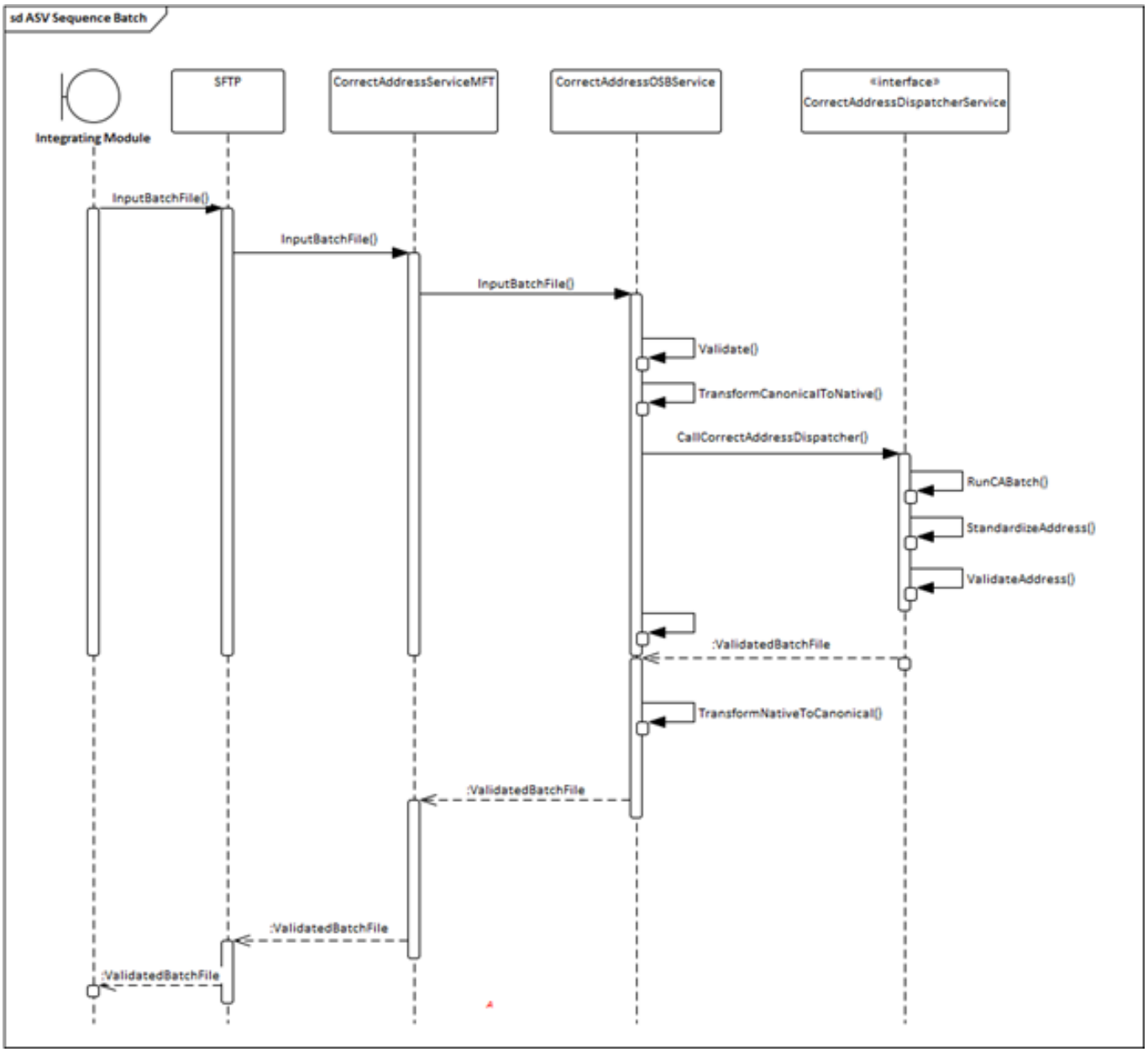
Figure 10: EAVS REST Request via SOAPUI

REDACTED DUE TO SECURITY CONCERNS

EAVS – Batch Validation Process

The EAVS batch validation process runs a batch of addresses through CorrectAWorld function according to specifications in the configuration file. A batch Input file is used when multiple address needs to be standardized and validated.

Figure 11: EAVS Batch Sequence Diagram



Sequence of Events

- The integrating module makes an address validation request (Batch) via the MFT hosted SFTP Server in canonical format
- OSB handles the validation and transformation both to and from CorrectAddressDispatcherService
- The CorrectAddressDispatcherService realizes the operations involving standardizing and validating the address in native Experian CorrectAddress format.
- ValidatedAddress batch CSV object is returned to the integrating module in canonical format.

Batch Process Input File (Canonical)

The following fields are on the header row of the input CSV. These fields are derived from the standardized canonical schema.

- RecordID – Unique Record Identifier
- Line 1 – Address Line 1
- Line 2 – Address Line 2
- Line 3 – Address Line 3
- City – The name of the city, town, village, or other community or delivery center

- County – A region created by territorial division for the purpose of local government. In the USA, a county (or parish in Louisiana) is the largest administrative district within a state
- Country – a nation as commonly understood or generally accepted
- StateCode – 2 letter alphabetic state codes for the US states (if applicable)
- PostalCode – A postal/zip code designating a region defined by the postal service
- PhoneNumber – Phone number in standardized XXX-XXX-XXXX format

Figure 12: EAVS Canonical Input in CSV

	A	B	C	D	E	F	G	H	I	J
1	RECORDID	LINE1	LINE2	LINE3	CITY	COUNTY	COUNTRY	STATECODE	POSTALCODE	PHONENUMBER
2	1	1001 bayhill drive			san bruno			ca	94066	111-111-1111
3										

Batch Process Input File (Native)

The canonical input file is transformed into the native Experian CorrectAddress supported format with the following header row columns:

- RecordID – Unique Record Identifier
- Address1 – Address Line 1
- Address2 – Address Line 2
- City – The name of the city, town, village, or other community or delivery center
- State– 2 letter alphabetic state codes for the US States (if applicable)
- Zip – A postal/zip code designating a region defined by the postal service
- PhoneNumber – Phone number in standardized XXX-XXX-XXXX format

Figure 13: EAVS Native Input in CSV

	A	B	C	D	E	F	G	H
1	RecordID	Address1	Address2	City	State	Zip	Phone Number	
2	1	1001 bayhill drive		san bruno	ca	94066	111-111-1111	

Batch Process Output File (Native)

The batch file, when processed by the Experian CorrectAddress function, it produces additional columns in the output CSV as following:

- RecordID – Unique Record Identifier (as given in the input file)
- Address1 – Address Line 1 (as given in the input file)
- Address2 – Address Line 2 (as given in the input file)
- City – The name of the city, town, village, or other community or delivery center (as given in the input file)
- State– 2 letter alphabetic state codes for the US states (if applicable) (as given in the input file)
- Zip – A postal/zip code designating a region defined by the postal service (as given in the input file)
- PhoneNumber – Phone number in standardized XXX-XXX-XXXX format (as given in the input file)
- DeliveryLine1 – Standardized Address Line 1
- DeliveryLine2 – Standardized Address Line 2
- City – Standardized city name
- State – Standardized state name (2 letter alphabetic code)
- ZipAddon – Zip With 4 digit Add-on
- DPVFlags – Delivery Point Verification flag
- County Name – Standardized County name
- Return Codes – Experian CorrectAddress Return code (check Section 2.2)
- Error Codes – Experian CorrectAddress Error code (check Section 2.2)

Figure 14: EAVS Native Output in CSV

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	RECORDID	ADDRESS1	ADDRESS2	CITY	STATE	ZIP	PHONE NUMBER	DELIVERY	DELIVERY	*CITY	*STATE	ZIPADDON	DPV FLAGS	COUNTY NAME	RETURN CODE	ERROR CODES
2	1	1001 bayhill drive		san bruno	ca	94066	111-111-1111	1001 BAYHILL DR		SAN BRUNO	CA	94066-3062	DN Y	SAN MATEO	1	1251

Batch Process Output File (Canonical)

The EAVS, then transforms the native output from Experian into the canonical format. The canonical output includes the following header row columns:

- RecordID – Unique Record Identifier (as given in the input file)
- Line 1 – Standardized Address Line 1
- Line 2 – Standardized Address Line 2
- Line 3 – Standardized Address Line 3
- City – Standardized name of the city, town, village, or other community or delivery center (as given in the input file)
- County – Standardized County name
- StateCode – Standardized 2 letter alphabetic state codes for the US states (if applicable) (as given in the input file)
- PostalCode – A postal/zip code designating a region defined by the postal service (as given in the input file)
- PhoneNumber – Phone number in standardized XXX-XXX-XXXX format (as given in the input file)
- Return Codes – Experian CorrectAddress Return code (check Section 2.2)
- Error Codes – Experian CorrectAddress Error code (check Section 2.2)

Figure 15: EAVS Canonical Output in CSV

	A	B	C	D	E	F	G	H	I	J	K	L
1	RECORDID	LINE1	LINE2	LINE3	CITY	COUNTY	COUNTRY	STATECODE	POSTALCODE	PHONENUMBER	RETURNCODE	ERRORCODES
2	1	1001 BAYHILL DR			SAN BRUNO	SAN MATEO		CA	94066	111-111-1111	1	1251
3												

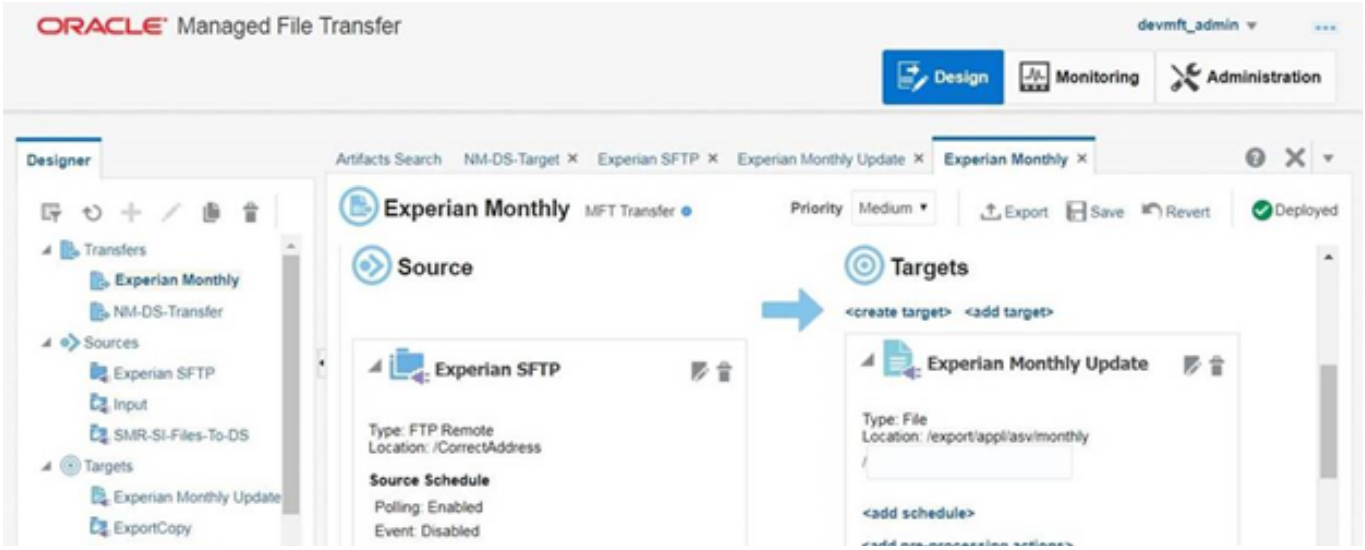
EAVS – Administrative Process

A subset of data files provided by Experian CorrectAddress FTP server, especially CASS data follows a release cycle and it expires at the end of each cycle (July 31st every odd year). Data files used by CASS also expire after 105 days from the release date of the data files. As part of the administrative process, the data files are replenished using a scheduler job configured on the Oracle MFT tool.

Configuration and Scheduling

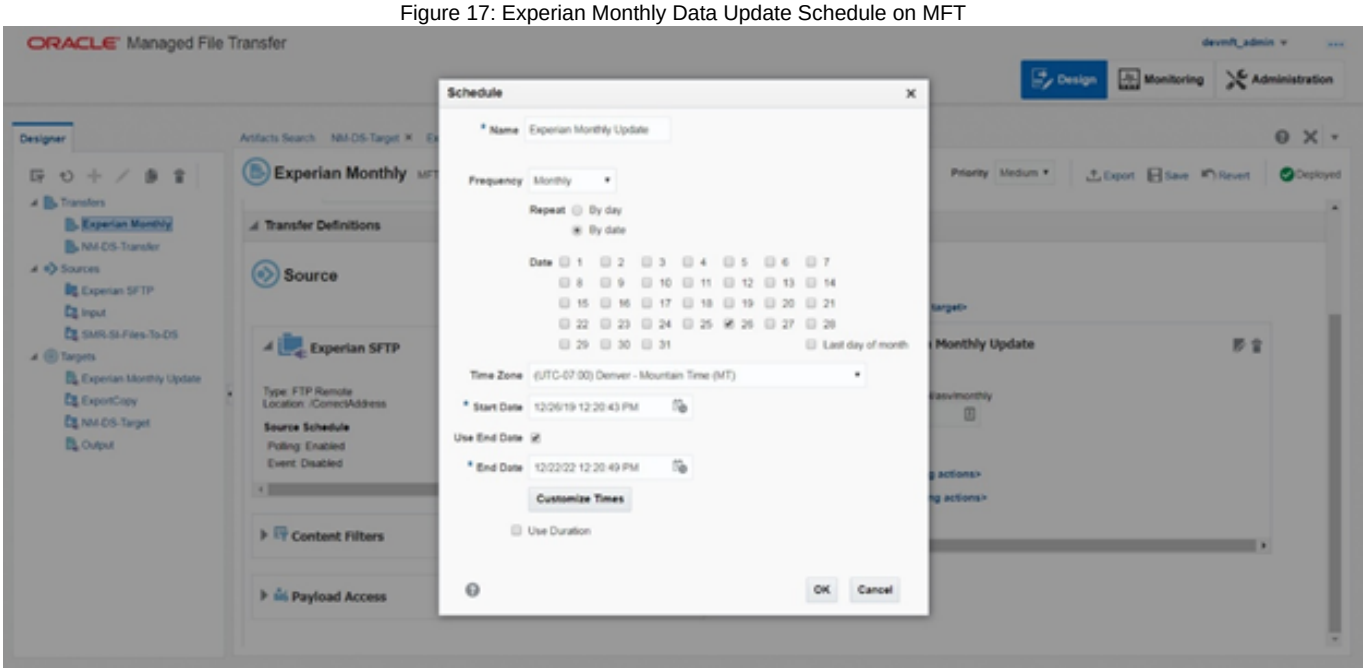
A CorrectAddress MFT job is created to update the local data files with the updated files from Experian SFTP server.

Figure 16: Experian Monthly Data Update Configuration on MFT



The following image shows the monthly schedule created to automate this process to run on a monthly schedule.

Figure 17: Experian Monthly Data Update Schedule on MFT



Sequence of Events

- The CorrectAddress MFT Job running on a schedule will wake up on monthly basis and initiate the operation.
- The operation involves retrieving the latest data files from the SFTP server hosted on Experian.
- Prior to replacing the latest files from SFTP, the MFT will perform a backup operation.
- The backup operation is to ensure that if the latest files lead to corruption or misconfiguration of existing data, we can revert to previous operational state.

- ASV - Security Checklist for QATProm06
- ASV Communication Diagram in QAT Environment
- ASV QAT Platform Firewall and Environment Details
- Design - Sample Scenarios