

Data Synchronization Engine

Interface Control Document (DSE ICD)

DSE Description

The Data Sync Engine (DSE) is a non-intrusive application with a goal to keep SDWIS/State database in sync with the CMDP database. First, the CMDP database will be updated with inventory and legal entity data changes in the SDWIS/State database. Secondly, the SDWIS/State database will be updated with sampling data changes in the CMDP database. The DSE is scalable, designed to copy high volumes of sample results from the CMDP to SDWIS/State, and to copy bulk changes to inventory and laboratory data from SDWIS State to the CMDP. The overall process is described in Figure 1 below.

The solution will be deployed by the State' database administrator (DBA) locally on the state primacy agency's server (same server as SDWIS/State). Note: the DSE will not change anything within the current SDWIS/State database or the State's existing XML Sampling business process. See "State Steps" section below for more details.

The DSE features three components.

The first component pulls changed inventory and legal entity data **from SDWIS/State to CMDP** (Fig. 1: Pushing LE & Inventory Data to CMDP from S/State). The DSE can connect to both Oracle and MS SQL Server databases with only "connection string" changes. LE and inventory data will only move from SDWIS/State to CMDP. Note: there are no external connections through the State's firewall. See *Appendix B for legal entity and Inventory data transferred.*

The second component pulls changed sampling data **from CMDP to XML Sampling to SDWIS/State** (Fig. 1: Pulling Sampling and Operational Data from CMDP to S/State). See *Appendix C for a listing of the SDWIS/XML Sampling Schemas it will use.*

The third component is the administration console. This tracks the status of web calls by the DSE in both directions (to CMDP and to the State) and records metrics on data sent (e.g. 100 samples attempted, 99 successful and 1 failure).

Diagram of Overall Technical Solution

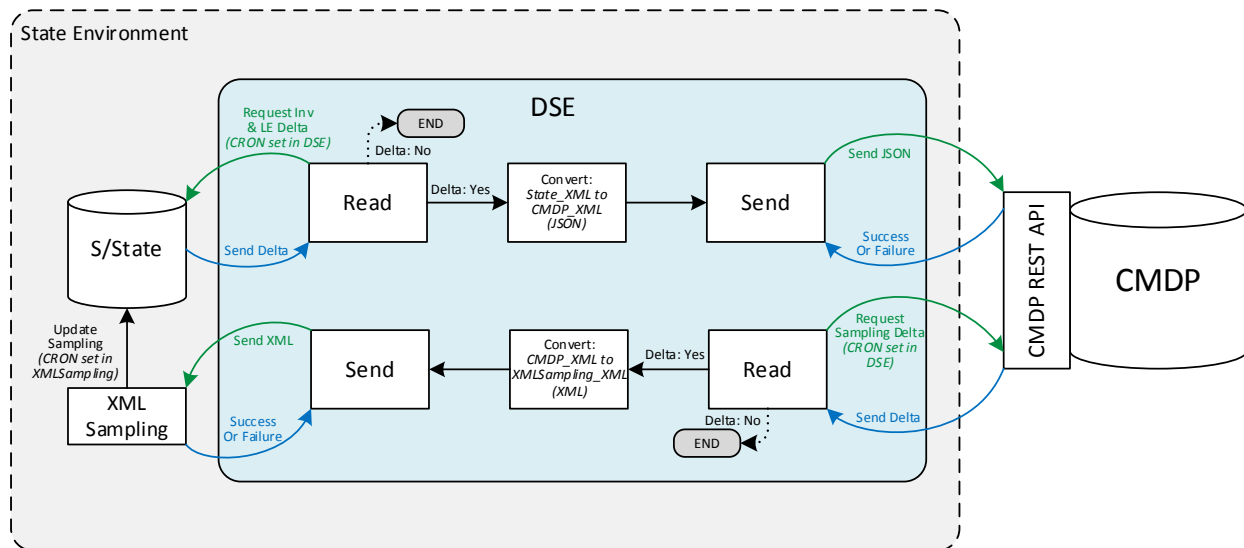


Fig.1 DSE Architecture

Note: Details on the CMDP REST API in the DSE Architecture diagram above can be found in the CMDP Web Services State DB Document on the CMDP Zendesk site.

These data transfers (in both directions) are triggered by a CRON job which “wakes up” the DSE at scheduled intervals (e.g. 1am every day, top of every hour, etc.). Once awake, the DSE cycles through the follow sequence:

Stage 1: SDWIS/State to CMDP (independent of Stage 2)

1. The DSE queries the SDWIS/State database for inventory and legal entity changes since the last time the DSE ran.
2. If changes exists, the DSE moves on to conversion as described in Step 3. If not, Stage 1 ends.
3. The DSE transforms the SDWIS/State database result set to an outgoing JSON format that is compatible with the CMDP schema.
4. The DSE transmits the data to the CMDP REST endpoint (the web service method that has been designed and built to accept the changed inventory and legal entity data into CMDP).
5. Feedback (success/failure) is captured and displayed in DSE admin console (log file).

Stage 2: CMDP to SDWIS/State (independent of Stage 1)

1. The DSE queries the CMDP database for new Sample Jobs that have been electronically signed and submitted since the last time the DSE ran.
2. If new jobs exist, the DSE moves on to conversion as described in Step 3. If not, Stage 2 ends.
3. The DSE transforms the CMDP new jobs in JSON to a compatible XML Sampling format per the XML schemas defined in XML Sampling 3.2.
4. The DSE sends the new Sample Jobs to XML Sampling, which then performs its current functions of migrating the XML file contents into the SDWIS/State database.

5. Feedback (success/failure) is captured and displayed in DSE admin console (log file). Note: Sample jobs successfully sent to XML Sampling via the DSE will appear as “Accepted” in the CMDP application.

State Application Requirements

SDWIS/State 3.32 (May also support 3.2 or older – stated in XML Sampling 3.21 installation Guide)

SDWIS/XML Sampling 3.21 (should also work with XML Sampling 3.1 and 3.0)

SDWIS/State Server with JDK 1.7

Test Environment: SDWIS/State Mirror environment that includes all relevant tables (inventory, legal entity, samples, etc.) necessary to produce the JSON files for CMDP and receive XML files via XML Sampling 3.1 for all data elements listed in Appendices B and C.

Security: State administrator will need to use SCS account to interact with production DSE.

State steps

1. Create a new schema in the SDWIS State database with a name of DSE_MATRIX. The permissions granted to the schema should be CRUD. *Note: New schema does not touch existing SDWIS/State schema.*

- a. CREATE USER DSE_MATRIX
- b. IDENTIFIED BY DSE_MATRIX
- c. DEFAULT TABLESPACE [PRIME_DATA] (State will replace PRIME_DATA with their table space name)

States will have to complete above steps (a-c) as part of DSE_MATRIX schema creation process

Attain will manage steps (d-g) as a part of the DSE Installation:

- d. TEMPORARY TABLESPACE TEMP
- e. PROFILE DEFAULT
- f. ACCOUNT UNLOCK
- g. Grant connect, resource, DBA to DSE_MATRIX

2. Download DSE application

- a. DSE can be found on the CMDP Help Desk website at: <https://cmdp.zendesk.com>
- b. Download installation files from CMDP Help Desk based on the States database type (Oracle/MS SQL Server).

3. Install using the instructions for your RDBMS type (MS SQL or Oracle)

- a. Load WAR (dseWebAdmin.war) file on same server as SDWIS/State.

- b. Database config: State DBA needs to config DSE (see DSE configuration) with State Database. Follow steps to give state DB location, login, transaction interval etc.
- c. States provide folder where XML sampling is monitoring for new data (aka "Input Folder" in SDWIS/XMLSampling configuration). See Appendix A for screen shot.

DSE Configuration

What's configurable now:

- CRON frequency:
 - found in dswWebContext.xml
 - *CRON frequency of XMLSampling is also configurable. See Appendix A.*
- URL specifications for CMDP
 - found in url_specs.properties
- Schema name of database for DSE
- Schema name of state from where the SQL is being read

TO BE (in production v1, we will most probably have following areas configurable)

- Output folders
 - will represent the folder where DSE will put the XML file for state application to pick up from
- Logging criteria configuration
- Setting Last Sync Date to sync all Inventory and Laboratory Profile information.

Appendix A - Input Folder for XML Sampling

The screenshot shows the 'Configuration' page for 'SDWIS/XMLSampling'. The page has a blue header with the title and a blue sidebar on the left with navigation links: Home, Scheduler, Submit Sampling File, View Jobs, Configuration, and View Log Files. The main content area is white and contains a form with the following fields:

- **Primary Contact Email Address: [Text Input]
- **Database: [Oracle] (Dropdown)
- **EmailServer: [Text Input]
- EmailUserId: [Text Input]
- EmailPassword: [Text Input]
- **XML Sampling ServerAddress: [localhost] [LookUp]
- **XML Sampling ServerPort: [8080] [LookUp]
- **URL Mapping: [XMLSampling] [LookUp]
- XMLSamplingEmailAccount: [Text Input]
- **Generate SamplingPointIndicator: [Yes] (Dropdown)
- **JobRequesterUserID: [ADMIN] (Text Input)
- **Input Folder: [Text Input] (highlighted with a red box)**

At the bottom of the form, there are 'OK' and 'Cancel' buttons.

Note: The Input Folder should contain a user-specified value, but defaults to Input Folder under the XML Sampling context.

Appendix B - Legal Entity and Inventory Data Copied from SDWIS/State to CMDP

The following data is transferred from SDWIS State to CMDP using the DSE. This information needs to be correct in SDWIS State before transferring to CMDP. The DSE will reject incorrect data based on CMDP validations explained in the CMDP User Manual.

Water System Profile (SDWIS STATE Tables and Columns)

Table Name	Column Name
TINWSYS	NUMBER0
	tinwsys_is_number
	NAME
	ALTERNATE_ST_NUM
	LOCAL_NAME
	STAGE2_CDS_ID
	GRND_WTR_RATIO
	GRND_WTR_PUR_RATIO
	GRND_WTR_UDI_RATIO
	GRND_WTR_UDI_PURCH
	SURF_WTR_RATIO
	SURF_WTR_PUR_RATIO
	PWS_ST_TYPE_CD
	D_FED_PRIM_SRC_CD
	owner_type_code
	activity_status_cd
	ACTIVITY_DATE
	D_ST_PRIM_SRC_CD
	MEMO_TEXT
	D_INITIAL_USERID
	D_LAST_UPDT_TS
	D_USERID_CODE
	TINWSYS_ST_CODE
TINWSLEC	Type_Code
	Active_Ind_Cd
	Tinwsys_Is_Number
	Tinwsys_St_Code
	Tinlgent_Is_Number
	Tinlgent_St_Code
TINLGENT	Tinlgent_Is_Number
	Tinlgent_St_Code
	Activity_Status_Cd
	D_LAST_UPDT_TS
	NAME

	ADDR_LINE_ONE_TXT
	ADDR_LINE_TWO_TXT
	ADDRESS_CITY_NAME
	ADDRESS_STATE_CODE
	ADDRESS_ZIP_CODE
TINWSIN	INDICATOR_VALUE_CD
	INDICATOR_NAME
	INDICATOR_DATE
	TINWSYS_IS_NUMBER
TINPOPSV	Avg_Daily_Cnt
	Type_Code
	Tinwsys_Is_Number
	Tinwsys_St_Code
TINAOPRD	Tinwsys_Is_Number
	Tinwsys_St_Code
	Eff_Begin_Dt
	Eff_End_Dt

Facility Profile ([SDWIS STATE Tables and Columns](#))

Table Name	Column Name
TINWSYS	TINWSYS_IS_NUMBER
	TINWSYS_ST_CODE
	NUMBER0
TINWSF	TINWSYS_IS_NUMBER
	TINWSYS_ST_CODE
	ST_ASGN_IDENT_CD
	TYPE_CODE
	NAME
	LOCAL_NAME
	ACTIVITY_STATUS_CD
	Activity_Date
	AVAILABILITY_CODE
	WATER_TYPE_CODE
	WATER_TYPE_CODE_DT
	TREATMENT_STAT_CD
	Tinwsf_Is_Number
	Tinwsf_St_Code
	D_LAST_UPDT_TS
TINWSFC	Tinwsf_Is_Number
	Tinwsf_St_Code
	active_ind_cd

	Tinlgent_Is_Number
	Tinlgent_St_Code
TINLGENT	Tinlgent_Is_Number
	Tinlgent_St_Code
	D_LAST_UPDT_TS
	ADDR_LINE_ONE_TXT
	ADDR_LINE_TWO_TXT
	ADDRESS_CITY_NAME
	ADDRESS_STATE_CODE
	ADDRESS_ZIP_CODE
tinpvals	TEXT_VALUE
	CODE_NAME

Sampling Point Profile (SDWIS State Tables and Columns)

Table Name	Column Name
TSASMPPT	tinwsf0is_number
	Tinwsf0st_Code
	D_LAST_UPDT_TS
	IDENTIFICATION_CD
	DESCRIPTION_TEXT
	ACTIVITY_DATE
	SOURCE_TYPE_CODE
	ACTIVITY_STATUS_CD
	TYPE_CODE
	D_USERID_CODE
	D_INITIAL_TS
	D_INITIAL_USERID
TINWSF	tinwsf_is_number
	Tinwsf_St_Code
	ST_ASGN_IDENT_CD
TINWSYS	Tinwsys_Is_Number
	Tinwsys_St_Code
	Number0

Laboratory Profile (SDWIS State Tables and Columns)

Table Name	Column Name
TSALAB	TSALAB_IS_NUMBER
	TSALAB0ST_CODE
	D_LAST_UPDT_TS
	D_USERID_CODE

	D_INITIAL_USERID
	LAB_ID_NUMBER
	TSALAB_ST_CODE
TSALLEA	ACTIVE_IND_CODE
	TYPE_CODE
TINLGENT	TINLGENT_IS_NUMBER
	TINLGENT_ST_CODE
	ADDRESS_STATE_CODE
	COUNTRY_CODE
	ADDR_LINE_ONE_TXT
	ADDR_LINE_TWO_TXT
	ADDRESS_CITY_NAME
	ADDRESS_STATE_CODE
	ADDRESS_ZIP_CODE
	INT_POST_CD
TINLGCOM	Electronic_Address
	tingent_is_number
	purpose_code
	tingcom_is_number
	Phone_number
	Electronic_Address

Referential data mapping from SDWIS State

State Table	Field Name in State Table	CMDP Ref Category
TINWSYS	PWS_ST_TYPE_CD	WS_STATE_TYPE
TINWSYS	D_ST_PRIM_SRC_CD	WS_STATE_WATER_SRC_TYPE
TINWSYS	ACTIVITY_STATUS_CD	WATER_SYSTEM_ACTIVITY_STATUS
TINWSF	ACTIVITY_STATUS_CD	FACILITY_ACTIVITY_STATUS
TSASMPPT	SOURCE_TYPE_CODE	SMP_PNT_WATERTREAT_STAT
TINLGENT	ADDRESS_CITY_NAME	US_CITY
TINWSLEC	TYPE_CODE	CONTACT_TYPE
TINPVALS	TYPE_CODE	CONTACT_TYPE
TINLGCOM	PURPOSE_CODE	LE_EMAIL_TYPE
TINLGENT	TYPE_CODE	LE_ORG_TYPE
TINLGCOM	PURPOSE_CODE	LE_PHONE_TYPE

Appendix C - XML Sampling Schemas Used

CMDP data elements will be compiled into an XML file compatible with the following SDWIS/XML Sampling Schemas:

- SDWIS_eDWR_v3.0.xsd
- SDWIS_Summary_v3.0.xsd
- SDWIS_SummaryResult_v3.0.xsd
- SDWIS_MDBPSummary_v3.0.xsd

Note: Please refer to Sampling Results to XML Sampling Application document and the Transmitting CMDP Operational Data to XML Sampling via DSE document (in Zendesk) for details on mappings between CMDP and XML Sampling.

Appendix D – DSE Setup Instructions.

Please refer to the “DSE Setup Guide.docx” document for instruction on setting up DSE.

Appendix E.

The following logic was implemented to populate “SampleRuleCode” tag.

- 1) If there is a single Analyte in Microbial Sample
 - A) If the Analyte is one of (3100, 3014, 3002, 3028), then "SampleRuleCode" tag = 'TC'
 - B) Else "SampleRuleCode" tag = 'GE'
- 2) If there are multiple Analytes in Microbial Sample
 - A) If all Analytes in Microbial sample are one of (3100, 3014, 3002, 3028), then "SampleRuleCode" tag = 'TC'
 - B) Else "SampleRuleCode" tag = 'GE'
- 3) If there is a single Analyte in Chems/Rads Sample
 - A) If the Analyte is one of (1020,1022), then "SampleRuleCode" tag = 'PB'
 - B) Else "SampleRuleCode" tag = 'GE'
- 4) If there are multiple Analytes in Chems/Rads Sample
 - A) If all Analytes in Chems/Rads Sample are one of (1030,1022), then "SampleRuleCode" tag = 'PB'
 - B) Else "SampleRuleCode" tag = 'GE'

The following logic was implemented to populate the “ComplianceIndicator” tag

- 1) If Sample_Type in ('RT', 'TG', 'RP', 'MR', 'CO') then Compliance_Indicator = 'Y'
Else Compliance_Indicator = 'N'