# 



Master Files Data Integration Interface Specifications  
- DTAC

Table of Contents

[1. Overview 6](#_Toc32217675)

[1.1 Interface Diagram 6](#_Toc32217676)

[1.2 File Feed List 6](#_Toc32217677)

[1.3 File Transfer Mechanism 7](#_Toc32217678)

[1.4 ODS Table Naming convention 8](#_Toc32217679)

[2. Assumptions 9](#_Toc32217680)

[3. Interface Details 10](#_Toc32217681)

[3.1 Details of Source Feed - MASTER FILES – USER\_DATA 10](#_Toc32217682)

[3.1.1 Source Feed Files 10](#_Toc32217683)

[3.1.2. Source Feed wise - Record Format 12](#_Toc32217684)

[3.1.3. Control Feed - Record Format 13](#_Toc32217685)

[3.1.4. Privacy Columns/Business Rules 14](#_Toc32217686)

[4. Interface SLAs & Methods 15](#_Toc32217687)

[4.1 Availability 15](#_Toc32217688)

[File Validation 15](#_Toc32217689)

[4.2.1 Duplicate File Check 16](#_Toc32217690)

[4.2.2 Missing file from Source 16](#_Toc32217691)

[4.2.3 Missing records or Incomplete File 16](#_Toc32217692)

[4.2.4 Alerts 16](#_Toc32217693)

[4.2.5 Invalid file format 17](#_Toc32217694)

[4.2.6 Erroneous data 17](#_Toc32217695)

[4.3 Source System Changes 17](#_Toc32217696)

[4.4 Unable to Collect or receive files from source 18](#_Toc32217697)

[4.5 Unavailability of Data Collection & Integration Layer 18](#_Toc32217698)

[4.6 Data Quality and Timeliness 18](#_Toc32217699)

[4.7 Exception Handling 18](#_Toc32217700)

[5. Appendix 19](#_Toc32217701)

[5.1. Email Notification for missing file: 19](#_Toc32217707)

**Revision History**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Version** | **Revision Date** | **Author / Reviewer** | **Summary of changes** | **Remarks** |
| 1.0 | 07.02.2020 | Hirak Sen | First Draft Version | For DTAC Review |
|  |  |  |  |  |
|  |  |  |  |  |

**Review and Approval**

|  |  |  |  |
| --- | --- | --- | --- |
| **Person** | **Role** | **Responsibility** | **Date** |
| Vorraluck Sarechuer Vorraluck.sarechuer@datac.co.th | Delivery Manager -DTAC |  |  |
| Porntip Soponnchai PorntipS@dtac.co.th | Sr Manager - DTAC |  |  |
| Rohit Umaraw  Rohit.umaraw@wipro.com | Manager - Wipro |  |  |

|  |  |
| --- | --- |
| Project Name | TELENOR - AEP |
| Account | Telenor - DTAC |
| Current Version | 1.0 |
| Author | Hirak |

Referenced Documents

| **#** | **Document Name** | **Document Description** | **Document Type/Drop#** | **Version** |
| --- | --- | --- | --- | --- |
| 1 | BUName\_SourceSystemSummary\_DTAC\_V 0.1.3.xlsx | This document captures the source feed technical details and its data structures modified for priority 1.1. feeds | Excel file | 1.3 |

Terminologies & Acronyms

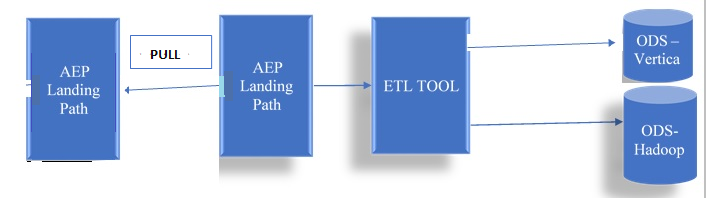
| **Acronyms Used** | **Description** |
| --- | --- |
| ODS | Operational Data store |
| AEP | Analytics Enablement Program |
|  |  |

# Overview

This document contains information and specification for all Master feeds (Business Team + Tech Team) Data Integration into ODS layer. The purpose of this document is to describe the design of output interfaces from Master files mainly CSV/TXT based to AEP platform (Data Integration).

The specification document will capture all the source feed details, their frequencies, naming conventions and their corresponding rules like exception handling, transformation rule, filtration rule if any, surrogate key and encryption logic.

## 1.1 Interface Diagram



Source

Server

## 1.2 File Feed List

The below table captures the list of files which will be integrated for Master Feeds as part of first phase of ODS go-live.

| **#** | **Feed Name** | **File Name** | **File Format** | **Source System Notification Mail\_Group** | **Source System Contact Point** |
| --- | --- | --- | --- | --- | --- |
| 1 | CBS\_ACCTTYPE\_RANGE\_DTN | cbs\_accounttype\_range | Csv | BSS-BILLING OPERATION-MEDIATION <BSS-BILLINGOPERATION-MEDIATION@dtac.co.th> | Chotevit@dtac.co.th 0894402238  pattranuch.nakprasert@dtac.co.th 0945601583  Karntima@dtac.co.th 0850553961  attaphon.ongvisit@dtac.co.th 0873200700 |
| 2 | CBS\_OFFER\_RANGE\_DTN | cbs\_offer\_range | Csv | BSS-BILLING OPERATION-MEDIATION <BSS-BILLINGOPERATION-MEDIATION@dtac.co.th> | Chotevit@dtac.co.th 0894402238  pattranuch.nakprasert@dtac.co.th 0945601583  Karntima@dtac.co.th 0850553961  attaphon.ongvisit@dtac.co.th 0873200700 |
| 3 | CBS\_PRODUCT\_RANGE\_DTN | cbs\_product\_range | csv | BSS-BILLING OPERATION-MEDIATION <BSS-BILLINGOPERATION-MEDIATION@dtac.co.th> | Chotevit@dtac.co.th 0894402238  pattranuch.nakprasert@dtac.co.th 0945601583  Karntima@dtac.co.th 0850553961  attaphon.ongvisit@dtac.co.th 0873200700 |
| 4 | CBS\_RESULT\_CODE\_DTN | cbs\_result\_code | csv | BSS-BILLING OPERATION-MEDIATION <BSS-BILLINGOPERATION-MEDIATION@dtac.co.th> | Chotevit@dtac.co.th 0894402238  pattranuch.nakprasert@dtac.co.th 0945601583  Karntima@dtac.co.th 0850553961  attaphon.ongvisit@dtac.co.th 0873200700 |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |

| **#** | **Feed Name** | **File Name** | **File Format** | **Source System Notification Mail\_Group** | **Source System Contact Point** |
| --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |

## 1.3 File Transfer Mechanism

1. Business users and Tech users will place CSV or TXT files to ingest into ODS layer. There are feeds doesn’t have control file associated. Following are the lists of files:

| **#** | **Feed Name** | **Feed File Name** | **File Format** | **Control File Name** | **Control File Format** |
| --- | --- | --- | --- | --- | --- |
| 1 | CBS\_ACCTTYPE\_RANGE\_DTN | cbs\_accounttype\_range | Csv | Control File NOT Available | NA |
| 2 | CBS\_OFFER\_RANGE\_DTN | cbs\_offer\_range | Csv | Control File NOT Available | NA |
| 3 | CBS\_PRODUCT\_RANGE\_DTN | cbs\_product\_range | csv | Control File NOT Available | NA |
| 4 | CBS\_RESULT\_CODE\_DTN | cbs\_result\_code | csv | Control File NOT Available | NA |

| **#** | **Feed Name** | **Feed File Name** | **File Format** | **Control File Name** | **Control File Format** |
| --- | --- | --- | --- | --- | --- |
| 1 |  |  |  |  |  |
| 2 |  |  |  |  |  |
| 3 |  |  |  |  |  |
| 4 |  |  |  |  |  |

1. There will be separate folder created for each day where the files will be Pulled inside AEP- Landing zone. Folder structure as below:

* For Raw Data files path should be
* /<root>/FIX\_DATA/CBS/DTAC/<YYYYMMDD>/
* /<root>/FIX\_DATA/CBS/DTN/<YYYYMMDD>/
* /<root>/USER\_DATA/CBS/DTAC/<YYYYMMDD>/
* /<root>/USER\_DATA/CBS/DTN/<YYYYMMDD>/
* For Control File path should be - NA

1. As soon as file gets picked for processing, the file\_copy\_wrapper script will move into working folder path and from there Loader\_util will move the feed file to HDFS feeds path. Subsequently Talend file validation scala job and vertica load jobs moves into DB layer. Once the file is processed successfully, the original file will be moved to ‘Archive’ folder where the file will be retained as per data retention period. Archive folder will contain all the raw files.
2. Few CSV feeds has header & no footer available inside the files.
3. Respective File owner is responsible to provide the correct data.
4. Available Source file path @ EDW Landing zone is mentioned in the below table-

| **#** | **Feed Name** | **File Name** | **Source File path - PROD** | **Source File path - SIT** | **Source File path-DEV** |
| --- | --- | --- | --- | --- | --- |
| 1 | CBS\_ACCTTYPE\_RANGE\_DTN | cbs\_accounttype\_range.csv | /EDW/nfsedw105/USER\_DATA/CBS/DTN/<YYYYMMDD>/cbs\_accounttype\_range.csv | /SIT/EDW/USER\_DATA/CBS/DTN/<YYYYMMDD>/cbs\_accounttype\_range.csv | /DEV/EDW/USER\_DATA/CBS/DTN/<YYYYMMDD>/cbs\_accounttype\_range.csv |
| 2 | CBS\_OFFER\_RANGE\_DTN | cbs\_offer\_range.csv | /EDW/nfsedw105/USER\_DATA/CBS/DTN/<YYYYMMDD>/ cbs\_offer\_range.csv | /SIT/EDW/USER\_DATA/CBS/DTN/<YYYYMMDD>/ cbs\_offer\_range.csv | /DEV/EDW/USER\_DATA/CBS/DTN/<YYYYMMDD>/ cbs\_offer\_range.csv |
| 3 | CBS\_PRODUCT\_RANGE\_DTN | cbs\_product\_range.csv | /EDW/nfsedw105/USER\_DATA/CBS/DTN/<YYYYMMDD>/ cbs\_product\_range.csv | /SIT/EDW/USER\_DATA/CBS/DTN/<YYYYMMDD>/ cbs\_product\_range.csv | /DEV/EDW/USER\_DATA/CBS/DTN/<YYYYMMDD>/ cbs\_product\_range.csv |
| 4 | CBS\_RESULT\_CODE\_DTN | cbs\_result\_code.csv | /EDW/nfsedw105/USER\_DATA/CBS/DTN/<YYYYMMDD>/ cbs\_result\_code.csv | /SIT/EDW/USER\_DATA/CBS/DTN/<YYYYMMDD>/ cbs\_result\_code.csv | /DEV/EDW/USER\_DATA/CBS/DTN/<YYYYMMDD>/ cbs\_result\_code.csv |

| **#** | **Feed Name** | **File Name** | **Source File path - PROD** | **Source File path - SIT** | **Source File path-DEV** |
| --- | --- | --- | --- | --- | --- |
| 1 |  |  |  |  |  |
| 2 |  |  |  |  |  |
| 3 |  |  |  |  |  |
| 4 |  |  |  |  |  |

## 1.4 ODS Table Naming convention

* Transaction Table : DWO\_DEF\_<feed\_name>
* Dimension / Master Table : DIM\_DEF\_<feed\_name>

# Assumptions

* Master data file will be pushed to the AEP landing path on a daily basis even if there is no new data also.
* Files are required to be send to existing platform and new platform concurrently until AEP is stabilize. However Existing file formats remains the same in both the platform.
* File count validation cannot be done for source feed which are not having any control file.  
  No process currently in DTAC.
* This Source data server path details are considered as is with current BI system, any change in setup will have impact on the data ingestion configuration mechanism.

# Interface Details

## 3.1 Details of Source Feed - MASTER FILES – USER\_DATA

|  |  |
| --- | --- |
| Interface Number |  |
| Interface Name | MASTER FILES – USER\_DATA |
| Interface Owner | Refer to Section 1.2 for Feed wise Owner Detail |

### Source Feed Files

The below table captures the list of files which will be integrated for master feeds as part of first phase of ODS go-live.

| **#** | **Feed Name** | **File Name** | **File Format** | **Char set** | **Conv?** | **File Type** | **Freq uency** | **Mechanism of File Transfer** | **No. of files EOD** | **DELI MITER** | **Extract type** | **SCD?** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 1 | CBS\_ACCTTYPE\_RANGE\_DTN | cbs\_accounttype\_range | csv | iso-8859-1 | Y | Master | DAILY | Pull | 1 | , | Full Dump | 1 |
| 2 | CBS\_OFFER\_RANGE\_DTN | cbs\_offer\_range | csv | iso-8859-1 | Y | Master | DAILY | Pull | 1 | , | Full Dump | 1 |
| 3 | CBS\_PRODUCT\_RANGE\_DTN | cbs\_product\_range | csv | utf-8 | N | Master | DAILY | Pull | 1 | , | Full Dump | 1 |
| 4 | CBS\_RESULT\_CODE\_DTN | cbs\_result\_code | csv | us-ascii | N | Master | DAILY | Pull | 1 | , | Full Dump | 1 |

| **#** | **Feed Name** | **File Name** | **File Format** | **Char set** | **Conv?** | **File Type** | **Freq uency** | **Mechanism of File Transfer** | **No. of files EOD** | **DELI MITER** | **Extract type** | **SCD?** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 1 |  |  |  |  |  |  |  |  |  |  |  |  |
| 2 |  |  |  |  |  |  |  |  |  |  |  |  |
| 3 |  |  |  |  |  |  |  |  |  |  |  |  |
| 4 |  |  |  |  |  |  |  |  |  |  |  |  |

#### 3.1.1.1 Interface Characteristic

|  |  |
| --- | --- |
| Source Feed Name | All source feeds |
| Source Feed Description | See Section 3.1.2 on Source feed files |
| File Type  (Transaction / Reference) | Master |
| Collection Mechanism (Push / Pull) | Pull |
| Extraction Criteria | Daily Full Dump/ Incremental [ Check section 3.1.1 on individual feed wise details] |
| Collection Protocol | **SFTP** |
| File frequency /Collection Frequency | Refer to Section 4.1 for details. |
| File Format | “|” delimited File & ‘,’ separated file  File extension is .csv /.txt |
| Control File available? (Yes / No) | Yes [refer to Section 3.1.3 for detail]] |
| Header Record Available? | Yes [refer to Section 3.1.1.4 for detail] |
| Footer Record Available? | No |
| Retention Period at Source/ File Archival policy | Non CDR -7days [ File Based] |

#### 3.1.1.2 Environment Details & Access Details

|  |  |
| --- | --- |
| Production IP Address | TBD |
| Production Port | TBD |
| Production User Name | TBD |
| Production Source Folder | Server : 172.17.4.23  Source File Path :-  /EDW/nfsedw105/USER\_DATA/CBS/DTN/<YYYYMMDD>/ |
| SIT IP Address | TBD |
| SIT Port | TBD |
| SIT User Name | TBD |
| SIT Source Folder | Server : nickel18  /SIT/EDW/USER\_DATA/CBS/DTN/<YYYYMMDD>/ |
| Dev IP Address | TBD |
| Dev Port | TBD |
| Dev User Name | TBD |
| Dev Source Folder | Server : nickel17  /DEV/EDW/USER\_DATA/CBS/DTN/<YYYYMMDD>/ |

#### 3.1.1.3 Information required for first time loading & data transfer

For Transaction feeds the data from ODS will start from date of GO-LIVE of ODS.

Historical data loading and ODS Data retention strategy captured in the below table

| **#** | **Feed Name** | **File Name** | **File Format** | **ODS Historical Load** | **ODS Data Retention** |
| --- | --- | --- | --- | --- | --- |
| 1 | CBS\_ACCTTYPE\_RANGE\_DTN | cbs\_accounttype\_range | Csv | N | Full Table |
| 2 | CBS\_OFFER\_RANGE\_DTN | cbs\_offer\_range | Csv | N | Full Table |
| 3 | CBS\_PRODUCT\_RANGE\_DTN | cbs\_product\_range | csv | N | Full Table |
| 4 | CBS\_RESULT\_CODE\_DTN | cbs\_result\_code | csv | N | Full Table |

| **#** | **Feed Name** | **File Name** | **File Format** | **ODS Historical Load** | **ODS Data Retention** |
| --- | --- | --- | --- | --- | --- |
| 1 |  |  |  |  |  |
| 2 |  |  |  |  |  |
| 3 |  |  |  |  |  |
| 4 |  |  |  |  |  |

#### 3.1.1.4 Header Record Format

Please find below details for feed wise Header record availability

| **#** | **Feed Name** | **File Name** | **File Format** | **Header present?** | **No of Row** |
| --- | --- | --- | --- | --- | --- |
| 1 | CBS\_ACCTTYPE\_RANGE\_DTN | cbs\_accounttype\_range | Csv | Y | 1 |
| 2 | CBS\_OFFER\_RANGE\_DTN | cbs\_offer\_range | Csv | Y | 1 |
| 3 | CBS\_PRODUCT\_RANGE\_DTN | cbs\_product\_range | csv | Y | 1 |
| 4 | CBS\_RESULT\_CODE\_DTN | cbs\_result\_code | csv | Y | 1 |

| **#** | **Feed Name** | **File Name** | **File Format** | **Header present?** | **No of Row** |
| --- | --- | --- | --- | --- | --- |
| 1 |  |  |  |  |  |
| 2 |  |  |  |  |  |
| 3 |  |  |  |  |  |
| 4 |  |  |  |  |  |

#### Footer Record Format

Footer not available for this Feed.

#### File count Check

Refer to Section 3.1.1 on Source feed files for column “No. of files EOD”.

#### Feed to Table Mapping

Following table depicts the proposed Master feeds mapping to TIER 1 (ODS) target table.

| **#** | **Feed Name** | **TIER1 (ODS) Table Name** | **Sizing (byte)** | **Data Volume Per Day** | **Type** |
| --- | --- | --- | --- | --- | --- |
| 1 | CBS\_ACCTTYPE\_RANGE\_DTN | DIM\_DEF\_ACCTTYPE\_RANGE | 2750.4 | 59 | Master |
| 2 | CBS\_OFFER\_RANGE\_DTN | DIM\_DEF\_OFFER\_RANGE | 6572 | 101 | Master |
| 3 | CBS\_PRODUCT\_RANGE\_DTN | DIM\_DEF\_PRODUCT\_RANGE | 764.8 | 17 | Master |
| 4 | CBS\_RESULT\_CODE\_DTN | DIM\_DEF\_RESULT\_CODE | 6743.2 | 92 | Master |

| **#** | **Feed Name** | **TIER1 (ODS) Table Name** | **Sizing (byte)** | **Data Volume Per Day** | **Type** |
| --- | --- | --- | --- | --- | --- |
| 1 |  |  |  |  |  |
| 2 |  |  |  |  |  |
| 3 |  |  |  |  |  |
| 4 |  |  |  |  |  |

### Source Feed wise - Record Format

#### DIM\_DEF\_ACCTTYPE\_RANGE

Master of account type range to classify type of prepaid and postpaid pocket.

|  |  |  |  |
| --- | --- | --- | --- |
| **S.NO** | **Field Name** | **Data Type** | **Sample Values** |
| 1 | ACCOUNTTYPE\_RNG\_ID | INTEGER | 14 |
| 2 | SUB\_ACCOUNT\_TYPE | VARCHAR2(254) | Voice Onnet |
| 3 | ACCOUNTTYPE\_ID\_RNG\_STRT | INTEGER | 5500 |
| 4 | ACCOUNTTYPE\_ID\_RNG\_END | INTEGER | 5549 |
| 5 | ACCOUNTTYPE\_GROUP | VARCHAR2(20) | F |

#### DIM\_DEF\_OFFER\_RANGE

Master of offer range to classify type of prepaid and postpaid offer.

| **S.NO** | **Field Name** | **Data Type** | **Sample Values** |
| --- | --- | --- | --- |
| 1 | OFFER\_RNG\_ID | INTEGER | 10 |
| 2 | OFFER\_TYPE | VARCHAR2(254) | Optional Offer |
| 3 | OFFER\_RGN\_DES | VARCHAR2(254) | SMS without pocket |
| 4 | OFFER\_ID\_RNG\_STRT | INTEGER | 20300000 |
| 5 | OFFER\_ID\_RNG\_END | INTEGER | 20399999 |

#### DIM\_DEF\_PRODUCT\_RANGE

product range master

| **S.NO** | **Field Name** | **Data Type** | **Sample Values** |
| --- | --- | --- | --- |
| 1 | PROD\_RNG\_ID | INTEGER | 12 |
| 2 | PROD\_RNG\_DES | VARCHAR2(255) | Data Receive Service Fee & Â Fax Receive Service Fee Â Â =>(Voice Service Fee) |
| 3 | PROD\_ID\_RNG\_STRT | INTEGER | 2600000 |
| 4 | PROD\_ID\_RNG\_END | INTEGER | 2699999 |

#### DIM\_DEF\_RESULT\_CODE

Result code and description

| **S.NO** | **Field Name** | **Data Type** | **Sample Values** |
| --- | --- | --- | --- |
| 1 | RESULT\_CODE | INTEGER | 102022166 |
| 2 | RESULT\_DESCRIPTION | VARCHAR2(255) | Failure recurring subscription due to RBT auto re deduct after have debt on RBT subscription |

### Control Feed - Record Format

Following table captures the CBS feed wise control file details-

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| # | Feed Name | Feed data file | Feed control file | Control File Format |
| 1 | CBS\_ACCTTYPE\_RANGE\_DTN | cbs\_accounttype\_range | No Control File Available | NA |
| 2 | CBS\_OFFER\_RANGE\_DTN | cbs\_offer\_range | No Control File Available | NA |
| 3 | CBS\_PRODUCT\_RANGE\_DTN | cbs\_product\_range | No Control File Available | NA |
| 4 | CBS\_RESULT\_CODE\_DTN | cbs\_result\_code | No Control File Available | NA |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| # | Feed Name | Feed data file | Feed control file | Control File Format |
| 1 |  |  |  |  |
| 2 |  |  |  |  |
| 3 |  |  |  |  |
| 4 |  |  |  |  |

### Privacy Columns/Business Rules

No data filtration and no exception handling for any columns. Direct one to one mapping and loading to ODS. Surrogate key will be generated only for MSISDN.

Only following type of fields will be encrypted while loading to ODS tables. Attached excel contains CBS feed wise consolidations

* Group 1 - MSISDN / Subscriber Number /Phone no /Fax No
* Group 2 - IMSI\_IMEI
* Group 3 - ADDRESS
* Group 4 - Name
* Group 5 – ID Key
* Group 6 - Email
* Group 7 - Credit Card Number
* Group 8 - Bank Account ID



## 3.2 Details of Source Feed - MASTER FILES – FIX\_DATA

|  |  |
| --- | --- |
| Interface Number |  |
| Interface Name | MASTER FILES – FIX\_DATA |
| Interface Owner | Refer to Section 1.2 for Feed wise Owner Detail |

### Source Feed Files

The below table captures the list of files which will be integrated for master feeds as part of first phase of ODS go-live.

| **#** | **Feed Name** | **File Name** | **File Format** | **Char set** | **Conv?** | **File Type** | **Freq uency** | **Mechanism of File Transfer** | **No. of files EOD** | **DELI MITER** | **Extract type** | **SCD?** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |

| **#** | **Feed Name** | **File Name** | **File Format** | **Char set** | **Conv?** | **File Type** | **Freq uency** | **Mechanism of File Transfer** | **No. of files EOD** | **DELI MITER** | **Extract type** | **SCD?** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |

#### 3.2.1.1 Interface Characteristic

|  |  |
| --- | --- |
| Source Feed Name | All source feeds |
| Source Feed Description | See Section 3.2.2 on Source feed files |
| File Type  (Transaction / Reference) | Master |
| Collection Mechanism (Push / Pull) | Pull |
| Extraction Criteria | Daily Full Dump/ Incremental [ Check section 3.1.1 on individual feed wise details] |
| Collection Protocol | **SFTP** |
| File frequency /Collection Frequency | Refer to Section 4.1 for details. |
| File Format | “|” delimited File & ‘,’ separated file  File extension is .csv /.txt |
| Control File available? (Yes / No) | Yes [refer to Section 3.2.3 for detail]] |
| Header Record Available? | Yes [refer to Section 3.2.1.4 for detail] |
| Footer Record Available? | No |
| Retention Period at Source/ File Archival policy | Non CDR -7days [ File Based] |

#### 3.2.1.2 Environment Details & Access Details

|  |  |
| --- | --- |
| Production IP Address | TBD |
| Production Port | TBD |
| Production User Name | TBD |
| Production Source Folder | Server : 172.17.4.23  Source File Path :-  /EDW/nfsedw105/USER\_DATA/CBS/DTN/<YYYYMMDD>/ |
| SIT IP Address | TBD |
| SIT Port | TBD |
| SIT User Name | TBD |
| SIT Source Folder | Server : nickel18  /SIT/EDW/USER\_DATA/CBS/DTN/<YYYYMMDD>/ |
| Dev IP Address | TBD |
| Dev Port | TBD |
| Dev User Name | TBD |
| Dev Source Folder | Server : nickel17  /DEV/EDW/USER\_DATA/CBS/DTN/<YYYYMMDD>/ |

#### 3.2.1.3 Information required for first time loading & data transfer

For Transaction feeds the data from ODS will start from date of GO-LIVE of ODS.

Historical data loading and ODS Data retention strategy captured in the below table

| **#** | **Feed Name** | **File Name** | **File Format** | **ODS Historical Load** | **ODS Data Retention** |
| --- | --- | --- | --- | --- | --- |
| 1 | CBS\_ACCTTYPE\_RANGE\_DTN | cbs\_accounttype\_range | Csv | N | Full Table |
| 2 | CBS\_OFFER\_RANGE\_DTN | cbs\_offer\_range | Csv | N | Full Table |
| 3 | CBS\_PRODUCT\_RANGE\_DTN | cbs\_product\_range | csv | N | Full Table |
| 4 | CBS\_RESULT\_CODE\_DTN | cbs\_result\_code | csv | N | Full Table |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |

| **#** | **Feed Name** | **File Name** | **File Format** | **ODS Historical Load** | **ODS Data Retention** |
| --- | --- | --- | --- | --- | --- |
| 1 |  |  |  |  |  |
| 2 |  |  |  |  |  |
| 3 |  |  |  |  |  |
| 4 |  |  |  |  |  |

#### 3.2.1.4 Header Record Format

Please find below details for feed wise Header record availability

| **#** | **Feed Name** | **File Name** | **File Format** | **Header present?** | **No of Row** |
| --- | --- | --- | --- | --- | --- |
| 1 | CBS\_ACCTTYPE\_RANGE\_DTN | cbs\_accounttype\_range | Csv | Y | 1 |
| 2 | CBS\_OFFER\_RANGE\_DTN | cbs\_offer\_range | Csv | Y | 1 |
| 3 | CBS\_PRODUCT\_RANGE\_DTN | cbs\_product\_range | csv | Y | 1 |
| 4 | CBS\_RESULT\_CODE\_DTN | cbs\_result\_code | csv | Y | 1 |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |

| **#** | **Feed Name** | **File Name** | **File Format** | **Header present?** | **No of Row** |
| --- | --- | --- | --- | --- | --- |
| 1 |  |  |  |  |  |
| 2 |  |  |  |  |  |
| 3 |  |  |  |  |  |
| 4 |  |  |  |  |  |

#### Footer Record Format

Footer not available for this Feed.

#### File count Check

Refer to Section 3.1.1 on Source feed files for column “No. of files EOD”.

#### Feed to Table Mapping

Following table depicts the proposed Master feeds mapping to TIER 1 (ODS) target table.

| **#** | **Feed Name** | **TIER1 (ODS) Table Name** | **Sizing (byte)** | **Data Volume Per Day** | **Type** |
| --- | --- | --- | --- | --- | --- |
| 1 | CBS\_ACCTTYPE\_RANGE\_DTN | DIM\_DEF\_ACCTTYPE\_RANGE | 2750.4 | 59 | Master |
| 2 | CBS\_OFFER\_RANGE\_DTN | DIM\_DEF\_OFFER\_RANGE | 6572 | 101 | Master |
| 3 | CBS\_PRODUCT\_RANGE\_DTN | DIM\_DEF\_PRODUCT\_RANGE | 764.8 | 17 | Master |
| 4 | CBS\_RESULT\_CODE\_DTN | DIM\_DEF\_RESULT\_CODE | 6743.2 | 92 | Master |
| 1 | CBS\_ACCTTYPE\_RANGE\_DTN | DIM\_DEF\_ACCTTYPE\_RANGE | 2750.4 | 59 | Master |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |

| **#** | **Feed Name** | **TIER1 (ODS) Table Name** | **Sizing (byte)** | **Data Volume Per Day** | **Type** |
| --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |

### Source Feed wise - Record Format

#### DIM\_DEF\_ACCTTYPE\_RANGE

Master of account type range to classify type of prepaid and postpaid pocket.

|  |  |  |  |
| --- | --- | --- | --- |
| **S.NO** | **Field Name** | **Data Type** | **Sample Values** |
| 1 | ACCOUNTTYPE\_RNG\_ID | INTEGER | 14 |
| 2 | SUB\_ACCOUNT\_TYPE | VARCHAR2(254) | Voice Onnet |
| 3 | ACCOUNTTYPE\_ID\_RNG\_STRT | INTEGER | 5500 |
| 4 | ACCOUNTTYPE\_ID\_RNG\_END | INTEGER | 5549 |
| 5 | ACCOUNTTYPE\_GROUP | VARCHAR2(20) | F |

#### DIM\_DEF\_OFFER\_RANGE

Master of offer range to classify type of prepaid and postpaid offer.

| **S.NO** | **Field Name** | **Data Type** | **Sample Values** |
| --- | --- | --- | --- |
| 1 | OFFER\_RNG\_ID | INTEGER | 10 |
| 2 | OFFER\_TYPE | VARCHAR2(254) | Optional Offer |
| 3 | OFFER\_RGN\_DES | VARCHAR2(254) | SMS without pocket |
| 4 | OFFER\_ID\_RNG\_STRT | INTEGER | 20300000 |
| 5 | OFFER\_ID\_RNG\_END | INTEGER | 20399999 |

#### DIM\_DEF\_PRODUCT\_RANGE

product range master

| **S.NO** | **Field Name** | **Data Type** | **Sample Values** |
| --- | --- | --- | --- |
| 1 | PROD\_RNG\_ID | INTEGER | 12 |
| 2 | PROD\_RNG\_DES | VARCHAR2(255) | Data Receive Service Fee & Â Fax Receive Service Fee Â Â =>(Voice Service Fee) |
| 3 | PROD\_ID\_RNG\_STRT | INTEGER | 2600000 |
| 4 | PROD\_ID\_RNG\_END | INTEGER | 2699999 |

#### DIM\_DEF\_RESULT\_CODE

Result code and description

| **S.NO** | **Field Name** | **Data Type** | **Sample Values** |
| --- | --- | --- | --- |
| 1 | RESULT\_CODE | INTEGER | 102022166 |
| 2 | RESULT\_DESCRIPTION | VARCHAR2(255) | Failure recurring subscription due to RBT auto re deduct after have debt on RBT subscription |

### Control Feed - Record Format

Following table captures the CBS feed wise control file details-

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| # | Feed Name | Feed data file | Feed control file | Control File Format |
| 1 | CBS\_ACCTTYPE\_RANGE\_DTN | cbs\_accounttype\_range | No Control File Available | NA |
| 2 | CBS\_OFFER\_RANGE\_DTN | cbs\_offer\_range | No Control File Available | NA |
| 3 | CBS\_PRODUCT\_RANGE\_DTN | cbs\_product\_range | No Control File Available | NA |
| 4 | CBS\_RESULT\_CODE\_DTN | cbs\_result\_code | No Control File Available | NA |

### Privacy Columns/Business Rules

No data filtration and no exception handling for any columns. Direct one to one mapping and loading to ODS. Surrogate key will be generated only for MSISDN.

Only following type of fields will be encrypted while loading to ODS tables. Attached excel contains CBS feed wise consolidations

* Group 1 - MSISDN / Subscriber Number /Phone no /Fax No
* Group 2 - IMSI\_IMEI
* Group 3 - ADDRESS
* Group 4 - Name
* Group 5 – ID Key
* Group 6 - Email
* Group 7 - Credit Card Number
* Group 8 - Bank Account ID



# Interface SLAs & Methods

## 4.1 Availability

Business will place these feeds to EDW landing path in a designated server (172.17.4.23) on daily basis using SFTP protocol. AEP has to pull from EDW Landing path the files.

| **#** | **Feed Name** | **File Name** | **Arrival Time** | **AEP Landing path - Data files** | **AEP Landing path - Control files** |
| --- | --- | --- | --- | --- | --- |
| 1 | CBS\_ACCTTYPE\_RANGE\_DTN | cbs\_accounttype\_range | 3.11AM | /<root>/USER\_DATA/CBS/DTN/<YYYYMMDD>/ | **NA** |
| 2 | CBS\_OFFER\_RANGE\_DTN | cbs\_offer\_range | 2.52AM | /<root>/USER\_DATA/CBS/DTN/<YYYYMMDD>/ | **NA** |
| 3 | CBS\_PRODUCT\_RANGE\_DTN | cbs\_product\_range | 2.40AM | /<root>/USER\_DATA/CBS/DTN/<YYYYMMDD>/ | **NA** |
| 4 | CBS\_RESULT\_CODE\_DTN | cbs\_result\_code.csv | 2.40 AM | /<root>/USER\_DATA/CBS/DTN/<YYYYMMDD>/ | **NA** |

## File Validation

* Following are the high-level validations being performed on master and transaction data feeds -

1. Orphan File check
2. Duplicate file check
3. Zero byte check
4. Header and Trailer
5. Record Count
6. Delimiter check
7. MD5 file Checksum (Applicable for CBS)

### 4.2.1 Duplicate File Check

If the same file is received at AEP landing area for the same day, the process has to mark it duplicate (FileName\_Duplicate.dat) and reject the file.

* Audit table will record name of each file. Compare the file name with previous 30 days filename (configurable), to identify duplicate files.
* Alerts for Duplicate file name to Source system owner over the email once a day (configurable).
* Move duplicate files in reject folder

### 4.2.2 Missing file from Source

An alert will be send to source owner if file is not received in given time frame.

In case of files which are coming in sequence and there is a sequence missing in file, an alert will be sent for missing file sequence.

Email Notification for missing file has been shared by DTAC as per below template



PS: Please see Section – 5 Appendix for email attachment.

### 4.2.3 Missing records or Incomplete File

Reconciliation can be done with source system as able to validate between feed file and control file available with every feed. Reconciliation can also be done based on source file wise count and the data loaded to ODS. Source owners will be notified in case of any anomalies.

### 4.2.4 Alerts

An alert mechanism will be in place to notify source and business owners in case of deviation from processes.

An Email will be triggered to source owners when

* Less number of Files received in expected time.
* When Files will be pushed to rejected folder.

\*\*A Governance process will be set to address any issues in the file transfer process including manual intervention when necessary.

**Action to be taken by source owner after alert - Once Notification alert is sent, Source owner should check and resend the correct data within 4 hours after the receipt of alert.**

**Sample Email Content for Missing file :-**

|  |  |  |
| --- | --- | --- |
| Source System | Source Feed | File Name |
| MASTER | CBS\_ACCTTYPE\_RANGE\_DTN | cbs\_accounttype\_range.csv |

**For respective list of Feed wise file name refer to section 3.1.1**

**Sample Email Content for Reject file :-**

|  |  |  |  |
| --- | --- | --- | --- |
| Source System | Source Feed | File Name | Reject Reason |
| MASTER | CBS\_ACCTTYPE\_RANGE\_DTN | cbs\_accounttype\_range.csv | Name of file is not correct |

**For respective list of Feed wise file name refer to section 3.1.1**

**Sample Email for Less number of files.**

|  |  |  |  |
| --- | --- | --- | --- |
| Source System | Source Feed | Previous Day File Count | Today’s File Count |
| MASTER | CBS\_ACCTTYPE\_RANGE\_DTN | 1 | 0 |

**For respective list of Feed wise file name refer to section 3.1.1**

### 4.2.5 Invalid file format

Files which don’t follow standards of file naming convention will be rejected. For example

* File name should exactly match with the name specified in sec 3.1.1.
* File should always have the .dat extension. Rest all other files with different extension will be rejected.

Refer the extension specified in sec 3.1.1.

### 4.2.6 Erroneous data

In case of any mandatory field missing, datatype mismatch, record separator is not proper, junk VARCHAR2 coming in any field or file name is wrong, then file will not be processed. In such a case notification to be sent for error file and source system need to correct the file and push the correct file into AEP landing path.

## 4.3 Source System Changes

In case of any changes in file format or data format, Source team will inform team well in advance. Any new column addition or changes in existing feed format will be considered as Change Request.

Guidelines for adding new columns are as follows:

**Addition/modification/deletion of columns in file:**

1. When new columns added after registered\* columns in the file. File handling process will ignore the newly added columns (until the columns are registered in the metadata table for that file ). File will not be rejected by the processes; however, these columns will be ignored for processing.
2. If columns are added in between registered\* columns the file will be rejected.
3. If number of columns received are lesser than registered\* columns the file will be rejected.

## 4.4 Unable to Collect or receive files from source

If the files are not received, then after one hour<configurable> auto email alert will be sent to the source system owner.

## 4.5 Unavailability of Data Collection & Integration Layer

In case of unavailability of Data Collection and Integration layer, AEP landing server will keep the files until Data Collection and Integration layer is restored.

However, in case of capacity issues DTAC will notify source owners, to stop pushing files to AEP landing area and manage the files at source end until normalcy is restored.

## 4.6 Data Quality and Timeliness

It would be responsibility of CBS team to have some files/data level checks before files are made available Platform Consumption. Source system team will inform in case there is any issues in CBS provided files.

## 4.7 Exception Handling

Exception handling is subject to scenarios. Different scenarios will be handled differently:

All such files those are required to be reprocessed will be kept in the same source folder from where files are to be collected.

For source files which do not have datetimestamp, duplicate file identification will not be possible. In such case if source push the file twice then the previous file will get overridden and latest overridden file will be processed.  
  
For re-processing of a file which has already been processed and data loaded into tables in such a case we will delete / unload the data from tables by identifying the data based on FILE\_ID and LOAD\_DATE.  
  
In case the file gets pushed with different name then notification via email / or on call will be sent so that the file can be considered for re-processing.

# Appendix



### Email Notification for missing file:

