November 21 Major Release

Data Cleansing Approach

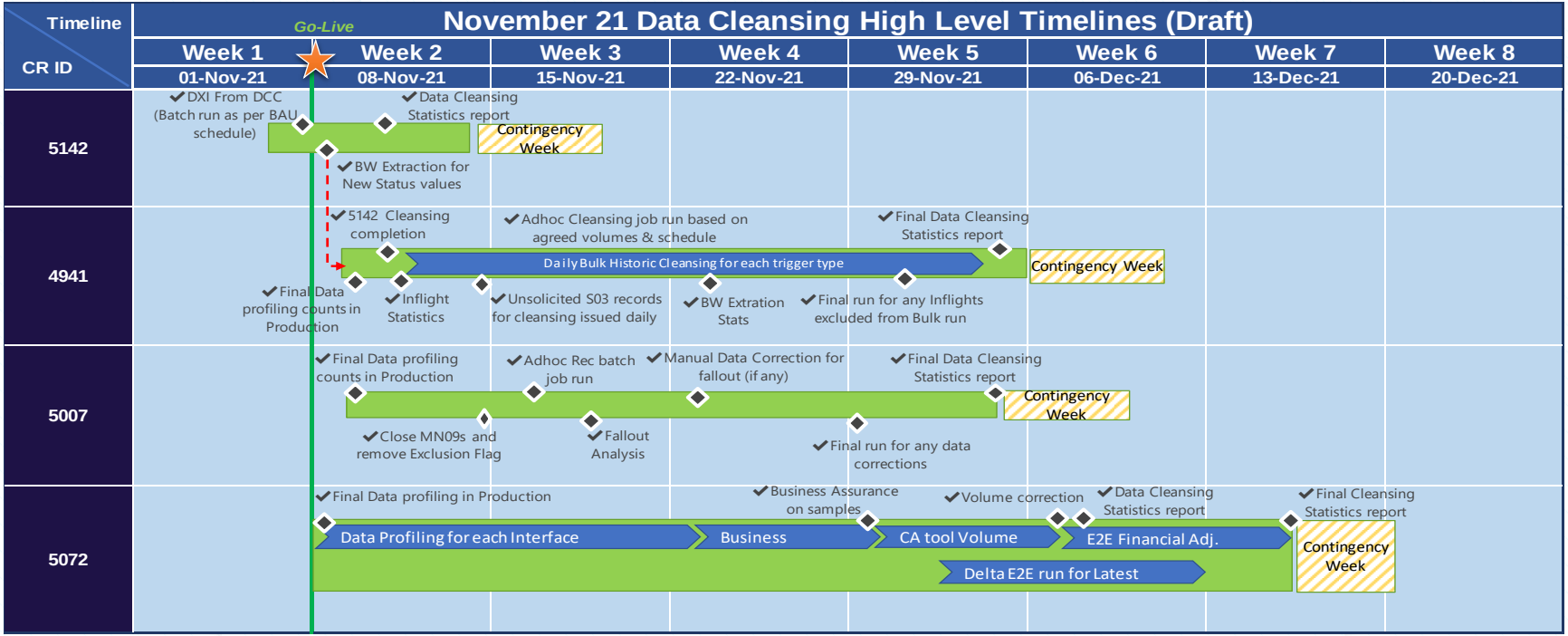
Background

- November 21 Release project is a major UK Link release consisting of 6 changes, approved by ChMC.
- The Detailed Design phase for all 6 changes has been completed, with the final testing phases being completed (Regression and Performance Testing).
- Out of 6 changes in scope of November 21 release, 4 changes require a data cleanse activity where existing UK Link data needs to be profiled and amended as part of project cutover/Post Implementation Support activities.
- These slides will cover the high-level approach to data cleansing for all affected November 2021 changes and aims to provide you with the approach being taken for each.
- This will include, list of planned activities, impacted applications, dependencies and a schedule for each change.
- Below is a complete list of all changes within November 21 and where data cleansing is required.

#	XRN Reference	Description	Data Profiling & Cleansing Required	Business Process Area
1	4941	MOD0692 - Auto updates to meter read frequency	**Y**	**SPA**
2	5007	Enhancement to reconciliation process where prevailing volume is zero	**Y**	**Metering**
3	5072	Application and derivation of TTZ indicator and calculation of volume and energy – all classes	**Y**	**Metering**
4	5142	New allowable values for DCC Service Flag in DXI File from DCC	**Y**	**SPA**
5	5180	Inner tolerance validation for replacement reads and read insertions	**N**	**Metering**
6	4780C	Inclusion of Meter Asset Provider Identity (MAP Id) in the UK Link system	**N**	**SPA, RGMA**

November 21 Data Cleansing - Proposed Timeline

- High Level Timelines for Data cleansing are shown below.
- These timelines will be finalised as part of Performance Test sign off (Targeted for week commencing 18-Oct-21, post completion of Performance Testing Cycle 1) which will consider final volumes for cleansing, job run times and available time slots to run the required jobs in Production



Data Cleansing Approach – XRN5142

SEC proposal (MP077) is seeking to remove the allowable values of “S” - Suspended and “W” – Withdrawn and replace them with values of “N” – Non-Active and “I” – Installed Not Commissioned. Therefore, this Change has been raised in parallel to ensure it is aligned in UK Link. The allowable values for the ‘DCC Service Flag’ data item in the DXI File Format, E46 record, need to be updated to be in line with those proposed by SEC MP077 ‘DCC Service Flagging’. The new DCC Service Flag states should align to what is written in the SEC and allow Users to tell the difference between Devices on a Smart Metering System (SMS) that are Active, installed but not commissioned or decommissioned.

The SEC proposal is seeking to remove the allowable values of “S” - Suspended and “W” – Withdrawn and replace them with values of “N” – Non-Active and “I” – InstalledNotCommissioned.

There is a dependency on the DCC to send the new values to UK Link to initiate the Data Cleansing activity (Approximate volumes are circa 60K).

In addition to the enduring solution, this change will also ensure that all data using the DCC service values is cleansed for the current data set in the system. This activity will be performed collaboratively with inputs from the DCC. This covers below:

- For data cleansing / updates, DXI files will be received from the DCC with the new values of ‘N’ & ‘I’ for identified sites.
    - Effective dates will be utilised as supplied by the DCC.
- The existing UK Link Business As Usual (BAU) job will be utilised to process the data cleansing as per BAU rules. This has the capability to split the total volume over a number of iterations.
- DXR response files will be sent back to the DCC based on the outcome Accepted / Rejected (AC/RJ).
- This will be a one-time activity and run in production during Post Implementation Support based on the BAU schedule. If Performance Testing (PT) results shows that DXI volumes are to be processed over multiple intervals, then timings to run the job for each day will be agreed as part of PT results assurance.
- There is the ability to pause and resume this activity if required to focus on other critical BAU activities.

XRN5142 – Details of CDSP Data Cleansing / Update Activities

#	List of Activities	System	Dependencies	Remarks
1	Final Data profiling before running the Historic DCC cleansing for Sites with DCC_Flag = A, S or W in production	SAP ISU	DCC to provide the proposed volumes during Performance Testing (PT).	
2	Utilise the existing BAU schedule for DXI file processing to update the provided DCC_SERVICE_FLAG values	SAP ISU	Final Volumes will be determined nearer to Go-Live (Current estimate is 60K sites to be cleansed)	Batch run time, no. of parallel work processes & schedule will be determined in PT
3	Hold/Pause the job if required for any BAU critical activity	SAP ISU	Final counts for cleansing	To be determined as result of PT
4	New values to be extracted and passed to BW as part of daily jobs and as a result will be reflected in DES	SAP BW	DCC data cleansing completion	
5	A daily execution summary will be completed by the Project team to validate that all planned cleansing activities have completed successfully (internal use only)	SAP ISU	Completion of data cleansing run	Daily and Final counts
6	Acknowledgment will be sent to Customers to advise that all data cleansing activities for this change have been completed	SAP ISU	Completion of all data cleanse activities.	

Data Cleansing Approach – XRN4941

This change will allow the CDSP to update the Meter Read Frequency (MRF) for Class 4 Supply Meter Point's (SMP) when one of the following trigger points is met and a site does not currently have a monthly MRF:

- Where a Supply Meter Point's AQ value is amended to 293,000 kWh or above on the Supply Point Register.
- Where a Supply Meter Point has AMR Meter installed on the Supply Point Register.
- Where a Supply Meter Point is operational smart meter where a DCC Service Flag is 'Active' on the Supply Point Register.

In the above circumstances, the MRF of the SMP with product class 4, will be amended to monthly by the CDSP.

There is a dependency on the Data Cleansing for XRN5142 to be completed prior to commencing with the activities for XRN4941.

As part of this change, along with developing an enduring solution, a new historic data cleansing program is being developed to cleanse existing data in the UK Link system, for which, the approach is outlined below:

- The Data Cleanse job includes the capability to split the volumes based on each of the three identified trigger points.
- There is the ability to pause and resume this activity if required to focus on other critical BAU activities.
- Exclude any SMP's that have inflight SPA activities where the effective date is within the next 2 days or have a 'CO' status from the bulk data cleanse activity. The MRF for these sites will be corrected post completion of the SPA activity.
- This will be an activity run in production during Post Implementation Support based on the schedule agreed as part of Performance Testing (PT) assurance.
- The unsolicited S03 records will be sent in a staggered manner to minimise impacts to Customers systems. From the volumes seen in the Implementation Dress Rehearsal and the planned timescales to complete the data cleansing for this change, it is forecast that circa 160k S03 records will be produced daily.
- Performance Testing output will determine if the current SCR generation schedule (4 times a day) will support the inclusion of the unsolicited S03 records without impacting current operational response timescales.
- If this is not feasible then additional instances of the SCR generation schedule (potentially 2 further instances) could be added temporarily for the period of data cleanse activities.

XRN4941 – Details of CDSP Data Cleansing Activities

#	List of Activities	System	Dependencies	Remarks
1	Final Data profiling in Production before running the Historic cleansing for Class 4 non-monthly MRF sites for: 1. AQ = >293,000 2. DCC_STAT = A 3. AMR device LIVE	SAP ISU	SQL queries will be developed & run to identify the exact volumes	These will be run nearer to Go-Live in Production (Estimated Volumes for cleansing based on the output from IDR is as follows 4.7M – AQ ~ 5K, DCC ~ 4.5M AMR ~ 160K
2	Schedule the Adhoc Historic Cleansing job in Production for each trigger in PIS period	SAP ISU	Post XRN5142 data cleansing completion & Volumetric for each of the triggers	
3	Identify Inflight SPA activity where Auto MRF update should be completed and exclude from the Data Cleanse batch process. Data Cleanse batch job to be re-run for these MPRNs once the activity is complete	SAP ISU	SQL queries will be developed & run to identify the exact volumes (Run prior to cleansing job)	Confirmation in 'CO'/ Contract change workflow at D-2.
4	SCR files with unsolicited S03 records will be sent to Customers daily	SAP ISU		
5	Monitor the bulk Auto MRF data flow to the SAP BW system	SAP BW	Historic job completion for a day (point 2)	
6	A daily execution summary will be completed by the Project team to validate that all planned cleansing activities have completed successfully (internal use only)	SAP ISU	Completion of historic job run	Daily and Final counts
7	Acknowledgment will be sent to Customers to advise that all data cleansing activities for this change have been completed	SAP ISU	Completion of all data cleanse activities.	

Data Cleansing Approach – XRN5007

This change was raised to address an issue being experienced by the CDSP, where a period has been reconciled to a zero position and then a valid read related to that period is received and re-reconciliation takes place. At this point a divide by zero error is encountered as the prevailing metered volume is zero, as a result an MN09 exception is generated. The scenarios identified that are causing this issue are as follows:

- Re-reconciliation of a zero reconciled period triggered by a site visit or replacement reading.
- A breaking rec (where a previously reconciled period is split as a result of an inserted read) on a non-consuming period.

The data cleansing approach for this change is outlined below:

- Post implementation of the enduring code changes in Production, for all sites with outstanding MN09's in the UK Link system and therefore are currently excluded from the Amendment invoicing process, the flag to exclude these from the Reconciliation process will be removed and they will be re-processed through the new reconciliation factor calculation process (as per the enduring solution).
- Reconciliation will be re-run for all the impacted MPRNs and will utilise the deemed volume rather than the prevailing volume to calculate the Reconciliation Factor (as per the enduring solution).
- In the instance of any remaining MN09 exceptions, these will be categorised into different scenarios which would have led to the site not being resolved, to be looked at manually. These scenarios may include:
  - A Data Issue.
  - Any other exception raised after MN09 resolution.
  - Any other exception already present for the site.
  - Corruption of data caused by a previous attempt to resolve the MN09 exception.

XRN5007 – Details of CDSP Data Cleansing Activities

#	List of Activities	System	Dependencies	Remarks
1	Final Data profiling will be completed prior to running the Auto MN09 exception resolution in Production for the 4 scenarios. The 4 impacted scenarios are as follows: • The zero Prevailing Volume is within a Check to Check reconciliation period • A Consumption Adjustment has already been applied for the same reconciliation period • Incorrect calculation of volume due to multiple Through The Zero's (TTZ)/Round The Clock (RTC) counts • The reconciliation period is for one day	SAP ISU	Reuse Existing report	These will be run nearer to Go-Live in Production (Approx. 2 to 3K)
2	Schedule the Adhoc Rec job in Production post removing the Rec flag from staging tables for sites with existing MN09s	SAP ISU		Batch run time & schedule will be determined in Performance Testing
3	Any MN09 sites with existing data issues or a new exception will be shared with internal business teams for standard resolution steps.	SAP ISU		To be determined as result of Performance Testing
4	Should the volume of sites with an MN09 exception be high (greater than circa 5K) the volume to include in each billing month will be managed by the CDSP and communicated to the impacted Shipper Users	SAP ISU	Business Operations to determine the split based on Pre-prod statistics	
5	A daily execution summary will be completed by the Project team to validate that all planned cleansing activities have completed successfully (internal use only)	SAP ISU	Completion of historic job run	Daily and Final counts
6	Acknowledgment will be sent to Customers to advise that all data cleansing activities for this change have been completed	SAP ISU	Completion of all data cleanse activities.	

Data Cleansing Approach – XRN5072

This change is looking to implement an enhanced formula to asses the Through The Zeros (TTZ) count and ensure the TTZ value is applied within the volume calculation correctly, based on the historic read activity for the MPRN in question.

The Data Cleanse activity will review all volume calculations that involved a TTZ since Nexus go live or Line in the Sand (LIS), whichever is later, which will identify any inaccurate calculations and seek to address them.

However, this change will not amend the TTZ derivation logic for RGMA flows.

In addition to the enduring solution, this change will also cleanse the existing data where the volume has been calculated incorrectly due to the incorrect application of the TTZ indicator for identified scenarios. This covers the following:

- Capability to split the volumes based on each of the different read sources (UMR, UBR, RGMA, and SFN)
- There is the ability to pause and resume this activity if required to focus on other critical BAU activities.
- Identify initial and final volume values for correction via Consumption Adjustment(s) for each of the triggers
- This will be an activity run in production during Post Implementation Support based on the schedule agreed as part of Performance Testing (PT) assurance.
- For clarity, there are no known issues with volumes calculated based on reads received via a fault notification file (SFN). Should any issues be found in this area between now and November 21 go-live the CDSP will seek to correct these manually.

XRN5072 – Details of CDSP Data Cleansing Activities

#	List of Activities	System	Dependencies	Remarks
1	Indicative Data profiling will be completed, for each trigger, within Pre-prod prior to running the TTZ Volume correction program in Production	SAP ISU	New Data cleansing program run completion for each trigger	Pre-prod statistics as a result of PT will indicate the volumes
2	Schedule the Adhoc Data profiling program in Production based on agreed run time and Volumes for each of the trigger type in PIS period	SAP ISU	Indicative volumes will be available in Performance Testing (PT).	
3	Information relating to Consumption Adjustment charges will be shared with Customers in advance of them being processed	SAP ISU	Process to be planned and agreed	Need to understand how Customers will be engaged to evaluate this information
4	Feed the identified Data cleansing sites to the Consumption Adjustment (CA) tool for volume correction	SAP ISU		To be determined as result of PT
5	Post processing of the volume correction in the CA Tool, the revised volume will feed into the following downstream processes. • Consumption Adjustment Invoicing • Rolling AQ correction • Formula AQ correction	SAP ISU		
6	Ensure that all data updated as part of the data cleanse activity flows to SAP BW successfully in order to support downstream reporting	SAP ISU	Completion of data cleansing job run (daily)	Daily and Final counts
7	A daily execution summary will be completed by the Project team to validate that all planned cleansing activities have completed successfully (internal use only)	SAP BW	Completion of data cleansing job run	Daily and Final counts
8	Acknowledgment will be sent to Customers to advise that all data cleansing activities for this change have been completed	SAP ISU	Completion of all data cleanse activities.	