

## POLICY ISSUES PAPER – CONTROL SHEET



Title of Paper	<b><i>Migration Strategy Paper</i></b>		
DA Issue Ref	xxx	Date:	22 September 2016
Issue Owner (Accountable)	<i>James Crump</i>		
Author of Paper (Responsible)	<i>John Wiggins and Jon Spence</i>		
Status of Paper	4 – Draft for EDAG Review 5 – Final Recommendation to DA		
Timing	<i>Completion of the migration strategy is planned to be progressed within the detailed design phase, however, opportunities do exist to identify activities that will add value to the next stage of delivery and decision-making, which have been identified within this paper.</i>		
Dependencies	TBC		

Circulation	Workstream Leaders / Design Team / User Group / EDAG / DA  Huddle / Website  <i>Design Team</i>		
-------------	---	--	--

Issue	<i>A migration of existing industry data from legacy systems will be required to support the deployment of a new CRS solution.</i>		
Impacts Domestic?	Yes	Impacts Non-Dom?	Yes
Policy Objective (and reference to ToM v2)	<i>Section 12.35 of Version 2 of the ToM states that "The Transition and Implementation Scheme will incorporate</i>		

	<ul style="list-style-type: none"> <li>• <i>Arrangements for the migration and cleansing of any data, including how data cleansing would be funded”.</i></li> </ul> <p><i>This paper provides a high level strategy for migrating data to the new CRS. Data cleansing arrangements are addressed in a separate product.</i></p>
Previous Positions on this/related Issues	<i>Work undertaken by the Business Process Design workstream, such as the Data Model and the Solution Architecture Model.</i>
Summary of Recommendations	<i>TBC</i>

<b>Internal and External Engagement</b>	
Business Process Design	<i>Jenny Boothe, Phil Bryan</i>
Regulatory Design	<i>Jon Dixon</i>
Delivery Strategy	<i>James Crump</i>
Commercial Strategy	<i>Natasha Sheel</i>
DIAT	<i>Andrew Wallace</i>
Legal	
Other Ofgem Teams	
<b>Meetings at which this paper has been discussed</b>	
Workstream Leaders	
User Group	<i>29 September 2016</i>
EDAG	<i>13 October 2016</i>
Other External	<i>Ditto</i>
Ofgem Design Authority	<i>Ditto</i>

## POLICY ISSUES PAPER – CRS Data Migration

### Introduction

1. This paper sets out our high-level Data Migration Strategy that will support the delivery of the Central Registration Service (CRS) and its component parts, the Switching System and the Market Intelligence Service, within the new switching arrangements. The Strategy forms part of the Blueprint Phase of the Programme. A more detailed approach to migrating data will be developed in the Detailed Level Specification phase of the programme, at which stage the strategy will be refined to reflect the selection of the target solution architecture and feedback from the Request for Information (RFI). Dependent on the solution architecture option chosen, data will be required to be migrated from source legacy systems to the new target CRS architecture.<sup>1</sup>
2. The scope of the Data Migration Strategy shall include;
  - a. The identification of the source systems and the individual data elements, identified within the data model that will be within scope of the migration;
  - b. Identification of how the source data elements map to the data elements required in the target CRS;
  - c. Identification of possible migration options;
  - d. Identification of interactions between data migration and products contained within other workstreams.
3. Following The Open Group Architecture Framework (TOGAF) for enterprise architecture will create an approach of how to move from the existing baseline to the target architecture through the creation of a detailed migration and implementation plan.

### Essential Background

#### Target Operating Model

4. Section 12.35 of the Target Operating Model v2 makes reference to the need for data migration within the Switching Programme:

“The [transition and implementation scheme] TIS will incorporate:

- Arrangements for the migration and any cleansing of data, including how data cleansing would be funded.

---

<sup>1</sup> Data elements required for faster and more reliable switching are currently held in the different legacy registration and enquiry systems operated by Xoserve, MRA, Gas Transporters, Independent Gas Transporters (iGTs), Distribution Network Operators (DNOs) and Independent Distribution Network Operators (IDNOs). Appendix 1 provides a view of the current, high level, system architecture model.

Arrangements for the transfer to new governance arrangements (ie from programme governance to the enduring arrangements under licences and industry codes).<sup>2</sup>

## Dependencies within the Switching programme

5. **Data Model:** An architectural data model for the new switching arrangements has been developed by the Business Process Design (BPD) workstream. It will be the reference, for indicative purposes, of data elements within the scope of the RFI which shall be presented to industry. A logical data model will be developed within the Detailed Level Specification (DLS) phase of the switching programme. It is proposed that the model be developed in accordance with ISO/IEC 11179, the international standard for Metadata Registries.
6. The logical data model will be determined by the CRS functional requirements and aligned to the solution architecture, each will be required as pre-requisites for the data migration products developed within the DLS phase. This product will be the definitive controlling product for defining the data elements that shall be mastered by the switching system or legacy systems and identifying the data steward (responsible for maintaining data quality) for each data element.
7. Within the Delivery Strategy workstream, the Transition Strategy and the Data Cleanse Strategy share a close dependency to data migration. In the context of decision-making around the most effective and lowest risk transition option, data migration options also need to be considered which will enable the optimal attainment of the delivery goals.
8. **Solution Architecture:** The solution architecture options will determine the migration requirements of the programme. We proposed that the RFI will seek feedback on the following options;
  - a. **Do nothing:** No changes will be made to the existing industry systems and processes.
  - b. **Do minimum:** This option will not feature the creation of a central 'Switching System' or 'Market Intelligence Service'. Existing legacy data systems will operate in a continuous manner for both gas and electricity, and as such no data migration will be required.
  - c. **Switching database with middleware:** A centrally-maintained 'Switching System' will be created, containing data used to effect a switch between suppliers. Creation of this system will require a migration of data from existing legacy electricity and gas systems. The extent and complexity of the migration would depend on the number of data items held within the 'Switching System'. Other data items would remain within legacy systems and will not be migrated.<sup>3</sup>

---

<sup>2</sup> "Target Operating Model and Delivery Approach, v2", pp55-56

<sup>3</sup> Appendix 3 provides a high level overview of the data model required to support this option.

- d. **Switching database and MIS database with middleware:** As above, a central 'Switching System' will be created and held centrally, requiring a data migration, the extent of which will be determined by the scope of data held within this database. In addition, a 'Market Intelligence Service' (MIS) will be created, which will allow interrogation of other data sources. Dependent on the solution, existing enquiry systems would likely be decommissioned post migration.
9. Existing data items can be divided into three groups;
    - a. Data items that shall be migrated from their existing legacy source systems and then mastered<sup>4</sup> within the CRS, with the assumption that all of these data items shall be mastered within the 'Switching System' but could also be made available to the MIS and still located within the legacy systems for information and data integrity purposes.
    - b. Data items that shall be migrated and are required to enable the switching operation, but not mastered within the CRS, remaining mastered within the existing legacy systems. These data items will be required by the 'Switching System', to enable the identification of the specific Registrable Measurement Point (RMP) which is to be traded between Energy Suppliers, they could also be held within the MIS for informational purposes.
    - c. Data items not required within the 'Switching System' to enable the trading of the RMP but are required within the MIS to enable a more effective customer validation and quotation experience prior to the switch. Existing legacy systems shall continue to be the master of this data.
  10. The work undertaken, to date, within the Blueprint Phase has not been performed to a level of detail that identifies how all of the existing data items utilised by the industry will map to the data element concepts catalogued within the data model. This will need to be completed in the DLS phase, at the point that a logical data model is created.
  11. The system architecture, data model and process model products delivered by the BPD workstream collectively describe the data elements that shall be required to enable both the 'Switching System' and the MIS and the design of the new database. Data analysis will define the scope of the data migration in terms of the existing data items that will be required to support each solution architecture option.
  12. **Transition Strategy:** The Transition Strategy will determine how the new Switching arrangements are brought to market. This will determine the order and timing with which the components of the solution architecture are

---

<sup>4</sup> In the context of the paper the term 'master' is used to identify the application that has control over a specific data element. It is the only application that is able to create, update or delete that data element.

'switched on'. This will have implications for how data is migrated and the timing of any migration.

13. **Data Cleanse:** A strategy is under development to cleanse data which is likely to feature in the CRS. This data cleanse strategy will focus, primarily, on address data.

## Analysis

### Migration Approaches

14. A migration of existing data will be an exercise conducted as part of the Switching Programme. The main actors will be the Registration Agent (a role performed by the DCC) and Network Operators. Other actors, such as Suppliers and Supplier Agents, will also be involved to a greater or lesser extent dependent on the solution.
15. In the gas market, the identified data elements will, after Project Nexus has been implemented, reside within a single system operated by Xoserve, on behalf of Licenced Gas Transporters, including Independent Gas Transporters.
16. In the electricity market, the system landscape is more complex, as each Licenced Electricity Distribution Network Operator provisions their own systems, although these systems are architecturally similar. The utilisation of a common data structure, the MRA Data Transfer Catalogue, also enables homogeneity of metadata.
17. Although different parties may take different approaches to operational and data quality, data management processes will need to be validated within the migration planning.
18. Several approaches exist for data migration, the correct approach will need to be determined within the migration design phase and in conjunction with the overall architectural and implementation solutions.
19. The data migration activity should be phased in the following way;
  - a. Design
  - b. Extraction
  - c. Cleansing
  - d. Load
  - e. Verification
20. Because of the scale and complexity of the migration, a design phase is particularly important. Architectural design is likely to be necessary to describe how systems will be devised to assist the migration from the current architecture to the target architecture, which may also include transitional

architectures, which in part, could be driven by data migration requirements identified within the design phase.

## Migration Principles

21. Any data migration must occur in a way that allows continuity of service across the old and new switching systems. Based on this, we have identified the following principles, which must be reflected in the conduct of the migration:
  - a) Data migration must be completed in a manner which minimises the impact on cost, consumers and competition, and reflects the overall transition strategy for the switching arrangements;
  - b) Data quality must be maintained from legacy to target systems and within any transitional systems if required.
  - c) The migration should be conducted in a way that minimises risk to existing systems or processes. Throughout the migration process, the key switching details must remain synchronised across the new CRS, the legacy MPAS and UKLink registration services, the DCC access control database and Data Aggregator systems;
  - d) Existing governance arrangements in place during the Design, Build and Test phase, combined with the System Integration function, should be configured to ensure that risks before and following migration are managed; and
  - e) A detailed migration plan should be drawn up once there is sufficient certainty around the solution architecture and data model.

## Data Stewardship and Mastering of Data

22. The data modelling product has identified indicative data stewards<sup>5</sup> for the individual data elements, and whether stewardship will change once the new system arrangements are operative (i.e. after migration). If stewardship changes pre- and post- migration, the party with responsibility for performing migration must be clearly defined. Changes to the stewardship and mastering of data will require the imposition of formal data governance framework under new arrangements, which will need to be considered within the new regulatory functions created to support the operation of the CRS.
23. Data stewards will be formally identified at the DLS phase of the project. At this point these parties will be required to participate in the development and delivery of the migration plan.
24. Data stewardship, and therefore responsibility for migration, will depend on the application or service that a particular data item is mastered within. The ability to change (create, update, or delete) data will only be possible within the master system. Where data will be mastered within the Switching System

---

<sup>5</sup> The actor responsible for managing the quality of data. This includes the definition and description of structured-data elements, the specification of data controls and the oversight of the stored data values.

(either within the CRS, or depending upon design, the Market Intelligence Service (MIS)), the registration agent responsible for operating the Switching System may become the data steward. Other data elements contained within the CRS which continue to be mastered within the legacy systems under the new arrangements will be stewarded by the owners of those systems. An enduring integration solution will be required to ensure ongoing alignment of data.

25. Any data elements held within the CRS are likely to be crucial to ensuring customers are able to switch quickly and reliably, and therefore to public confidence in the new switching arrangements and the operation of the UK retail energy market. For this reason, it is essential that migration activity is subjected to an adequate governance regime, as the impact of data migration issues will be very high.

### **Data Conversion**

26. A majority of data elements currently identified within the data model product are based on existing data items, harmonisation has been attempted between gas and electricity datasets where possible. Some data elements may require conversion when they are migrated to any new switching arrangements, this requirement will be investigated further within the DLS phase.
27. Other data elements, such as Premises Address, may not be based on existing industry data items. For example, a service such as Ordnance Survey's AddressBase could be utilised. The migration and conversion of external data will also need to be investigated within the migration design phase.
28. The options for data conversion will be determined by the system architecture and the data model that will only be confirmed post completion of the RFI, to be held in 2017. It will not be possible to offer detailed and specific options for the delivery of the migration until the specific solution is known without a significant amount of uncertainty and duplication of work. The utilisation of middleware within a solution could provide data conversion capabilities and will need to be investigated within the design phase.
29. However, a data prototyping exercise could be conducted on existing industry data to verify the outcome of combining or converting the source data into new data sets to support the CRS. This exercise could be performed utilising data extracted from industry systems.

### **Governance and Assurance**

30. The migration process will require oversight to ensure that market participants are undertaking their roles effectively, that the migration is completed without exposing the Programme as a whole to undue risk.
31. The migration will take place in the Design, Build and Test phase of the Switching Programme. During the development of the plan, it must be decided whether the existing governance and assurance processes for the DBT phase will be adequate to ensure that the migration is completed with minimal risk. The plan should set out responsibilities and overall control for



the migration process, and whether there is a role for a system integration, or other parties operating a system integration function. These requirements should be considered within the decision making associated to the potential creation of a new retail code.

## Recommendation

32. Whether a data migration will be necessary, and the extent of that data migration, is dependent on the design of the final Solution Architecture adopted. A data migration will only be necessary if the Solution Architecture includes an MIS and/or Switching System. If migration is necessary, the migration itself will take place within the Design, Build and Test phase of the Switching programme.
33. The need for a data migration will become apparent when a decision is taken on the chosen solution architecture, and (if necessary) on the data elements that will form part of the Switching System and/or MIS. When this certainty is achieved, a detailed plan or plans for the migration should be created.
34. The decision on solution architecture is not planned to be made until December 2017. It would be possible to begin detailed work in regards to data migration sooner than this date if certain assumptions are made. For example, an assumption that architecture option 3 would be selected could allow the activities identified below to begin sooner. With the appreciation that the scope would need to be expanded if option 4 is eventually selected or the work undertaken not utilised if option 1 or 2 selected. Although this creates a risk that the investment could be written off it would also significantly de-risk the overall programme delivery timescales.
35. The products delivered within the DLS phase, or sooner if solution architecture assumptions are made, should encompass the following:
  - a. **Data Catalogue:** –defining all data elements, their metadata and identify the data steward and mastering application for each.
  - b. **Historical Data Assimilation:** An assessment of what historical data (for example historical supplier registrations, agent appointments and settlement data, used for reconciliation in electricity) will be migrated into the CRS and which will be retained in the legacy registration systems (such as MPRS and UKlink).
  - c. **High-Level Source System to Target System Mapping Specification:** An initial view of the data elements and their relationships that will be linked during the migration, including an identification of which elements will require conversion.
  - d. **High-level Volumetric and Scoping Report:** Analysis that will confirm the volume and scope of data to be migrated to the CRS.
  - e. **Risk and Issue Management Control:** An assessment of risks and issues associated to individual data items within the migration.

- f. **Data Quality Management Process and Impact:** Determining specification for data quality for each data element required for the CRS and the MIS (if required). Also determines how issues occurring ahead of, during and after migration will be identified, managed and resolved.
- g. **Change of Data Application Mastering:** Working in conjunction with the solution architecture product, this product will identify and address risks in order to ensure that legacy systems will continue to operate post migration and that the enduring interfaces are operationalised to ensure ongoing data quality.
- h. **Data Prototyping:** If used, analysis based on existing industry data sets, testing scenarios to validate the CRS data model, quality of existing data, data conversion options and optimal migration approach.
- i. **Migration Governance and Assurance:** An assessment of which parties will be responsible for governance and assurance of the data migration.
- j. **Migration Approach:** An assessment of the merits, costs and benefits of identified options to support the migration of data within the wider solution and transition design.

36. In addition to this development of the strategy, a detailed operation plan will be developed, explaining technical aspects of data migration, including (but not necessarily limited to) the following aspects.

<b>Migration Aspect</b>	<b>Description</b>
Extract Design	How data will be extracted from legacy registration systems, exchanged and verified
Migration Design	How data will transformed into the target structure
Conversion Plan	Listing of data elements within the CRS (and/or the MIS) that will require conversion as an activity within the migration, and a specification of how these will be converted.
Mapping Rules	Detailed level details of migration mapping
Testing	Testing any refresh/delta update functionality
Quality control	Assessment that data migration is complete and of a sufficient quality to allow cut-over from old to new switching arrangements (and responsibility for signoff)
Data Security	Controls in place to ensure data security during migration

Recovery	Recovery options in the event of catastrophic failure for each stage of the migration
Acceptance Criteria	Conditions required for migration to be considered complete (and responsibility for signoff)

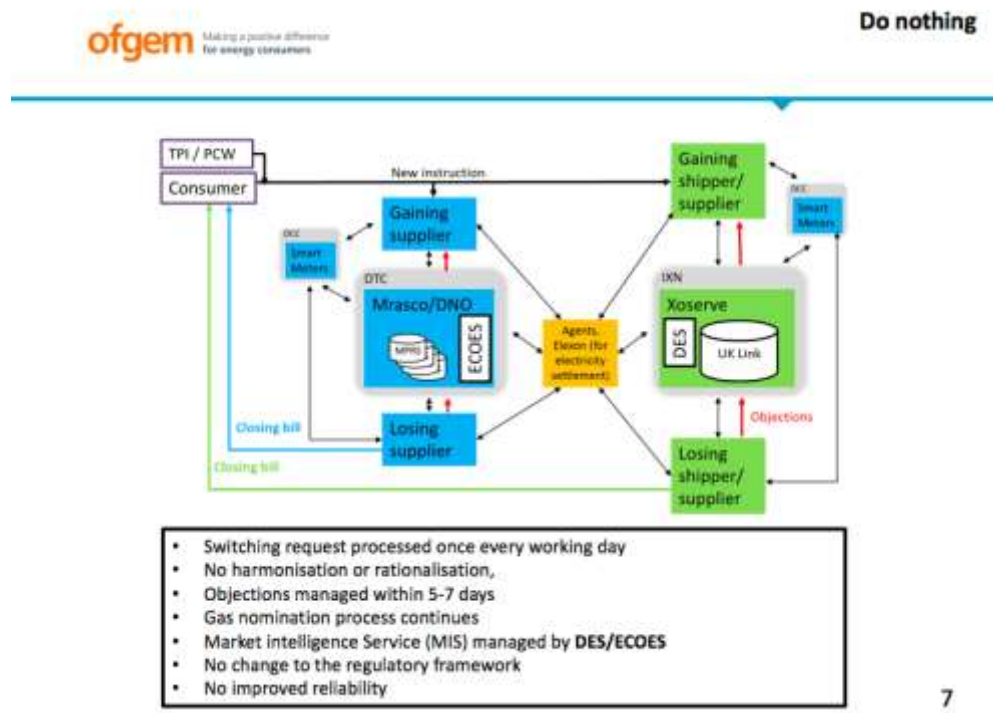
### Questions for the User Group

37. We ask respondents to the RFI to consider the following questions:

- *Do you agree that the content of this paper is of sufficient scope and is factually correct?*
- *Do you agree with our contention that a detailed strategy and plan for data migration should be developed during DLS?*
- *And that the DLS should include a data migration design phase, which if assumptions are made in regards to the solution architecture, could be started sooner than the current planned start date of December 2017?*
- *Are there any other elements, other than those identified, which should be considered within the detailed strategy and plan to be created during the DLS phase of the project?*

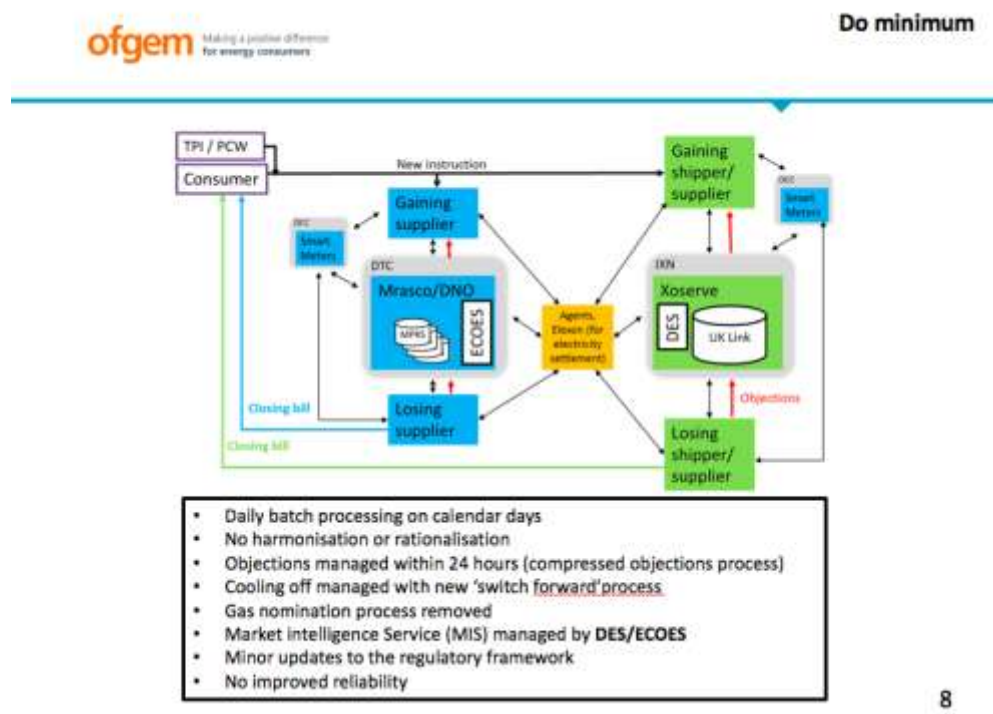
## Appendices

### 1. Legacy Solution Architecture / Do Nothing Architecture (Option 1)



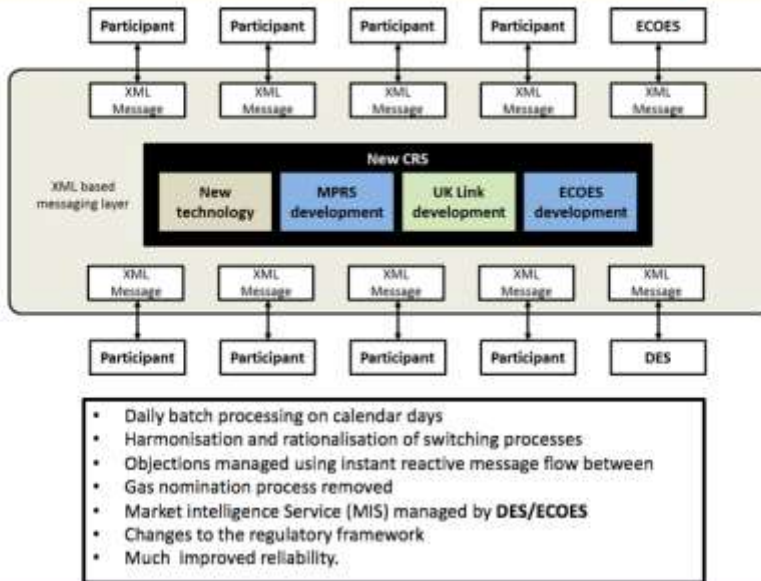
### 2. Solution Architecture Options

#### Option 2 – Do Minimum



#### Option 3 – Switching data base with middleware

### Switching database with middleware (XML communications) layer



### Option 4 – Switching data base and market intelligence service with middleware

### Switching database and MIS database with middleware (XML communications) layer

