OLDHAM BOROUGH COUNCIL AND BRIMSTONE
Providing Reporting Capabilities Across Disparate Data Sources 2016/2017

Prepared by: Steve Jones,
Brimstone Advantage Ltd.
www.brimstone.it
28 June 2017

OVERVIEW

Background

In May 2016, Oldham Council published a tender for the provision of database consultancy services and support to tackle two issues. The first of these was to provide technical assistance to an existing data analysis team and the second task was to research, demonstrate and ultimately identify a data discovery tool for the purposes of enhancing local facing services across Oldham in a variety of data contexts from crime and health data to school census and early help assessments.

Brimstone Advantage Limited ("Brimstone IT") has over 20 years experience in database development and design, coupled with strong Business Intelligence knowledge. Although well versed in Microsoft RDBMS technologies, Brimstone IT also has experience of alternative database products such as PostgreSQL and MySQL. Infrastructure skills are also available, with complete server builds including the Windows Server product family and the Linux family (CentOS, FreeBSD) forming a large part of the Brimstone IT skill set. For these reasons, Brimstone IT was best positioned to bid for this project.

In May 2016, Brimstone successfully bid for and won the tender and has been working with Oldham since this point on a variety of data tasks and developments. These are outlined in the "Technologies and Skills" section of this document to inform the reader as to the types of roles and responsibilities Brimstone IT undertook as part of this tender. It is hoped that these experiences inform decisions relating to future partnership opportunities.

The Challenge

As part of the Council's Troubled Families program, various disparate data sources were being imported into a cloud based third party website (the "Matching Engine") for the purposes of matching data into a single view of a person and/or family (based on standard fields defining a person or family). Whilst this Matching Engine allows for matching and other related features, it does not perform the role of a reporting platform or BI self-discovery tool.

As stated, these data sources originate from a multitude of different internal and external agencies, from crime and social care data to health, schools and benefits. They range (in format) from ODBC databases to Excel files, text files and CSV file formats.

The challenge set by the tender was to implement a reporting solution, which could provide a holistic, service driven view of the data (based on the matched data provided by the Matching Engine) for a given person or family. This solution would need to provide the means to search for a person or family and show the services they had engaged with (data profiling) or the same from a top level reporting aggregation such as a location based search, a KPI value or a defined cohort of people (data discovery). This reporting solution would have to provide the means to explore the data in a non-linear fashion with not only drill-down features but also drill

across. Therefore, a solid understanding of all BI tools, used both in the Council and externally in the market place, would need to be attained and discussed.

Constraints

As stated, the Matching Engine provides the ability to load and match data to current "Golden" people or family records. This means the Matching Engine can present a view of a person (or family) and show all related data about that person (e.g. show all police calls that have originated from that person's address and also any early help assessments for that individual, and so on).

Whilst a powerful matching tool in it's own right, the Matching Engine does not allow for reporting from the data in a flexible, meaningful fashion. The underlying database (for commercial reasons) does not allow for optimisations (such as index application) or pre-defined aggregations to be developed. Also, given the Matching Engine is cloud based, this introduces a critical dependency on the Council's Internet connectivity.

Also, during the lifetime of the project, two members of staff with key involvement in the Matching Engine system and the local Hub database maintenance left the organisation at separate times. This resulted in a requirement for Brimstone IT to manage the transfer of skills and knowledge to the BIS team at a later date. Key to this process was the attendance of handover meetings and creating documentation relating to the various roles and responsibilities of the departing staff members.

THE SOLUTION

Concept

To mitigate the constraints imposed by the Matching Engine, the first role of Brimstone was to design and build a LAN based database (the "Hub") within Oldham Council's network to allow for reporting, optimisations and aggregations to be performed. This concept was introduced to mitigate the risks associated with the cloud based system around the main points noted in the previous section. The Hub database employs the Microsoft SQL Server 2014 platform.

The concept design was therefore that the relational Hub database would be periodically loaded (as the raw data is made available to the Council) with the raw source data (as discussed, from various agencies including the police and from housing and health) and then, combined with the Matching Engine tool, this data would be "knitted" together by importing the local identifiers (present in the raw data files) and their corresponding global identifiers into the Hub database. This "knitting" together of the data therefore presents us with a holistic view of a person or family, giving the possibility of exploring the various Oldham local services a person or family has engaged with.

Loading of the raw data into the Hub would be developed using Microsoft SQL Server Integration Services ("SSIS") as an ETL tool (since the Council already is licensed to use this product), using Microsoft Visual Studio 2012 as the IDE. A SSIS package would exist for each data source and the package would 1) load data into the Hub database and 2) output a flat file which would be used to import into the Matching Engine for the purposes of matching data against people and families. The final component of the design is to extract the matching data from the Matching Engine and import into the Hub to complete the connectivity of the raw data to the single unified person view.

This therefore allows the opportunity to connect BI tools to the Hub database such as Microsoft Power BI, Tableau or any other BI tool which is fit for purpose. Internet connectivity dependancy is removed, and also database enhancements can be made to optimise the data tables for reporting purposes. A future reporting solution therefore would depend on the Hub database for reporting purposes.
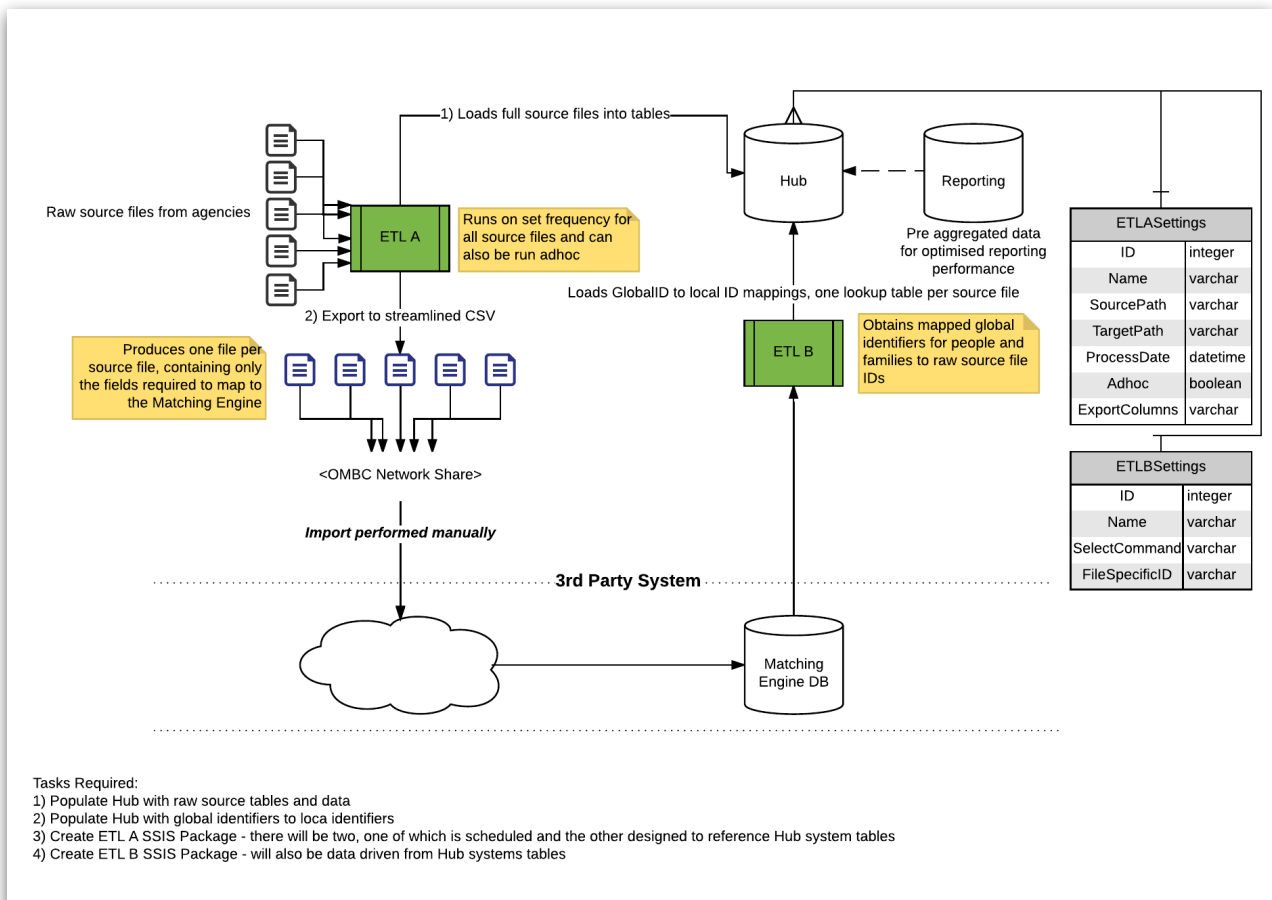
Technologies and Skills

Both the Matching Engine and the Hub database are implemented using Microsoft SQL Server 2014 edition. Access to the hosted Matching Engine database is provided via the third party using a firewalled ODBC connection. In order to provide synchronisation between these two databases, the following skills and technologies were employed:

- Detailed knowledge of the Microsoft Business Intelligence stack
- T-SQL scripts for data manipulation and import
- Microsoft SQL Server Integration Services for creating data driven packages to provide the ETL toolkit
- In-depth knowledge of networks and ODBC connectivity

- Microsoft Visual Studio 2012 with SQL Product components installed
- In-depth knowledge of Microsoft SQL Server Management Studio
- Development of stored procedures and views for reporting purposes in the Hub
- Experience and database knowledge pertaining to the raw data files and their behaviour
- Understanding of the Matching Engine database and it's structure

The following diagram explains the concept of synchronising the Hub database with that of the Matching Engine database and was one of the earlier designs:

Tasks Required:
1) Populate Hub with raw source tables and data
2) Populate Hub with global identifiers to loca identifiers
3) Create ETL A SSIS Package - there will be two, one of which is scheduled and the other designed to reference Hub system tables
4) Create ETL B SSIS Package - will also be data driven from Hub systems tables

TODAY'S LANDSCAPE

Deliverables

As a result of the work performed by Brimstone IT, the following deliverables are now in place:

- The Hub SQL Server database, implemented on the Council's LAN, containing the tables and structure necessary to support business as usual reporting for the organisation
- Microsoft SQL Server Integration Services packages for each data source requiring load and transformation as part of the Hub maintenance required for regular reporting
- Knowledge and experience of the Matching Engine system across the Business Intelligence key personnel
- Detailed documentation and handover materials relating to all aspects in maintaining the Hub, including business as usual processes and also change management
- Training and experience delivered to personnel in the Business Intelligence team, focusing on database development, using T-SQL and Microsoft SQL Server Integration Services
- Wider understanding of the Council's development, change and assessment processes through their existing IT supplier arrangement
- A plan to drive the future development road map relating to the Matching Engine
- Understanding of Business Intelligence tools and their comparative fit with the Council's requirements for a reporting solution

In relation to the identification of a reporting solution to provide data discovery functionality for the Council, a separate project is in the process of initiation at the time of writing.

Exit Strategy

As part of Brimstone IT's exit strategy, a comprehensive training handover plan was designed for the personnel assuming the responsibilities of the Hub and it's maintenance, alongside knowledge transfer relating to the Matching Engine.

This took the form of skills training to complement the BAU tasks, including T-SQL and SSIS skills, since those staff required their skills bolstering in these areas. This skills training was delivered in the form of workshop style led sessions. Full, detailed documentation was provided to ensure business as usual reporting tasks would continue and change to those tasks could be implemented using practical, step by step guides.

Further Information

For more information relating to this case study and how Brimstone IT can help your data connectivity and discovery project, please contact Steve Jones at Brimstone Advantage Limited on steve@brimstone.it or via 07905 344970.

Nicola White, Data Integration Manager within the Business Intelligence team managed Brimstone IT during the project. Nicola can be contacted for information on 0161 770 1696 or via nicola.white@oldham.gov.uk.