# A Cloud Based Data Integration Framework

Nan Jiang[1], Lai Xu[1], Paul de Vrieze[1], Mian-Guan Lim[2], Oscar Jarabo[1]

[1] Software Systems Research Centre,
School of Design, Engineering and Computing,
Bournemouth University, United Kingdom
{njiang, lxu, pdvrieze}@bournemouth.ac.uk
ojarabo@gmail.com
[2] Future Computing Group, School of Computing
University of Kent, United Kingdom
m.g.lim@kent.ac.uk

**Abstract.** Virtual enterprise (VE) relies on resource sharing and collaboration across geographically dispersed and dynamically allied businesses in order to better respond to market opportunities. It is generally considered that effective data integration and management is crucial to realise the value of VE. This paper describes a cloud-based data integration framework that can be used for supporting VE to discover, explore and respond more emerging business opportunities that require instant and easy resource access and flexible on-demand development in a customer-centric approach. Motivated by a case study discussing power incident management in the Spanish Electricity System, an effective on-demand application is also implemented to demonstrate how to use this framework to solve real world problems.

**Keywords:** Cloud computing, Situational application, Mashup, Virtual enterprise, Data as a Service.

## 1 Introduction

Virtual enterprise (VE) is a temporarily and dynamically formed alliance of businesses where these organisations collaborate and share their skills, core-competency and resources in order to better respond to market opportunities [1]. It is commonly understood that the success of VE relies on the flexibility and agility of resource sharing across its member organisations in the virtual network (VN) [2]. Therefore, the key to actualise this lies in the effective data integration and management of each autonomous organisation in the network. VE is intended to meet a market opportunity that cannot normally be answered by individual organisations. However, the current financial and economic situation creates a highly competitive market condition where market opportunities are often unpredictable, short-lived and fast-changing in a wide social context which calls for the development and fast adoption of new information and communication technology (ICT) [8]. Thus, new challenges of maintaining sustained competencies have been brought to enterprises and VE [7]. In this paper, we propose a cloud based data integration framework which allows VE to better respond these market opportunities and customer needs that require instant and easy access to resources with fast and flexible development in a customer-centric approach. Motivated by a case study focusing on a real world problem, the implementation of this framework is also discussed where an effective on-demand application is actualised to tackle real-time power incident management in the Spanish Electricity System.

First, a motivational case is described in the Section 2. Next, the requirements of implementation are analysed in Section 3. A proposed cloud-base framework is then described in Section 4 followed by the presentation of application in Section 5. Last, the conclusion and the future work are drawn.

## 2 A Motivational Case Study

A national electricity system is formed with a high-voltage electric power transmission network and grid connecting power stations and substations to transport electricity from where it is generated to where it is needed in the country. In an industrial perspective, there are three key stakeholder areas in the system: generators, distributors and suppliers [3]. When an incident occurs in the system, effective communications from these stakeholders to clients become crucial. However, this is not normally well managed for various reasons which results in negative impacts on all stakeholders in the system and the community.

### 2.1 The Spanish Electricity System and Key Stakeholders

As shown in Fig. 1, Act 9(1) in Law 17/2007 defines six key stakeholder areas in the nation's electricity system: generators, distributors, system operator, market operator, suppliers and end users [4]. In practice, a utility company commonly plays multiple roles in the system. For example, Endesa (E), Iberdrola (I) and Gas Natural Fenosa (GNF) are the three major energy companies in Spain. They are not only the main suppliers but also the principle generators as well as distributors. Red Electrica De España (REE) is the system operator and carrier for operating the nation's power transmission system and electricity grid and Compañía Operadora del Mercado Español de Electricidad (OMEL) is the market operator dealing with electricity wholesales. End users include industrial users and domestic users where the former is often connected to the high voltage network directly.

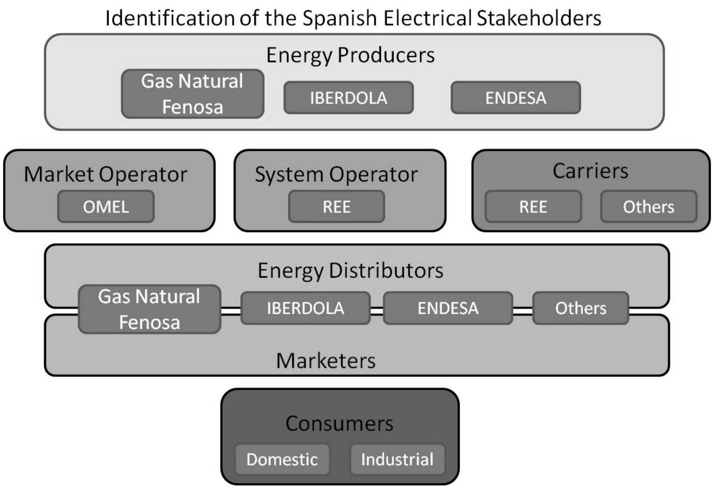


**Fig. 1.** The Spanish Electricity System in a market perspective.

The electricity transmission starts from power stations where energy is generated from various sources. The production is later transformed to a high-voltage and transported to REE, the system operator, through the transmission network. After that, it is transmitted from power substations, through an output line substation, to a transforming centre and is finally transformed to the needed voltage level for different consumption needs. Since electricity cannot be stored in large quantities, the whole process must work continuously without any interruptions. Moreover, it also needs to consider balancing the system to make sure that demand is met by supply. Consider Spain is the fourth biggest wind power producer. With intermittent generation, it is becoming even more difficult.

### 2.2 Incidents and the Situational Environment

The whole system is complex. Thus, several incidents can occur during the process to transport electricity from energy production plants to customers. These incidents will affect energy supply, lead to power cut and eventually generate negative impact to some stakeholders and the community. A 2007 blackout on 23rd July 2007 in Spain affecting 323,337 customers living in Barcelona area for more than 56 hours is seen as a prime example which resulted in huge fines, severe punishments and supplier switch [5].

Incidents have to be coordinated at a system level, typically involves the system operator REE and all energy distributors in the system where REE works as a coordinator to provide knowledge of problem to all distributing companies. This is because distributing companies may use each other's substations to provide services to their own clients in others' serviced areas and regions. Industrial users who are usually connected to the high voltage are also informed for the incidents by their suppliers bound to an "interruptibility" contract (ITC 2370/2007) but domestic customers are often ignored. As a result, call centres could be overwhelmed with an unprecedented number of calls in the event of an incident, leading to customer dissatisfaction. Therefore, following a customer-centric approach, all customers, rather than only the important ones, must be actively informed with the incidents and progress. However, energy suppliers show little interest in doing this due to technical difficulties and cost issues.

A misunderstanding is that informing clients should not be a difficult task for suppliers who also work as distributors in the system. The problem is that energy distribution and supply are operated as separate businesses due to strongly different market focus so the inter-connection between them does not exist. In addition, for a business where scattered information systems are normally used for different operations and functions, information interchange is also difficult. These issues subsequently make inter-organisational information exchange become more complicated.

## 3 Business Opportunities and Requirements

Business opportunities lie in the enterprise situational awareness (SA) which relies on a fast response to the emerging and/or changing situations. This requires instant access to resources of different systems and owners and rapid and flexible customer-oriented development. Consider the fact that key stakeholders in the Spanish Electricity System use different information systems for supporting their operations and activities. The following requirements for implementing an effective incident management application have been formed:

1. A list of affected streets should be obtained from a distributor.
2. A list of the affected customers should be obtained from suppliers.
3. Customers must be effectively informed with the problem, the forecast and the progress through available communication channels provided by both suppliers and distributor.

In addition, the key requirement of situational driven enterprise applications is that their initial development until in a working stage is reasonably simple and cheap, which means that little time must be spent in the development [7].

## 4 Cloud-based Data Integration Framework

As shown in Fig. 2, a cloud based data integration framework using a DaaS (Data as a service) model [9][10][11] is proposed. From the bottom up, it includes three different parts. First, *enterprise data*, which is supplied by different VE partners, is wrapped from the owners' different information systems and served as services embedded into an open/private cloud. Second, *a cloud infrastructure* is adopted to include all different functional services which can be later used to process data or can be composed for different data processes. Third, *mashups* are used to specify situational demands in the real business environment. Implemented mashups are also served as services within the cloud infrastructure.

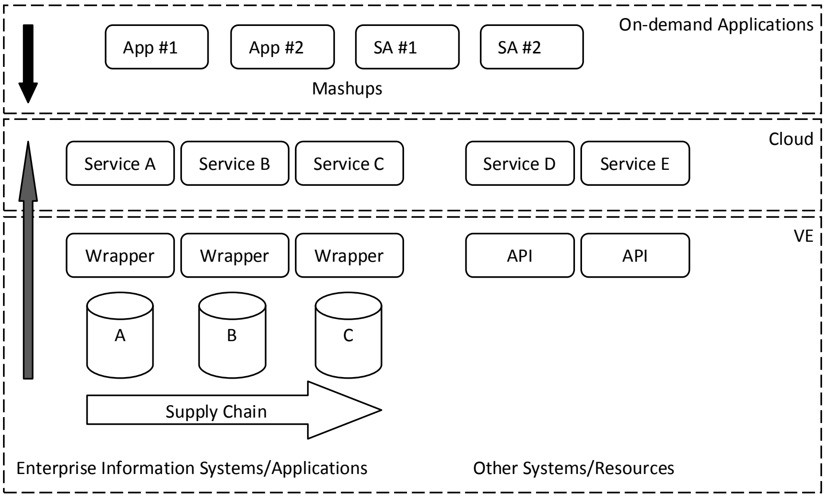


**Fig. 2.** The cloud-based data integration framework for VE.

*First Part: Data Wrapped as Services:* An organisation's master data are normally stored in different information systems and business applications used by the organisation. Such data are often shared and exchanged intra-organisationally to support the organisation's business functions and activities. With a number of appropriate processes such as data integration and MDM, some data can be extracted as individual services directly and made available in the virtual network (e.g., *Service A*, *B* and *C*). Since data access is controlled through the data services, it tends to improve data quality in an end-user perspective.

Sometimes the data services can be provided in a more convenient way with the integration of third-party applications. For example, Web mapping applications (e.g., Google Maps [6]) have been widely used by organisations to provide location-based services. Therefore, organisations may also consider creating their service in an open architecture through the support of third-party APIs (e.g., *Service D* and *Service E*) with business mashups.

*Second Part: Cloud as Service Infrastructure:* Cloud is considered as the service hosting infrastructure for VN to address two common limitations of VE: maintainability and flexibility [1]. First, a virtual network (VN) is a temporary collaborative network where participating organisations are dynamically allied in a customer-centric approach. This brings challenges to the long-term maintenance of the network as participants can stay, join or leave the network at any time. The loose-coupling feature of cloud makes changing the presentation layer of the virtual network is very cost-effective and much more feasible. Second, although VE is highly flexible as it optimises supply chain in a wider context, it aims to provide value-added services over existing services/activities rather than creating new services. This means that a traditional VE is not highly flexible to respond emerging business opportunities that require instant access of different resources. A cloud empowered VN can provide good agility due to the simplicity of the data access without the need for extensive knowledge of the underlying data. Additionally, cloud also makes it possible to merge

VEs to enter a new market as long as data access can be maintained at the presentation layer of the new cloud.

*Third Part: Mashups as Interface:* Mashups are used to manage inter-organisational data communications/activities. With common data access protocols provided by mashups, organisations and customers can access all available services on the cloud and create applications on-demand through the mix-n-match of different services that represent different business opportunities [12]. Moreover, mashups remove the internal boundaries in a dynamic supply chain formed in the VE so that any member organisation in the virtual network can access any available services from anywhere and form its own customised applications.

# 5 Implementation

In this section, we demonstrate how the proposed framework can facilitate on-demand applications, which actualises effective incident management for the Spanish Electricity System in the case described in the Section 3. In Section 5.1 we explain different data sources which related to the applications. Final implementation of the applications can be found from Section 5.2.

## 5.1 Data Sources

A distributor typically operates two distinct information systems, named SGC (*Sistema de Gestión y Control*) and BDI (*Base de datos integrada*). SGC is a management and control system containing a list of transformer centres (CT) where each associates with customers and a supplier. This allows the distributor to charge a supplier directly for the energy consumed by its customers and in turn allows a supplier to charge their customers with this information. BDI is an integrated database containing substation detail where each substation includes a list of positions inside a substation in which a CT list is attached to each position. As shown in Fig. 3, a substation-customer service can be created and published through data extraction and integration from SGC and BDI.

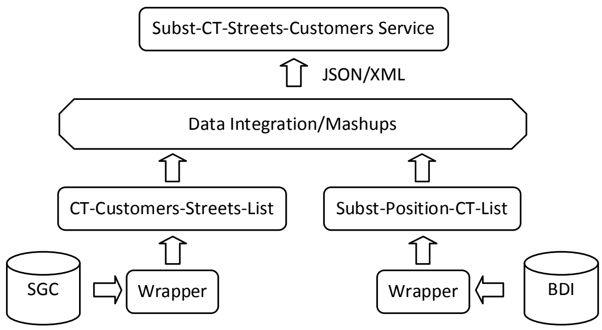


**Fig. 3.** Distributor's service integration for substation-customer.

Incidents are normally discovered through a distributor's SCADA system which can also be extracted as a situation service trigger as shown in Fig. 4.

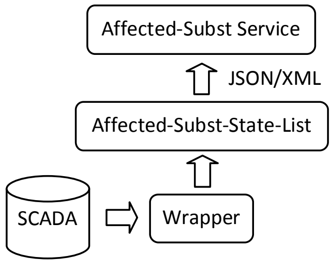


**Fig. 4.** Distributor's service integration for substation incident triggers.

Supplier stores information about their customers and marketing offers etc in its CRM, which can be extracted to form a customer notification service in conjunction with supplier's existing communication channels as shown in Fig. 5.

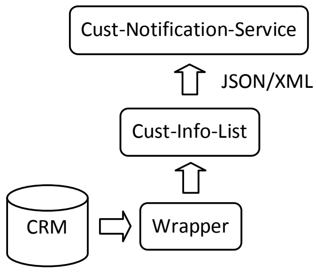


**Fig. 5.** Supplier's customer notification service.

### 5.2 On-Demand Application

Situational application [1]can be created by using mashups to mix-n-match the above services to notify customers and report progress when an incident occurs (Fig. 6).

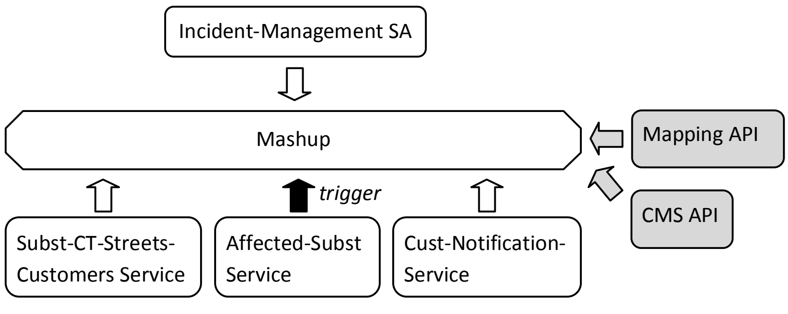


**Fig. 6.** Incident management SA based on the framework.

A better application can be further actualised in conjunction with third-party services.

[1] A working demo of the SA using this framework has been developed through the use of a public service cloud hosting and IBM Mashup Center 2.0.

For example, a CMS API can be integrated to provide incident information on all suppliers' websites and a mapping API can be integrated to provide interactive location-based views of incidents and progress. Additionally, the situational application can extend its communication channels through the integration of other social Web service APIs (e.g., twitter/facebook updates).

## 6 Conclusion

The core concept of the framework is that it uses DaaSs and mashups to help VE member organisations respond to immediate customer needs that require instant and easy resource access and rapid and flexible on-demand development. The case study has demonstrated how the framework facilitates in solving a real world problem effectively. Whilst it is based on the Spanish Electricity System but it can also be applied to other electrical systems especially for the European Union countries in compliance with Act 17/2007. Moreover, the situational application shown in the case study focuses on power incident management which can also be extended to other emerging application areas and industrial sectors where immediate communications to clients are needed.

**Acknowledgments.** This work is made possible by the support of the Natural Science Foundation of China (NSFC) under Grant No.61150110484, ESSENTIAL: Enterprise Service deSign based on ExistiNg software Architectural knowLedge.

## References

1. Martinez, M.T., Fouletier, P., Park, K.H., Favrel, J.: Virtual Enterprise – Organisation, Evolution and Control. Intern. Journal of Production Economics, 74(1-3), 225-238 (2001)
2. Camarinha-Matos, L.M., Afsarmanesh, H.: A Comprehensive Modeling Framework for Collaborative Networked Organizations. The Journal of Intelligent Manufacturing, 18(5), 527- 615 (2007)
3. National Grid, http://www.nationalgrid.com
4. Red Eléctrica de España, http://www.ree.es
5. Barcelona blackout may last weeks for 80,000, www.msnbc.msn.com/id/19921787/ns/world_news-europe/t/barcelona-blackout-may-last-weeks
6. Google Maps API, https://developers.google.com/maps
7. de Vrieze, P. T., Xu, L. and Xie, L.: Situational Enterprise Services. In: Encyclopedia of E-Business Development and Management in the Digital Economy, pp. 892-90. Idea Group Publishing, Hershey, PA (2010)
8. O'Reilly, C. A., Harreld, J. B., Tushman, M. L.: Organizational ambidexterity: IBM and Emerging Business Opportunities. California Management Review, 51(4), 75-99 (2009)
9. Mateljan, V., Cisic, D., Ogrizovic, D.: Cloud Database-as-a-Service (DaaS) – ROI. In: 33rd International Convention on MIPRO, pp.1185-1188. IEEE Press, New York (2010)
10. Wang, L., von Laszewski, G., Younge, A, He, X., Kunze, M., Tao, J., Cheng, F.: Cloud Computing: a Perspective Study. New Generation Computing, 28(2), 137-146 (2010)
11. Youseff, L., Butrico, M., Da Silva, D.: Towards a Unified Ontology of Cloud Computing. In: Grid Computing Environments Workshop (2008)
12. de Vrieze, P. T., Xu, L., Bouguettaya, A., Yang, J. and Chen, J.: Building Enterprise Mashups. Future Generation Computer Systems, 27 (5), 637-642 (2011).