Integrating Distributed Data Storage of a Smart City Public Administration

Carmelo Pino, Salvatore Ravidà and Santo Scibilia
University of Catania, Department of Electrical, Electronics and Computer Engineering (DIEEI), V.le Andrea Doria, 6, 95125 Catania, Italy

Keywords: Smart City, e-Government, Cloud Computing, Data Storage Integration, RDF.

Abstract: The developments and increasing solutions offered by Cloud Computing technologies represent a great opportunity for e-governance of smart-cities. The benefits provided by Cloud-based technologies allow us to re-design new services to solve many problems afflicting e-governance and public administration in each country. In particular, one of the primarily need of Public Administration (PA) is the interoperability between the different areas and districts, which is rather complex due to the lack of standard schemes for data modeling. In this paper we propose a cloud-based approach for supporting such interoperability by a customizable DBMS mapping tool especially dedicated to smart city governance.

1 INTRODUCTION

As known, a smart city arises when citizens are ready to use technologies to improve relevant city life aspects such as governance, local development, cultural events, mobility and logistics, environment and health services (Chourabi et al., 2012), (Berthon et al., 2011). Cloud computing can be considered as a collection of technologies to store or process data through the use of distributed hardware/software over a virtualized Network. It's a model for structuring IT resources that redefines the way to manage computer systems (Rimal et al., 2009), (Dillon et al., 2010). Nowadays, the diffusion of Cloud Computing and the perception of its benefits represents an opportunity to increase the capabilities of the e-governance in term of offered services, cost reduction, distributed data storage, security management (Mukherjee and Sahoo, 2012), (Ramgovind et al., 2010), scalability, accountability, interoperability (Cellary and Strykowski, 2009), and reliability (Chandra and Bhadoria, 2012).

Among the different challenges in e-governance the interoperability between the existing hardware, software and data storage is a major issue (Smitha et al., 2012) that often systems providing e-services in different administrative units are not able to guarantee.

There exists several e-governance applications and services that can be categorized at different levels: Government to Government (G2G), Government to Business (G2B), Government to Citizens (G2C) and Government to Enterprise (G2E) (Jeong, 2007), (Giordano et al., 2013), (Faro et al.,2008), (Constanzo et al., 2012). Each category is related to a specific set of applications and the common need is the communication between multiple data sources (e.g. the storage related to "municipal maintenance" has the need to retrieve data about a citizen from the "registry office storage"). The miscellaneous of data contained in different archives suggests the use of RDF or OWL to define a common semantic schema to map the different terms used in the various data bases with the same meaning to a standard vocabulary of concepts.

Interoperability is therefore one of main issue for the development of future e-government systems, called in following as e-gov systems, which can be tackled with the use of cloud computing (Tripathi and Parihar, 20119) and the advantages provided by its model of services (SaaS, PaaS, IaaS) and distribution (Public, Private, Hybrid) (Armbrust et al., 2010).

In this paper we propose a private cloud architecture for supporting the interoperability and integration between distributed e-gov data storages. The idea behind this work is to collect each data storage

Pino C., Ravidà S. and Scibilia S..
Integrating Distributed Data Storage of a Smart City Public Administration.
DOI: 10.5220/0004407802390243
In *Proceedings of the 2nd International Conference on Smart Grids and Green IT Systems* (SMARTGREENS-2013), pages 239-243
ISBN: 978-989-8565-55-6

belonging to a Administrative Unit (AU) in a private cloud architecture (without changing the storage structure) and to create a layer hiding the storage details so that it is possible to view the distributed data storage as a centralized storage that exposes the data according to a standard data model provided by a RDFs/XML schema.

This data schema for e-gov data management has been created analyzing the structures of several data storages belonging to different AUs. Then, the cloud has been used as in the Software as a Service (SaaS) to offer to citizens and employees the possibility to manage and retrieve data residing on different nodes using a single front end. In section two, the details about the developed architecture are discussed, whereas some examples of how using the proposed cloud are shown in section three. Concluding remarks are given in the last section.

2 THE PROPOSED ARCHITECTURE

The proposed system is based on a general architecture that supports the interoperability between distributed data storages belonging to the same or different AUs, e.g. registry office, garbage fee, cadaster, etc. It consists of two main modules: 1) Data mapping and exposing, 2) Data querying. The first module creates a transparent layer hiding the details of each storage in terms of DBMS and data storage schemata. The second module aims at: a) retrieving the data exposed by the previous module, and b) integrating such data according to the user's query. More in details:

- **Data Mapping and Exposing (DME) Module:** it executes the mapping and returns the data according to a standard vocabulary, i.e. also called data ontology. It's installed on each storage node to interface the query issued by the users to the local DBMS. Currently, the local DBMSs considered in our approach are MySql, Oracle and DB2, but other DBMSs will be considered in future by simply attaching the relevant driver to the DME module.
- **Data Querying (DQ) Module:** This module behaves as a front end for the user and represents the consumer of the service installed on each storage. It communicates with each VM in order to perform a specific query and to obtain the results. The XML data returned by the different VMs represent different views and are parsed, collected and stored locally.

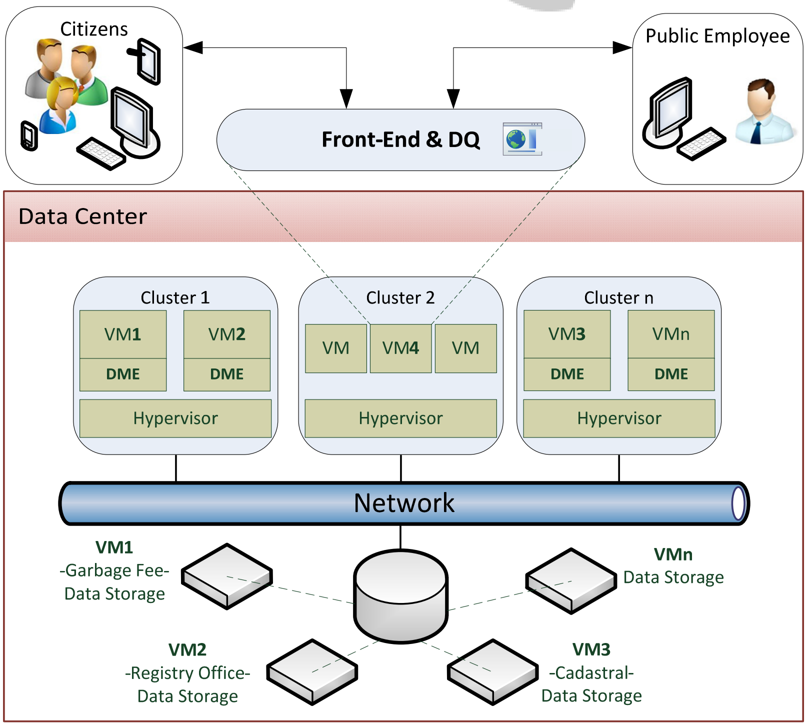

Figure 1: The proposed architecture: each VM has its storage and hides storage details through the DME tool. The DQ modules resides on the Front-End using the DME tools.

In our private cloud architecture these two modules are used by each Virtual Machine (VM) to expose the data through the web service described above. Another VM houses the service that communicates with each VM in order to perform the requests (Front-End). It interacts with the user through a specific interface to authenticate the request and to query the distributed data storage. Each request is redirected from the front end to the relevant VMs that retrieve the data and return them to the front-end which takes care of the data integration.

Figure 1 shows the proposed architecture. An example of usage for public employees could be: *“search for cadastral situation related to a citizen called ‘Mario Rossi’ and how many people live together with him”*. The front end implemented by a specific VM, i.e., VM4 in figure 1, receives and parses the request identifying the VMs which expose the cadastral and registry data, i.e., VM2 and VM3 in figure 1. The data querying module residing on the front end integrates the retrieved data and provides the information (always updated to the latest change) to the user in real time.

Our private Cloud infrastructure is based on Microsoft System Center 2012[1] and is composed by an HP C7000 enclosure with 8 blades for a total of 64 processors and storage capabilities of 20 TB. Each VM is managed by Hyper V for a total of 4 VMs respectively related to: the garbage fee storage (VM1), the registry office storage (VM2), cadastral storage (VM3), the front end (VM4). In VMs 1, 2 and 3 we have installed the DME tool. DQ is obtained by a module residing on the front end cooperating with the DMEs residing on the VMs. The account, privileges and access policy are configured and managed with system center 2012.

3 SOME EXAMPLES

In this section we illustrate two examples of use. The first example regards the configuration of a DME related to a storage. This is an essential step to select which data will be exposed and mapped.

Fig. 2 shows the DME configuration step: an AU administrator configures the DME through a graphical interface specifying the DBMS type and a set of queries that will be exposed as views. The queries that will be exposed can be specified using SQL, and will be executed when the front-end sends a request to the DME. The data obtained from each query are mapped and returned according to an ontology.

Fig. 3 shows an example of interaction between the front-end and different DME modules. When a user (citizens or public employees) performs a request through the graphical interface, the front-end identifies which are the DMEs useful for the data retrieval. For example, if the request is: *“search for garbage fee and cadastral situation related to a citizen called ‘Giuseppe Bianchi’”*, the Front-End understands that three storages are involved in this search, i.e., garbage fee, cadastral and registry fee.
For this reason it sends a request to three DMEs for a specific query. Each DME return the data related to the request in RDF format. The returned data are integrated by the front-end thanks to the common RDF schema used.

4 CONCLUSIONS AND FUTURE DIRECTIONS

In this work we have sketched a cloud-based e-gov system for the integration of PA storages which use different DBMS and schemas. Our solution has been tested in public district of Catania (Italy) for the following archives related to: registry office, garbage fee and cadaster. This architecture is a practical solution to integrate and retrieve data from multiple sources.

As future work we plan to extend the ontology for any type of data involved in e-government applications and to publish this data as part of linked data. Moreover a hybrid cloud solution will be tested in order to possibly improve the efficiency of the proposed approach thanks to the advantages of a consolidated private cloud in term of performance and computing power.

[1] http://www.microsoft.com/en-us/server-cloud/system-center/default.aspx

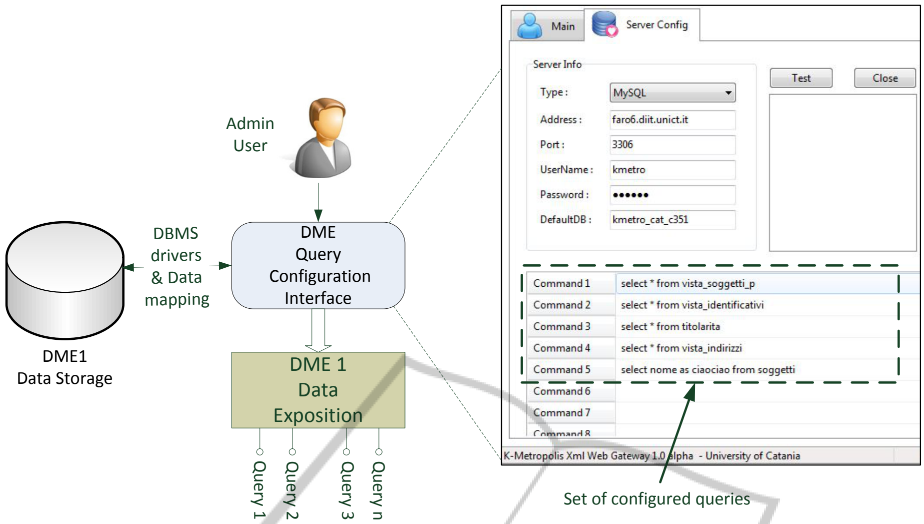

Figure 2: Example of DME configuration.

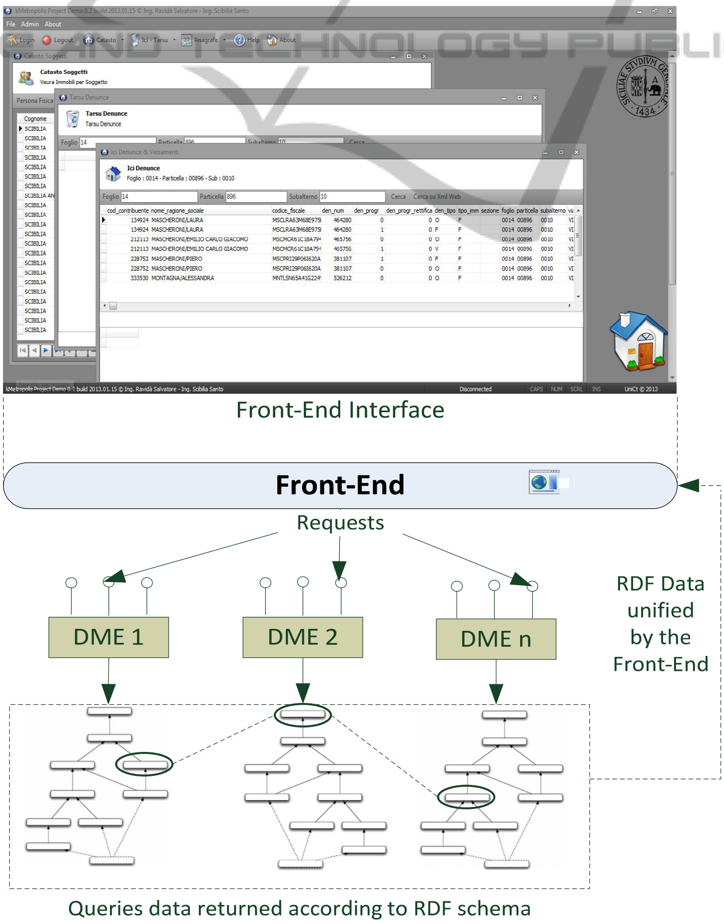

Figure 3: Example of Interaction between Front-End and DME modules.

REFERENCES

Chourabi H., et al., "Understanding Smart Cities: An Integrative Framework", 45th Hawaii International Conference on System Sciences, 2012

Berthon B., Massat P., Collison S., "Building and Managing an Intelligent City", http://www.accenture.com/SiteCollection Documents /PDF/AccentureBuilding-Managing-Intelligent-City.pdf, 2011

Rimal, B.P.; Eunmi Choi; Lumb, I., "A Taxonomy and Survey of Cloud Computing Systems", Fifth International Joint Conference on INC, IMS and IDC. NCM '09 Aug. 2009

Dillon, T.; Chen Wu; Chang, E., "Cloud Computing: Issues and Challenges," 24th IEEE Int. Conf. on Advanced Information Networking and Applications (AINA), April 2010

Mukherjee, K.; Sahoo, G. , "A novel methodology for security and privacy of cloud computing and its use in e-Governance" , World Congress on Information and Communication Technologies (WICT), 2012

Ramgovind, S.; Eloff, M.M.; Smith, E., "The management of security in Cloud computing", Information Security for South Africa (ISSA), 2010

Cellary W., Strykowski S., "E-government based on cloud computing and service-oriented architecture", 3rd International Conference on Theory and practice of electronic governance (ICEGOV '09), Tomasz Janowski and Jim Davies (Eds.). ACM, New York, NY, USA, 5-10, 2009

Chandra D.G., Bhadoria R.S., "Cloud Computing Model for National E-governance Plan (NeGP)", Fourth International Conference on Computational Intelligence and Communication Networks (CICN), 2012

Smitha K.K., Thomas T., Chitharanjan K., "Cloud Based E-Governance System: A Survey", Procedia Engineering, Volume 38, Pages 3816-3823, 2012

Jeong Chun Hai @Ibrahim, "Fundamental of Development Administration", Selangor Scholar Press, 2007

Giordano D., Torre A., Samperi C., Alessi S., Faro A., "An Ontology based Approach to Join E-Government Data in a GIS framework: a Case Study", IEEE 4th International Conference on Software Engineering and Service Science, ICSESS, 2013

Faro A., Giordano D., Spampinato C., "Evaluation of the traffic parameters in a metropolitan area by fusing visual perceptions and CNN processing of webcam images", IEEE Transactions on Neural Networks, vol.19 (6), 1108-1129, 2008

Costanzo A., Faro A., Giordano D., Venticinque M., "Wi-City: A federated architecture of metropolitan databases to support mobile users in real time", International Conference on Computer and Information Science, ICCIS 2012, A Conference of World Engineering, Science and Technology Congress, ESTCON 2012

Tripathi A., Parihar B., "E-Governance challenges and cloud benefits", IEEE International Conference on Computer Science and Automation Engineering (CSAE), 2011

Armbrust M. et al., "A view of cloud computing", Communications of the ACM, Vol.53(4), 2010