Volume 5, Issue 4, April 2015 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

Open System Architecture Platform for Big Data: An Integrated Emergency Disaster Response System Architecture

Quist-Aphetsi Kester[1, 2], William Sam-Aggrey[1]
[1]Ghana Technology University College
[2]Lab-STICC (UMR CNRS 6285), European University of Brittany, University of Brest, France

Abstract— ***In situations such as emergency where urgent and active response is needed to access the case and deploy relief or support with effective service in real-time and notices drawn and effectively distributed, easy to use applications with disparate devices and wide range coverage are necessary for sharing information as well as reporting incidence based activities. Such activities may be accidents like plane crashes, train collision, road accidents, etc., natural disasters such as landslides, earthquakes, tsunamis, hurricanes, floods, volcanic eruptions etc. An integrated application is needed to accommodate different forms of reports involving the engagement of disparate devices, and hence we designed an approach of integration based on concepts of service oriented architecture and service busses with capabilities of accessing different information from disparate devices for effective analysis. In this situation, areas without telecommunication infrastructure that support internet facilities can also be hooked into the grid as well as devices without internet capabilities. We proposed*** **an** ***Open System Architecture Platform for Big Data: An Integrated Emergency Disaster Response System Architecture for an effective integration of applications and software systems with open standard format to help manage and made easily available disaster response systems data for easy usage by systems in an integrated manner and made available for access usage globally by software systems and publicly as well, so that other information systems can have access to the available or published data in real-time.***

Keywords— software architecture, disaster response system, service oriented architecture, service bus, disparate devices, open system, big data, information systems

I. INTRODUCTION

Disaster prevention, early warning systems and disaster preparedness is one of the top priorities of human institutions in working towards spontaneous likely occurrences of natural disasters. Our institutions aim to understand nature and document observations overtime so as to help us predict occurrences and avoid them. Humans have tried to learn the secrets of nature through science and have built computers in understanding and simulating data to forecast the next natural activities with the high hopes of avoiding them from having negative impact on our human ecosystem or trying to reduce the impact by making use of counterforce by engaging the same forces of nature in our defences. But our current activities overtime in advancing our world through technological advancement has triggered natures counter balance forces in resolving these issues with its forces and hence rendering a sharp and highly unpredictable devastating effect on our ecosystem. Studies carried overtime has shown that our activities as humans has increased the effect of global warming and forcing natures forces to cause effects like flooding, high temperatures etc. Disaster warning systems pick data from observations to predict the next occurrence of an activity and provide an early warning of an activity. These systems may pick data over satellite communications, earth observatory systems, sensor networks etc and analysed it to provide visuals or indications of occurrences. Responses before, during or aftermath of the effects has to involve continues human communications in order to help deploy relief, report the extent of damage, current developments, as well as likely events. These systems that help convey these important messages have to be integrated and open in such a way that data can be made available for systems as well as humans to help ascertain as well as contain the situation and its impact. The system should be able to have integrated feature to receive both formatted and unformatted data and have capabilities of processing them and presenting them in an open standard way so as to make easy mining of such data by both machines and humans.

During disaster and most often, news has to be obtained by people via television stations globally or locally as well as via other forms of media such as radio, print media, social network platforms, government communications etc. These processes of communication delay rallying of global support as well ascertaining the real situations on the grounds effectively so as to support with the right aid. And also real time information about current events, needs of people and geographical information have to be obtained differently from different platforms hence making availability of vital data very difficult in summarizing the exact situation occurring on the grounds. Services such as aid urgencies, security services, medical services, other NGOs and other manufacturing organisations etc need these data and also have to make use the data in providing an effective service for responses. For instance in a situation of an earthquake, the known geographical area needs to be made available with the exact coordinate as well as building information, road networks, weather conditions, water availability, population estimate of affected area, periodic update of evacuation plans, public and local plans of resolving the situation, progress of work minus aid estimate, possible affected people, hospital aid required etc. These amounts of data is necessary for making an effective budget for aid agencies, making orders easily accessible, information on nearest available stock of goods, need for sanitary aids to control spread of another case of contamination or spread of disease, water treatment technologies needed due to availability of water source, nature of environment so as to support it with the right clothing, tents, medication etc. This is because; an earthquake activity in West Africa will require a different aid of from the one at northern Europe. Data available in a pull of other big data environment can make an easy access, easy deployment as well as easy understanding o the physical environment as well as needs.

In providing an effective way of information processing for this situation in order not to waste resources as well as not to under provide relief, an information system plat form is needed to remedy situation. Henceforth in our work, we provided an architectural platform based on open standards and with the concepts of service oriented architecture to provide a data situation in which such information from such systems and during a disaster are published and made available and easily connected to the big data grid in such a way that other systems can easily mine data from it and use it to effectively resolve cases.

II. RELATED WORKS

Data analytics is an effective way of mining information from huge deposits of data residing in a distributed environment. Data generation within today's cyberspace is enormous and greatly important for decision making. Daily real-time generation of data from sensor networks, social media, mail messages, text messages, business, satellites, airlines, travel agency systems, health information systems, etc can play a very active role in effective decision making, understand migration trends, discover new opportunities etc.

McAfee, A., & Brynjolfsson in their work of Big data: the management revolution acknowledged in their writing that Big data is far more powerful than the analytics of the past and executives can measure and therefore manage more precisely than ever before making decisions compared to their previous practices. They can make better predictions and smarter decisions in a minimum amount of time and also forecast ahead. They can target more-effective interventions in areas that so far have been dominated by gut and intuition rather than by data and rigor. The differences between big data and analytics are a matter of volume, velocity, and variety: More data now cross the internet every second than were stored in the entire internet 20 years ago. Nearly real-time information makes it possible for a company to be much more agile than its competitors. And that information can come from social networks, images, sensors, the web, or other unstructured sources. The managerial challenges, however, are very real. Senior decision makers have to learn to ask the right questions and embrace evidence-based decision making. Organizations must hire scientists who can find patterns in very large data sets and translate them into useful business information. IT departments have to work hard to integrate all the relevant internal and external sources of data. The authors offer two success stories to illustrate how companies are using big data: PASSUR Aerospace enables airlines to match their actual and estimated arrival times. Sears Holdings directly analyzes its incoming store data to make promotions much more precise and faster [1]. Emergence of an era of Big Data is critical in this phase of technology evolution. This big data phenomenon is taking place in an environment of uncertainty and rapid change, current decisions will shape the future. With the increased automation of data collection, processing and analysis as well as algorithms that can extract and illustrate large scale patterns in human behaviour. [2]. Lessig argues that social systems are regulated by four forces: market, law, social norms, and architecture – or, in the case of technology, code [3]. When it comes to Big Data, these four forces are frequently at odds. The market sees Big Data as pure opportunity: marketers use it to target advertising, insurance providers use it to optimize their offerings, and Wall Street bankers use it to read the market. Legislation has already been proposed to curb the collection and retention of data, usually over concerns about privacy (e.g. the US Do Not Track Online Act of 2011). Features like personalization allow rapid access to more relevant information, but they present difficult ethical questions and fragment the public in troubling ways [4]. Scalable database management systems both for update intensive application workloads as well as decision support systems for descriptive and deep analytics are a critical part of the cloud infrastructure and play an important role in ensuring the smooth transition of applications from the traditional enterprise infrastructures to next generation cloud infrastructures. A system in the cloud must possess some features (often referred to as cloud features) to be able to effectively utilize the cloud economies. These cloud features include: scalability, elasticity, fault-tolerance, self-manageability, and ability to run on commodity hardware. Figure 1 provides a schematic representation of scalable database management systems to support large applications [5].

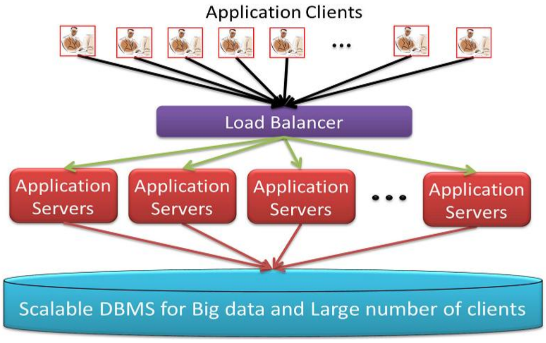

Fig. 1 Scalable database management systems to support large applications with lots of data and supporting hundreds of thousands of clients [5]

Big data is a data sets having large, more varied and complex structure with the difficulties of storing, analysing and visualizing for further processes or results. The process of research into massive amounts of data to reveal hidden patterns and secret correlations named as big data analytics. This provides useful information for companies or organizations with the aim of gaining richer and deeper insights and getting an advantage over the competition [6]. The next innovations phases in technology are around technology architecture and this yields three major trends. The first one of them is innovation trend and this is to integrate hardware, software, and service into one box. This type of innovation can hide or provide several abstraction layers to the complexity of the enterprises' IT systems and make them cloud ready. Most importantly, the services aspect of the box can capture best practices of operations in different scenarios and also support disparate applications and devices. This box also provides a development toolkit or platforms to enable developers to import knowledge and experiences into the integrated black box. The second innovation trend is to provide open cloud platforms with open standards by leading Internet service providers. Application engines, big data analysing algorithms, cloud storage APIs, and prediction and translation APIs are gradually becoming important tools and assets for open innovations. These provide effective integration platforms for easy configuration and interoperability. The third technology architecture innovation comes from mobile Internet and social networking services. Technology architecture is used to realize application architecture and data architecture, which further supports business architecture. Massive data is generated on these platforms everyday which runs in line with the true activities of the happenings in society to some extent and the data harnessed in this domain can be effectively applied in many ways for boost businesses and to forecast other variables and monitor certain trends in the society. The last innovation is around architecture governance for cloud computing. It includes the standardization of individual architectural building blocks within business architecture, application architecture, data architecture, and technology architecture [7].

Big data approaches are based on the concept of adopting acronyms of Vs as shown in figure 2 below with five Vs [8]. But as time passes the basis of the Vs continue to increase in order to accommodate new approaches to adopt its effective usage.

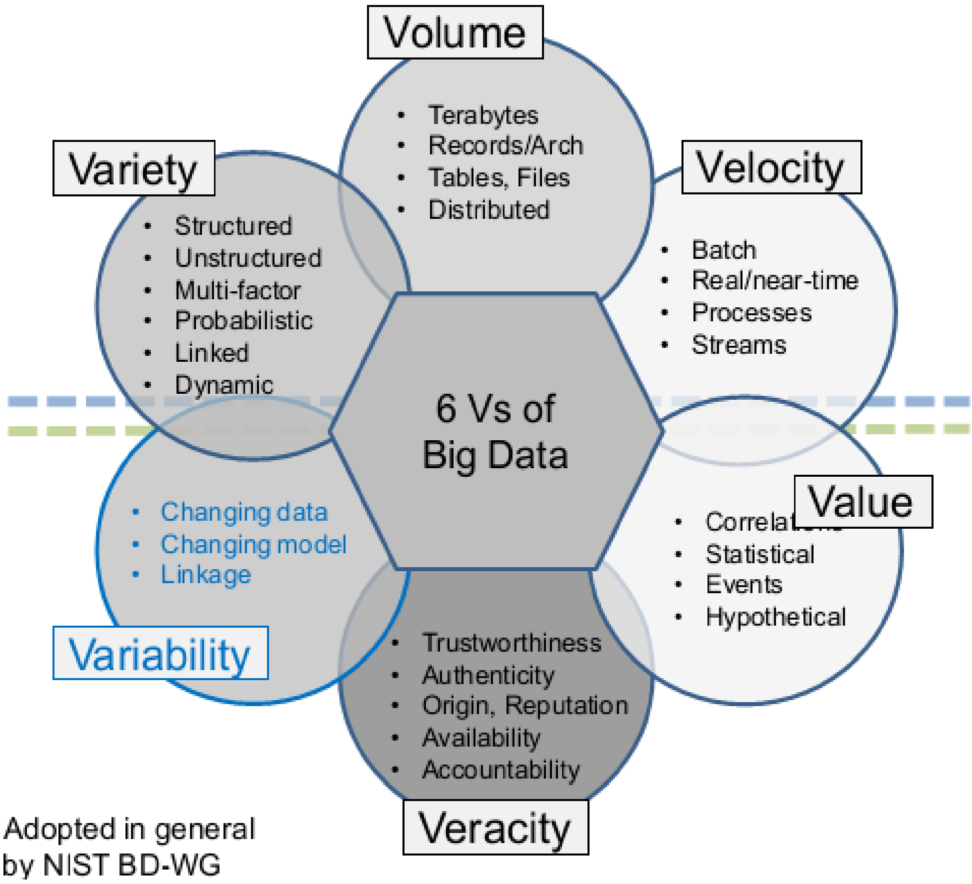

Fig. 1 The 6 Vs of big data

Refining Gartner definition "Big data is high-volume, high-velocity and high-variety information assets that demand cost-effective, innovative forms of information processing for enhanced insight and decision making" Big Data (Data Intensive) Technologies are targeting to process high-volume, high-velocity, high-variety data (sets/assets) to extract intended data value and ensure high-veracity of original data and obtained information that demand cost-effective, innovative forms of data and information processing (analytics) for enhanced insight, decision making, and processes control; all of those demand (should be supported by) new data models (supporting all data states and stages during the whole data lifecycle) and new infrastructure services and tools that allows also obtaining (and processing data) from a variety of sources (including sensor networks) and delivering data in a variety of forms to different data and information consumers and devices [8][9].

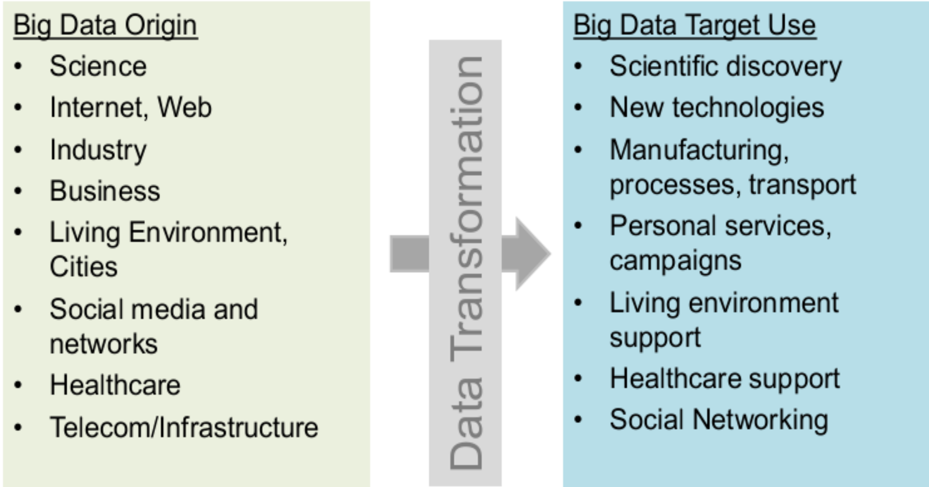

Fig. 3 Big Data Nature: Origin and Target (consumers)

Big data can have its source generated from different sources in a distributed environment across different geographical areas. These data sources can efficient in the analysis of different cases. For instance a hospital data of birth and death can give a raw figure about population in a geographical area. The number of mobile phone users of a particular geographical area can be determined by tacos and banks can produce savings and other forms of data such as work force, education etc can provide a lot of insight of that geographical area. All these data are crucial in opening up new businesses, rendering certain services etc. Below are Big Data Nature: Origin and Target (consumers)[9]

III. METHOD

Data acquisition is very vital in processes of understanding situations on the grounds and preparing for them. Disasters are unpredictable or foreseen but uncontrollable events in or ecosystem. Preparedness for them puts us at the advantage of avoiding losses and reducing its impacts. Early warnings are very critical as well as monitory and tracking of disaster effects too are very important. Hence data acquisition before, during and after its occurrence is very vital to the developments of productive of effective counter measures. Urban and rural areas affected by natural disasters suffer related effects. Different technological reporting systems are adopted in their management. In providing part of the solution via software platforms and approaches, we proposed an architectural platform based on open standards and with the concepts of service oriented architecture to provide a data situation in which such information from such systems and during a disaster are published and made available and easily connected to the big data grid in such a way that other systems can easily mine data from it and use it to effectively resolve cases. With this approach, disparate applications can easily connect and send data as well as receive information as to what to do and how to avoid the effects as well. We focused more on the architectural platform. Below are figures as well as the architecture adopted for our subsystems as well as the grid.

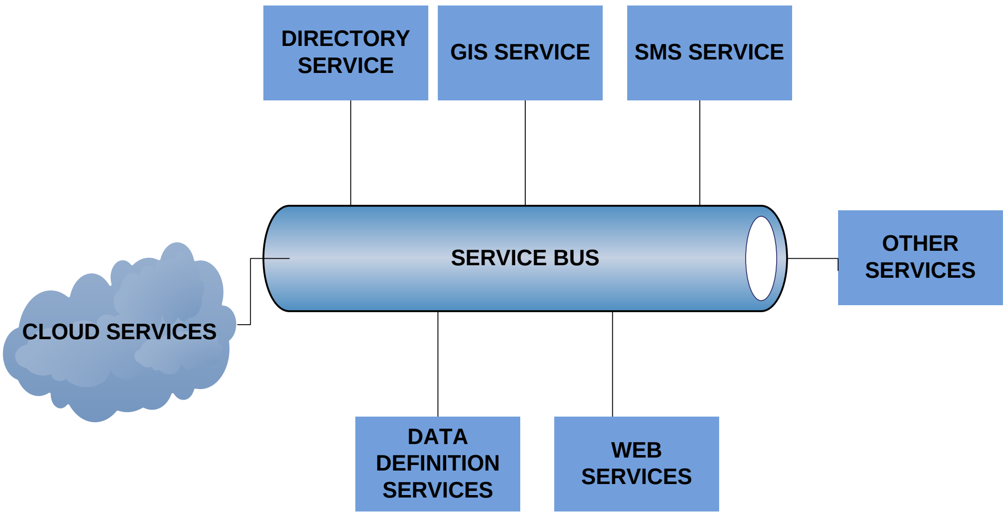

Fig. 4 Service Bus Architecture for integration of services

The service bus in figure 4 provides an effective service to users and other systems. The web service deliver request to systems that wants web based formats of data and it also publishes information in web formats. The data definition service manage request services and check formats to see whether they conform to the services that one is trying to access. The Cloud service provides the platform for which data is exchange between the service bus and other services based in the cloud. The directory service is responsible for forwarding request to the appropriate services based on requests and routine schedules. The GIS service provides geographical services to applications as well as map names with coordinates. The SME service is responsible for receiving and processing of information sent by mobile phones via such service. It also sends data through same. This service can also resolve short codes and names of locations against geographical locations for web service publications. For instance a location where internet service s not working but SMS service is available can send updates to the server via short codes and location name. The location name is then resolved with geographical coordinates to locate the region on a visual map produce by the web service or other application service. This information can then be stored in a database, relayed to other people or systems but accessed via the cloud etc.

Since our approach used was to get all the independent and disparate systems residing within each geolocations and spread over to get connected so that they can easily exchange data irrespective of the platform upon which they are running on and make the services available for usage by external systems during transactions. A multi-layer architecture approach based on service oriented architecture (SOA) was used to underpin the platform architecture. With this approach, upgrade becomes easy since modularity, orchestration and choreography plays a key for an easy integration of new services.

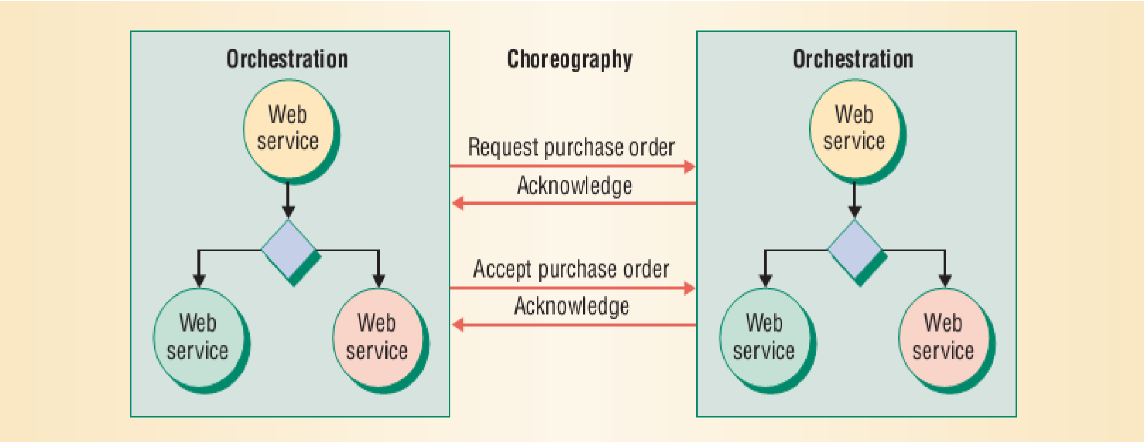

Fig. 5 Orchestration versus Choreography

With the approach higher-level services from existing orchestrated processes exposes these processes through their own Web service interfaces and Business Process Specification Schema (BPSS) protocol can define both the choreography and communication protocols between Web based services as shown in figure 5[10]. In order to deal with collections of services in case they became cumbersome, we adopted a concept and framework of service-level abstraction so that future different modules or services can deal with collections of services, rather than individual services. Service-level abstractions deal exclusively with services and they define all the important elements of the services. This approach was based on the work of Kester et al [11] and shown in figure 7 below. A service bus based on SOA is responsible for providing integration, message brokering, and reliability functionalities to the entire network. The systems bundle together adapters for connectivity, data transformation engines convert data to an appropriate format for usage by requesters, modular integration engines handle many different complex routing scenarios simultaneously, and other components present a unified integration solution. A central integration mechanism was used to further consolidate tasks within the service oriented integration architecture shown in figure 7. This create more room for a more flexible architecture, where new components can be added and removed as needed, by changing the configuration of the provider by simply developing modular components making a single service to be easily reused by multiple applications [12].

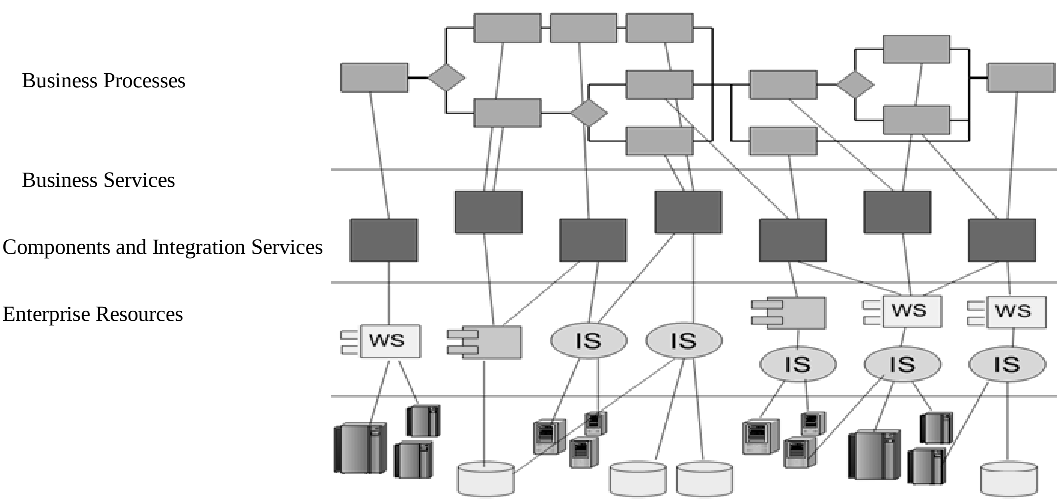

Fig. 6: Service Oriented Integrated Architecture

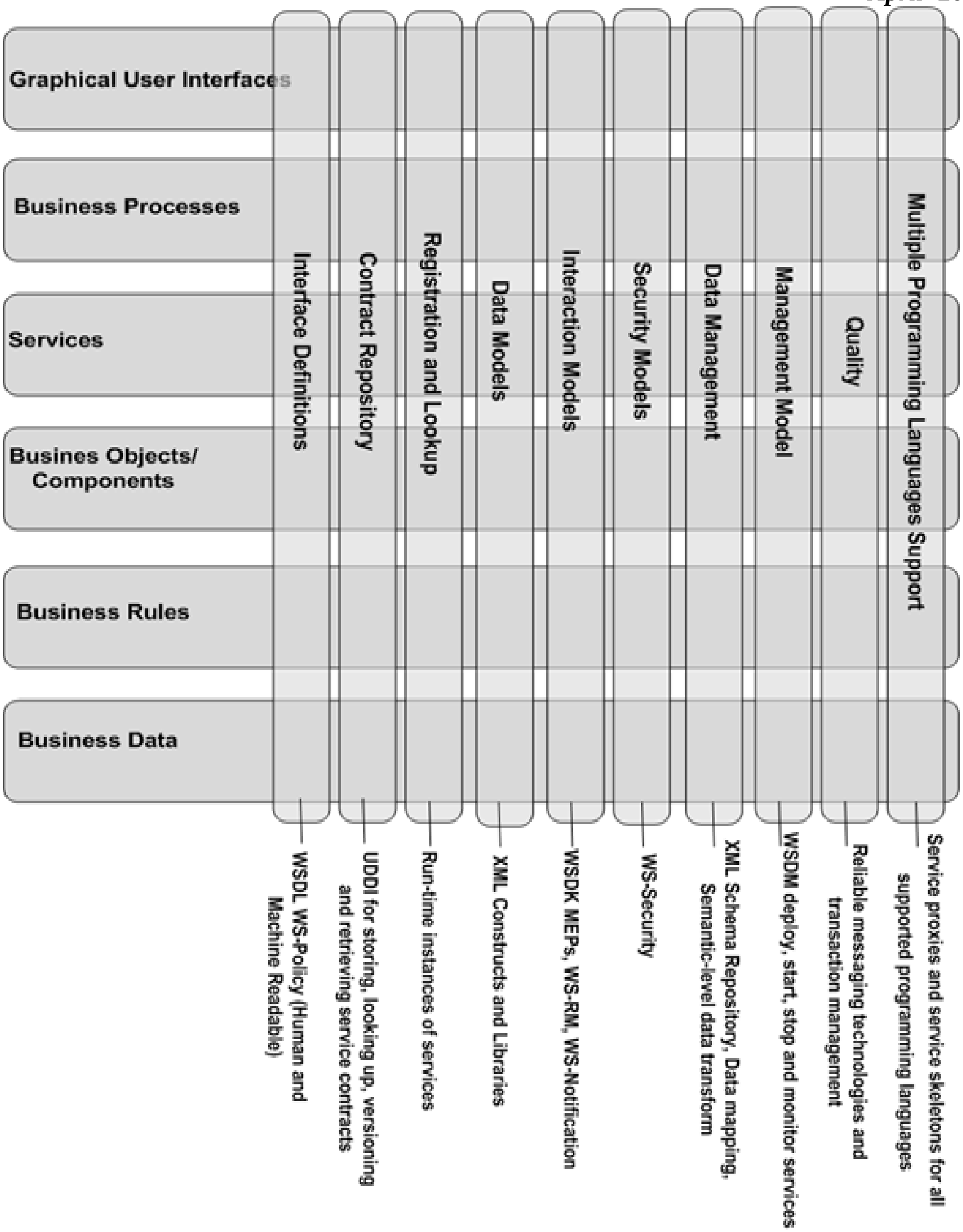

Fig.7: the service-level abstraction

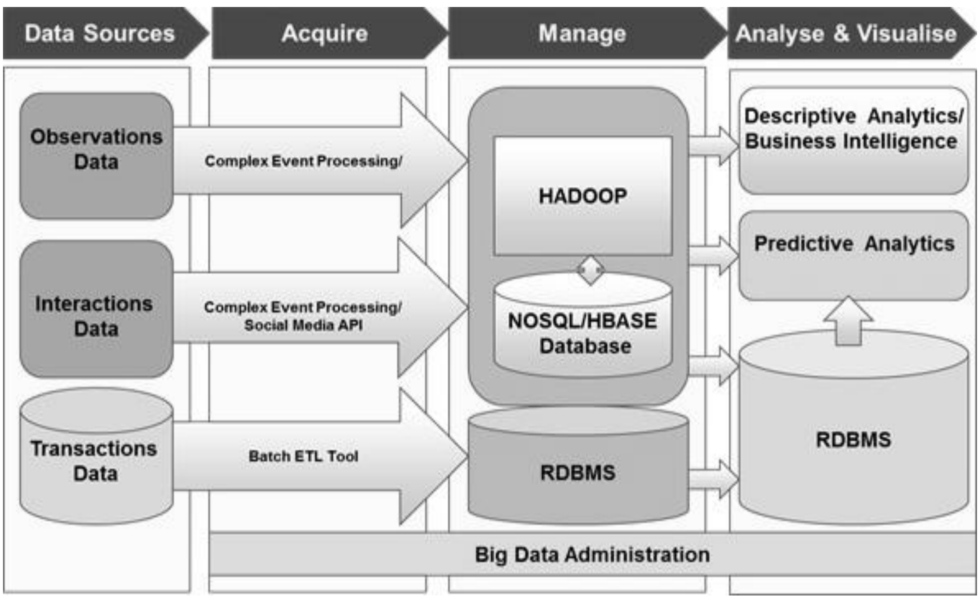

Fig.8: A high level typical Big Data solution architecture

The data acquisition of the approach is achieved via cloud to make way for data to be extracted as a transaction data for the high level architecture above in figure 8 for usage [13]. Transactions Data may be a historical transactions data from core observations, application systems etc Interactions Data may be social media data, Web content etc. And observations Data may be from sensor networks, GPS, RFID, Mobile, etc. Below is the depiction of the network environment.

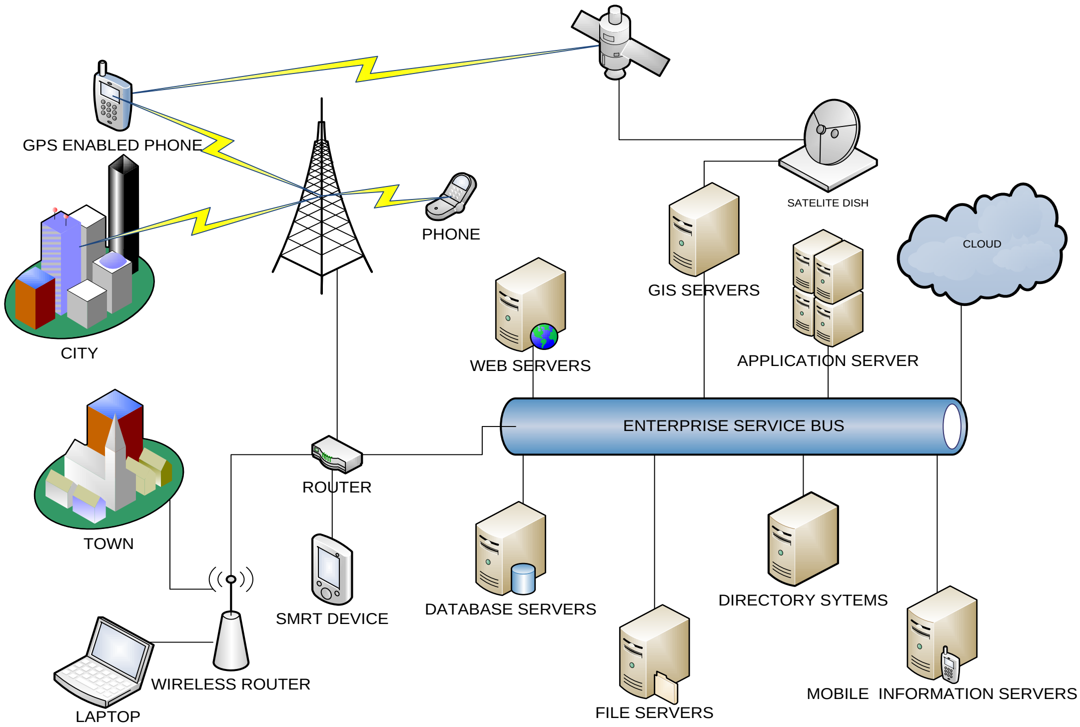

Fig.9: Depiction of the network environment

Our adoption of Open Systems Architecture (OSA) was based on the fact that "Open Systems Architecture integrates business and technical practices that yield systems with severable modules which can be competed. The essence of OSA is organized decompositions of complex systems, using carefully defined execution boundaries, layered onto a framework of software and hardware shared services, and a vibrant business model that facilitates competition. It requires publishing of key interfaces within the system and design disclosure. A key enabler for OSA is the adoption of an open business model which requires doing business in a transparent way that leverages the collaborative innovation potential of numerous participants across the enterprise, permitting shared risk, maximized asset reuse, and reduced total ownership costs. OSA yields modular, interoperable systems which more easily support component addition, modification, or replacement by different vendors throughout the lifecycle, driving opportunities for enhanced competition and innovation." [14]

OSA is composed of five fundamental principles [14]:

1. Modular designs based on standards, with loose coupling and high cohesion, that allow for independent acquisition of system components;
2. Enterprise investment strategies, based on collaboration and trust, that maximize reuse of proven hardware system designs;
3. Transformation of the life cycle sustainment strategies for software intensive systems through proven technology insertion and software product upgrade techniques;
4. Dramatically lower development risk through transparency of system designs, continuous design disclosure; and
5. Strategic use of data rights to ensure a level competitive playing field and access to alternative solutions and sources, across the life cycle.

IV. CONCLUSIONS

Based on this approach and architecture, an effective of disparate and heterogeneous devices can easily interoperate for an effective exchange of data. The engagement of the concepts of SOA with web services did not only reduce the amount of deployed code, but it also reduced the management, maintenance. Also, new services and applications can be created quickly and easily used with a combination of new and old services. Data can easily be mined and pulled for analysis due to the open nature of the architecture engaged and all the open principles with SOA. The approaches adopted based on open system architecture renders a vendor-independent, non-proprietary, and computer system or device design based on all kind of standards integration to enable system interoperability. It allows all to create add-on products that increase a system's flexibility, functionality, interpretability, potential use, and useful life. And enables the system users to fine tune or customize system's capabilities to suit their requirements.

REFERENCES

[1] McAfee, A., & Brynjolfsson, E. (2012). Big data: the management revolution. Harvard business review, (90), 60-6.

[2] Boyd, D., & Crawford, K. (2012). Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. Information, communication & society, 15(5), 662-679.

[3] Lessig, L. (1999) Code: and Other Laws of Cyberspace, Basic Books, New York, NY

[4] Pariser, E. (2011) The Filter Bubble: What the Internet is Hiding from You, Penguin Press, New York.

[5] Divyakant Agrawal, Sudipto Das, and Amr El Abbadi. 2011. Big data and cloud computing: current state and future opportunities. In Proceedings of the 14th International Conference on Extending Database Technology (EDBT/ICDT '11), Anastasia Ailamaki, Sihem Amer-Yahia, Jignesh Pate, Tore Risch, Pierre Senellart, and Julia Stoyanovich (Eds.). ACM, New York, NY, USA, 530-533. DOI=10.1145/1951365.1951432 http://doi.acm.org/10.1145/1951365.1951432

[6] Sagiroglu, S.; Sinanc, D., "Big data: A review," Collaboration Technologies and Systems (CTS), 2013 International Conference on , vol., no., pp.42,47, 20-24 May 2013 doi: 10.1109/CTS.2013.6567202

[7] Zhang, L. J. L. (2012). Editorial: Big services era: Global trends of cloud computing and big data. IEEE Transactions on Services Computing, (4), 0467-468.

[8] Demchenko, Y., de Laat, C., & Membrey, P. (2014, May). Defining architecture components of the Big Data Ecosystem. In Collaboration Technologies and Systems (CTS), 2014 International Conference on (pp. 104-112). IEEE.

[9] Big Data defintion, Gartner, Inc. [Online]. Available: http://www.gartner.com/it-glossary/big-data/

[10] Peltz, C. (2003). Web services orchestration and choreography. Computer, 36(10), 46-52.

[11] Kester, Q. A., Gyankumah, G. N., & Kayode, A. I. (2012). Using Web Services Standards for Dealing with Complexities of Multiple Incompatible Applications. International Journal of Information Technology and Computer Science (IJITCS), 4(11), 34.

[12] Kester, Q. A., & Quaynor, N. N. (2013). Integration of Independent Applications and EAI Systems using Service Oriented Enterprise Bus and Open System Application Development Standards. International Journal of Information Technology and Computer Science (IJITCS), 5(7), 1.

[13] Shaping Technology, Shaping Business. Big Data: What on earth is it?. http://blog.au.fujitsu.com/?p=441.

[14] Open systems architecture (osa). http://www.navy.mil/navydata/fact_display.asp?cid=2100&tid=450&ct=2.

