Cloud Business Intelligence – Is What Business Need Today

Yuvraj Singh Gurjar, Vijay Singh Rathore

Abstract - The present economic crisis experienced by all the states of the world orients more and more the information technology industry towards efficiency. Organizations are striving to become intelligent and achieve competition advantages through the use of Business Intelligence (BI) solutions. One of the instruments that can bring about the technology requirements of evolving BI solutions is Cloud Computing. The present paper identifies the key factors responsible for evolution of New Business Intelligence on the Cloud, the various models available to port BI solution on Cloud, the primary drivers for Cloud BI, the impact of implementing Cloud BI as well as issues around it.

Index Terms—Business Intelligence, Cloud, Cloud BI, BI in the Cloud.

I. INTRODUCTION

The past decade has seen rapid evolution of the business landscape. Business environment has totally changed today. The nature and the structure of the current dynamic world cause that nowadays, in times of uncertainty, risks and incomplete information, the crisis becomes a feature of modern business, not the state of emergency. There is more discontinuity observer today then continuity. In this changing world business organizations are increasingly realizing the need for a more scalable and flexible information technology architecture. The ever mounting burden of regulation and compliance is further amplifying business expectations from IT while at the same time tightening the noose [1].

Each organization tends at becoming an intelligent organization and at gaining competition advantage on the market by the use of new and innovative Business Intelligence (BI) solutions. Today, business intelligence (BI) has been under mounting pressure to evolve as an all pervasive information and analytics agent. But the reality is that the failure rate of BI projects is very high and the organizations most often are not able to realize the full potential of a BI project. This failed BI implementation could stem from a number of reasons; including unclear business requirements, multiple and diverse source data systems, long time-to-market or the proprietary technology standards that some off-the-shelf BI solutions demand. The problem gets compounded when we factor in investments that the organizations would have already made on hardware, software and manpower where the cost of failure is too high [1]. On the other side of the coin, in the wake of the present economic crisis and the pursuant business environment, IT has concretized its strategic relationship with business with the reintroduction of grid technology in the form of cloud computing. The cloud model enabled by SOA provides flexibility and scalability (infinite in certain cases) using external computing and processing power in the form of real time e-services. The primary benefits driven by this model are business agility with lower costs, enabling organizations to respond quickly and effectively to the ever changing business environment [1].

Taking into view the prospects provided by Cloud Computing, large investments in traditional BI solutions are often unpractical and unattractive, while popular solutions based on Cloud Computing, called Cloud BI or BI services on demand are increasingly popular. Cloud BI solution has special interest for organizations that desire to improve agility while at the same time reducing IT costs and exploiting the benefits of Cloud Computing.

II. DEFINITIONS

A. Cloud Computing

The National Institute of Standards and Technology (NIST) defines cloud computing in a specific manner, by this we can understand the cloud computing in a better way, that: “Cloud computing is a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction.”[3] Gartner defines, “cloud computing is a style of computing where massively scalable IT-enabled capabilities are delivered ‘as a service’ to external customers using Internet technologies.” [4] According to Quoting Chan, the concept may be defined starting from its name (cloud) as “common, location-independent, online, utility that is available on demand”. This approach emphasizes the fact that any shared resource is statistically multiplied on several applications and clients. Thus, no matter where the client is geographically located, he can access the information in the “cloud”. The “on demand” feature means resources have to be dynamically allocated [5].

Proposed Definition

“Clouds are a large pool of easily usable and accessible virtualized resources (such as hardware, development platforms and/or services). These resources can be dynamically reconfigured to adjust to a variable load (scale), allowing also for an optimum resource utilization. This pool of resources is typically exploited by a pay- per-use model in which guarantees are offered by the Infrastructure Provider by means of customized SLAs.”

B. Business Intelligence

Manuscript Received on January, 2013.
Yuvraj Singh Gurjar, Department of Computer Science, Pacific University, Udaipur -313024, India.
Prof. (Dr.) Vijay Singh Rathore, Director, Shree Karni College, Jaipur and Professor-CSE, Rajasthan College of Engineering For Women, Jaipur -302021, India.

Retrieval Number: F0431021613/2013/BEIESP
“Business intelligence (BI) is the ability of an organization to collect, maintains, and organizes knowledge. This produces large amounts of information that can help develop new opportunities. Identifying these opportunities, and implementing an effective strategy, can provide a competitive market advantage and long-term stability” — [6] “Business intelligence (BI) is an umbrella term that includes the applications, infrastructure and tools, and best practices that enable access to and analysis of information to improve and optimize decisions and performance – [7].

Proposed Definition
“Business Intelligence is a set of methodologies, processes, architectures, and technologies that transform raw data into meaningful and useful information used to enable more effective strategic, tactical, and operational insights and decision-making.” When using this definition, business intelligence also includes technologies such as data integration, data quality, data warehousing, master data management, text and content analytics, and many others that the market sometimes lumps into the Information Management segment.

C. Cloud Business Intelligence
Cloud BI is a revolutionary concept of delivering business intelligence capabilities “as service” using cloud based architecture that comes at a lower cost yet faster deployment & flexibility.

Software as a Service (SaaS) BI is also being used by many small and medium sized enterprises who seek to speed up their businesses with BI and analytics tools.

III. CLOUD, BI AND CLOUD-BI IN TOP TECHNOLOGY P Priorities

A. Cloud and Business Intelligence
Looking at trends shown in Gartner’s list of CIOs technology priority (Compiled on the basis of Annual surveys by Gartner Inc.:s), we can get an idea where Cloud Solutions and Business Intelligence stands in technology priorities for current CIOs [8]. Fig -1: depicts movements of Cloud and Business Intelligence in priorities since 2007 to 2012.

Business Intelligence topped the list from 2007 to 2009, ranked 5th in 2010 and again took top rank in year 2011-12. Cloud Computing hold 1st rank in 2010 and remain in top 5 during last 5 years.

B. Cloud Business Intelligence
According to a recent Market Study of 859 respondents [9], there is a current strong investment in cloud based BI and growing interest in tapping into the cloud’s benefits; while Gartner survey throws the fact that almost one-third of the BI platform users surveyed (27 percent, to be exact) are using or planning to use the cloud / SaaS model to expand their business intelligence capabilities in the next 12 months [10].

These statistics strongly suggest that cloud based BI implementation is on an upswing among organizations.

IV. REQUIREMENT ANALYSIS

A. Need of Mastering "Acquisition-to-Action" Cycle
In the present economy, organizational competitiveness is defined by how quickly companies can synthesize the many sources of information coming their way. To achieve this, they need to be able to master what we call the "acquisition-to-action" cycle. In other words, "How fast can data be captured, stored, queried, analyzed, shared and acted upon?" Traditional BI solutions, massive data warehouses / data marts and databases that were originally designed to crank out pre-configured reports, like sales and financials history, are simply not agile enough to handle today's urgent analytic needs, especially in the persistently expanding domain of machine-generated data [2]. More complex data analysis in the form of ad-hoc analysis (to figure out what to do now) and predictive analysis (to understand what to do next) is the requirement of today’s changing business requirements. Also, organizations are under incredible pressure to do all such analysis at ultra-high tech speed and there is no space for any lag in any term.

B. Arrival of New Data Source
When we talk about information analysis, it's important to think about where the data actually resides. While the scope of traditional BI is limited to structured data that can be stuffed into columns and rows on a data warehouse, the fact is that over 90% of today's data is unstructured in the form of images, MP3 files, videos and social media snippets. Also, much of the data that organizations need to look at is not necessarily "owned" by them - it exists within various social computing services like Twitter and Facebook, it's hidden within Web logs, sensor output, and call detail records. Forward-looking businesses are desperate to tame and gain competitive advantage from unstructured business data floating around within and beyond their enterprise. Finding exactly what you need within such an enormous stream can be like finding a needle in a haystack. But imagine if you could define a set of queries and get the summary information needed in a much smaller, more digestible form. This requires technology that's able to use knowledge about the data itself to intelligently isolate the relevant information and make queries more efficient [2]. To analyze such a big new source of data companies need resources with unlimited processing power, which is
not possible for every organization, thus limiting the scope to analyze new data sources for many.

C. Evolution of New Business Intelligence

In recent times, business intelligence (BI) has been under mounting pressure to evolve as an all pervasive information and analytics agent. Through business intelligence it is possible to improve the decision making process in virtually any department, organization or industry. More and more businesses are turning to analytic applications to provide critical business insights. Whether focused on achieving higher ROI, better understanding of the competitive landscape, improving product and service quality, BI is one of the few technologies that can equip organizations to more effectively prepare for tomorrow today [1]. Organizations have seen how business intelligence has changed over this time, how the tools have evolved offering more functionality to the analysts, and at the same time, providing solutions for more users. Today's IT professionals need to help their organizations capture, track, analyze and share more information than ever before. From mass quantities of transactional data, Web data, and huge and growing volumes of “machine-generated” information, such as sensor and log data, volumes are expanding into the terabyte (and even the petabyte) range. At the same time, the way end users consume information is rapidly changing. Information requirements have grown exponentially: while only a few gigabytes of data were needed some years ago, now data warehouses are populated with terabytes of data and rapidly moving to the petabyte range.

With the advent of new technologies new channels are arriving with lots of data to analyze. In today’s environment, lots of social computing technologies like Facebook, Twitter and LinkedIn, are spreading like wildfire, and companies are starting to look at how to harness social networks, blogs, wikis and more to share business intelligence and collaborate more effectively. As the data center strains under the need for more storage and faster performance (all while keeping costs in check,) cloud computing, open source technologies and other emerging approaches are presenting compelling new ways to manage data and consume IT services [2].

D. Need of Collaborative BI solutions

The value of business intelligence lies in its ability to shape and enhance decision-making throughout the enterprise. Yet all too often, there's not enough context associated with the analyses generated by information management experts, leaving business end users unable to make sense of it all, or at least make sense of it quickly enough to take action. Today's social computing technologies offer a great opportunity for intelligence to be digested in a more collaborative atmosphere. Imagine combining analytics (a quarterly sales report tracking online purchases, for example) with capabilities like search, bookmarking, tagging, commenting and rating capabilities. Now imagine accessing the intelligence via a Web-based portal where any number of enterprise stakeholders can look at the report collaboratively and engage in a conversation that adds context to the content. This socially powered approach to BI can significantly speed decision-making. More important, it enables a far richer understanding of corporate data, which enables better decision-making as well. Collaborative BI solutions that integrate social technologies directly into the analytic environment are increasingly available, and smart companies would do well to take a look at how these can enrich their BI efforts. As organizations continue to be bombarded by data, old BI strategies are increasingly giving way to more innovative approaches. The businesses that thrive will be those that succeed at adopting the best new ideas and technologies.

E. Moving BI in the Cloud

Considering the present trends in adopting Cloud Solutions, data center is not going to disappear anytime soon and cloud computing is certainly democratizing information access. The strengths of the cloud model e.g., accelerated speed-to-market, reduced TCO, scalability, etc., have led many BI vendors to introduce cloud services as a clear and distinctive extension to the on-premise and on-demand BI applications [1]. Companies like Amazon and Google offer unlimited processing power and storage thus allowing any business to cater to its increasing information stack while keeping the IT related costs under control.

For example, smaller companies that previously couldn’t afford to build huge server farms to process mass amounts of data can turn to providers like Amazon and Google to support large-scale analysis efforts. In addition, a number of innovative SaaS and “cloud-friendly” BI and analytic solutions are cropping up, which means that organizations can take advantage of the cloud to not only store their data, but also crunch it. There are, of course, some key considerations. Security and data privacy get the most press, but uptime, performance and openness/portability are also important. Depending on the organization’s specific requirements, there's more than one flavor of cloud, ranging from public (affordable and highly scalable), private (offering greater security and control) and hybrid (combining aspects of both). The best approach will ultimately depend on what's most important to the target organization. And, when it comes to BI, cloud solutions will only be as good as the performance they can deliver.

V. BASIC ARCHITECTURE

The Basic Architecture - The basic architecture needed to run business intelligence solution in the cloud is depicted in Figure 2 [12].

![Fig 2: BI on the Cloud: Architecture](image)

The lower layers are formed by hardware and software systems. These are the minimum elements that have to be offered by the cloud.
computing provider. Hardware refers to processing, storage, and networks, while software refers to the operating systems and drivers required to handle the hardware.

The Data integration box refers to the tools needed to perform the ETL and data cleansing processes. The database box refers to the relational or multidimensional database systems that administer the information. It is important to note that there are new devices called "data warehouse appliances", which integrate hardware, software and databases elements in just one box. However, they should be considered as an integrated part of the architecture.

Data warehousing tools are the set of applications that allow the creation and maintenance of the data warehouse. BI tools are the set of front-end applications that enable the final users to access and analyze the data.

Finally, since all the architecture is going to be accessed through the Internet, there is no need for thick clients or preinstalled applications, because all the content and configuration can be reached through traditional internet browsers.

VI. DEPLOYMENT MODELS

The following models are available while deploying BI components on Cloud -

1) Public cloud-based IaaS for a BI system: This option involves subscribing to an IaaS model for data storage and processing power. Companies can then buy and deploy their own ETL, DBMS and BI tools on top of this. Vendors serving IaaS Cloud solution are Amazon, Rackspace or GoGrid pro.

2) Public / Hybrid Cloud based PaaS for BI and DW: This option involves deploying the BI/DW system on a public or hybrid cloud to host own cloud-based BI system. This option is best suited to implement BI systems for SMBs, Custom Analytic applications, Enterprise BI systems, Data Mining, Prototyping etc. Vendors in this area are—AsterData MPP on Amazon EC2, IBM Cognos Express on Amazon EC2, Teradata Express on Amazon EC2 and RightScale/ Talend/Vertica/Jaspersoft on Amazon EC2.

3) Analytics on the Cloud: These products are public cloud based solutions. Vendors provide pre-build solutions for analytics. Vendors include — SAP Business Objects, Microsoft BI Tools, LucidEra, IBM Cognos, Cloud9 Analytics etc [14].

4) Public or Hybrid cloud based SaaS BI – This option provides cloud based reporting system on operational data either from SaaS based transaction processing system or from internal transaction processing system. Vendors available are – SAP Business Objects, Microsoft Azure etc.

5) Private cloud-based BI system running internally - The largest private cloud-based BI system is IBM’s internal Blue Insight which is based on IBM System Z and IBM Cognos 8 BI. IBM has also launched the Smart Analytics Cloud, a private cloud offering for large enterprises based on the same technologies [14].

VII. DRIVERS FOR CLOUD BI

There are several operational and financial factors that work in favor of Cloud Business Intelligence (BI), the key being:

1) Time Saving with speed of implementation and deployment: Immediate availability of environment without any dependence on the long periods associated with infrastructure procurement, application deployment, etc. drastically reduces the BI implementation time window [11][14].

2) Lower Total Cost of Ownership: Convert some part of capital expenditure to operational expenditure, cost-effective pricing models, pay per use model, etc. On a long term, these solutions help organization in reducing operational costs, IT support expenses and much more. [14]

3) Elasticity: Leverage the massive computing power available on the Web, scale up and scale down based on changing requirements. [14]

4) On-demand Availability: Support mobile and remote users, Browser-based access to control everything from the cloud platform to database management, from the data warehouse layer to the analytics platform. [14]

5) Expertise support: Organizations choosing third-party Cloud BI solutions getting expertise backing in the form of cloud and tech-savvy professionals who can help their internal IT and business team to deal with managing BI and analytic systems. [9]

6) Focus on Core Strength: Outsource running of BI apps to professionals and focus on core capabilities.

VIII. IMPACT OF TRANSFORMING BI IN THE CLOUD

Cloud computing is transforming the economics of BI and opens up the opportunity for enterprises to compete using the insight that BI provides. Cloud-based solution will impact BI by:

Fig 3: Benefits provided by Cloud BI Solutions

1) Easier evaluation of Technology: Cloud enables software companies to make new technology available to evaluators on a self-service basis, avoiding the need to download and set up free software downloads or acquire hardware’s fitting to the technology. [13]

2) Increased short-term ad-hoc analysis: Where short term needs (weeks or months) for BI is required, cloud services are ideal. A
data mart can be created in a few hours or days, used for the necessary period, and then cancel the cloud cluster, leaving behind no redundant hardware or software licenses. The cloud makes short term projects very economical. [13][15]

3) Increased flexibility: Due to the avoidance of long term financial commitments, individual business units will have the flexibility to fund more data mart projects. This is ideal for proof of concept, and ad-hoc analytic data projects on-demand. This agility enables isolated business units to respond to BI needs faster than their competitors and increase the quality of their strategy setting and execution [13][15].

4) Drive data warehousing in MB markets: Medium-size businesses often have very large volumes of data for analysis, yet only a few IT resource at their disposal to analyze tons of terabytes of historical data to fine tune market strategies. Cloud-based solution can enable such businesses to warehouse and analyze terabytes of data in spite of these resource constraints [13].

5) Drive the analytic SaaS market: Companies that collect economic, market, advertising, scientific, and other data and then offer customers the ability to analyze it online will be able to bring their solutions to market with much less risk and cost by utilizing cloud infrastructures during the early stages of growth [13][16].

6) “Scale-out” shared-nothing architecture: To handle changing analytic workloads as elastically as the cloud. Auto-scaling of Virtual Machine (VM) can be used to provide necessary compute power required during heavy workload and an efficient algorithm needs to be worked out in order to auto-scale in VM’s when not required.[13]

7) Aggressive data storage: Cloud provides an appropriate infrastructure for storing large amount of data at low cost. No further additional overheads are required to store data on cloud thus helping achieve manpower savings for operations like data backup and server maintenance. For example in case of Windows Azure, Table Storage is designed to be massively scalable and a typical Azure Table can contain billions of records amassing to Terabytes of data. Blob Storage provides a means to store unstructured data much in the same way that would store a bunch of images on the File System of a server. Blobs can be mounted as XDrives on the Virtual Machine instance where a particular service is running and accessed exactly like a file system would. [1][13]

8) Automatic replication and failover: This will provide high availability in the cloud. In case of Windows Azure, data is stored on 3 nodes to enhance both access speeds and reduce data redundancy.

According to recent research [17], Cloud based BI solutions were viewed as beneficial with 78% saying they would see value. Several participants took the time to point out that the demand for data analysis in a BI solution is very uneven, which makes it a particularly good fit for a cloud solution. Figure-3 lists the various benefits by Cloud BI as accepted by various participants of the research.

IX. ISSUES WITH CLOUD BI

While business intelligence can benefit from cloud computing, it is not a silver bullet, and there are several potential challenges, such as:

1) Moving data to the cloud – Large data sets in silos sitting on premises need to get to the cloud before they can be completely used. Moving large data volumes over the Internet can be a challenge for most companies. Upload speeds are often slower than download speeds and network pipes are often divided up into channels to accommodate voice and data. Initial data loads often require the sending of media and coordination with service providers. This is often recommended by cloud providers like Amazon. Although this introduces latencies, it is often acceptable for most BI options. With load windows shrinking in many environments, this can pose a significant barrier. While it is true that we could load data more frequently, many data integration architectures aren’t currently designed to support this change. [15][18]

2) Data Security – For some organizations, the concerns over security may be a barrier that is impossible to overcome today. The majority of data is core and proprietary to enterprises. Secure storage of this data is essential and, in quite a few cases, it might be mandatory to keep this data on premise. There are several options to secure data in the cloud including during storage and at rest, which include standard security measures like encryption keys, SSL and certificates. Compliance and regulatory reasons also require data to be stored securely. However in many cases, the Cloud vendors provide a more secure environment than what exists at customer sites.[14][18]

3) Speed of data access: Since storage resources are separate from server resources, there is likely to be a significant latency in accessing large amounts of data. Large scale data warehouses require high speed backplanes and dedicated data storage nodes. Until Cloud providers can provide this level of service or data warehouses can exist entirely in memory, there will be limits to the size and performance of data warehouses in the Cloud. There will also be significant latency if BI applications exist in the Cloud but the data exists at a site client, especially when processing and returning large amounts of data. [18]

4) BI components as a service – So far, only a limited set of services are available from established BI vendors. This includes some reporting capabilities and ability to do visualizations. Most established vendors are yet to introduce complete product features over the cloud. [15]

5) Integration with on premise data – It is challenging to integrate on-premise data with cloud components, as it continues to exist in silos and requires access to data behind the firewall. [14][15]

6) Lack of control: Tough to get Service Level Agreements (SLAs) from cloud providers. Data control and data ownership, reliability of service challenges are some of the main reasons for client concern. To mitigate this, organizations should already have in place thorough IT governance.
and service delivery standards and models. [14]
7) Vendor Maturity: Too many cloud BI vendors, hosting providers with varying offerings, etc. makes it confusing to choose the right vendor based on required needs and vendor capabilities. [14]
8) Reliability of service: The services offered by different vendors vary a lot and back support provided by vendors is limited. Amazon Web Services doesn’t have a customer service phone number. All contact is via email with a 48 hour service level agreement. This can be especially frustrating during service disruptions. On October 7, 2007, Amazon’s EC2 went down and an undisclosed number of customers lost their servers and their data. On February 2, 2008, both Amazon’s EC2 and S3 were inaccessible for more than 3 hours. [18]
9) Limited ability to scale-up: Some BI software, such as SMP databases, in-memory analytics and OLAP scale better on a single server versus across servers. Currently, the ability to provision a server beyond 4 CPU’s (8 cores) and 16GB’s is limited. As more providers support 64-bit platforms and virtualization becomes more advanced, this limitation will begin to go away.[18]
10) Performance: Limits to the size and performance of data warehouses in the Cloud, significant latency if BI applications exist in the Cloud but the data exists at a client site, especially when processing and returning large amounts of data. [14]
11) Pricing: It is difficult to budget for computing resources when those costs are variable in nature. It is especially difficult when pricing is based on multiple components such as network traffic, storage size and IP address requests as well as different pricing tiers for each component. Also, different vendors base their pricing on different components. For example, GoGrid charges, in part, based on memory (each Gigabyte (GB) of RAM utilized is equivalent to 1 RAM hour) as well as network traffic. [14][18]

X. CONCLUSION
Cloud is a big part of future Business Intelligence and offers several advantages in terms of cost benefits, flexibility of implementation, availability and speed of implementation. BI on the Cloud offers huge possibilities for removing barriers to decision making by integrating high volume and mission critical business processes. Therefore, a Cloud BI solution may be a feasible answer to the challenges of the economic crisis. By such a solution, the economic organization – small, middle-sized or large – may use market opportunities that under normal conditions (other than adopting BI or Cloud BI) would not be accessible. Irrespective of the age of a BI landscape the cloud model can drive increased BI adoption, improved end-user experience, better access to analytics and reduced IT dependence.

REFERENCES

AUTHOR PROFILE
Yuvraj Singh Gurjar, MCA, MCTP – SQL Server, MCTS – Business Intelligence, Oracle Certified Associate. He received his MCA degree from University of Rajasthan in the year of 2002. He started his career as MIS Programmer with UIDP (An Asian Development Bank funded project in Rajasthan (India)). In 2003, he joined renowned bank Punjab National Bank as Officer-IT. In 2008 he shifted to Union Bank of India as Sr. Manager (Software). Later on, in year 2011 he founded KURIOS - A firm working in the area of Business Intelligence and Performance Management Solutions. Currently he is working as Chief Technical Consultant with KURIOS. He has more than 10 years of experience in IT field with specialization in MIS and BI applications. He is also a student of Ph.D. in the Department of Computer Science in the Faculty of Computer Science of Pacific University, Udaipur, India.

Prof. (Dr.) Vijay Singh Rathore, MCA, M.Tech. (C.S.) PhD (Computer Science), MBA, ICADA-USA, ICAD-USA, Diploma French & German. Presently, He is Director in Shree Karni College, & Shree Karni Institute of Science, Management & Technology (KISMAT), and Professor (CSE) in Rajasthan College of Engineering For Women, Jaipur (India). He has 12 years of experience in Technical Education, having done the First PhD in Computer Science after MCA from University of Rajasthan, Jaipur, written 12 Books on various subjects of Computer Science, published more than 50 research papers at national & international level, supervised 06 PhD candidates, and currently supervising 08 PhD research scholars in Computer Science. He is active member of various committees of UGC, RTU, UOR, UOK, JNU, SOVU, VMOU, NIMSU, AMITYU, MIRPU, IJISU, JRSU, JNRVU, PU etc. He is reviewer in some reputed International Journals. His research area is networking & security with special emphasis on Cloud Computing.

Published By: Blue Eyes Intelligence Engineering & Sciences Publication
Retrieval Number: F0431021613/2013@BEIESP