BUSINESS INTELLIGENCE IN THE CLOUD

Shimaa Ouf
Faculty of Computers & Information
Helwan University, Egypt.
shimaaouf@yahoo.com

Mona Nasr
Faculty of Computers & Information
Helwan University, Egypt.
m.nasr@helwan.edu.eg

Abstract: **Business Intelligence (BI) deals with integrated approaches to management support. Currently, there are constraints to BI adoption and a new era of analytic data management for business intelligence these constraints are the integrated infrastructures that are subject to BI have become complex, costly, and inflexible, the effort required consolidating and cleansing enterprise data and Performance impact on existing infrastructure / inadequate IT infrastructure. So, in this paper Cloud computing will be used as a possible remedy for these issues. We will represent a new environment atmosphere for the business intelligence to make the ability to shorten BI implementation windows, reduced cost for BI programs compared with traditional on-premise BI software, Ability to add environments for testing, proof-of-concepts and upgrades, offer users the potential for faster deployments and increased flexibility. Also, Cloud computing enables organizations to analyze terabytes of data faster and more economically than ever before. Business intelligence (BI) in the cloud can be like a big puzzle. Users can jump in and put together small pieces of the puzzle but until the whole thing is complete the user will lack an overall view of the big picture. In this paper reading each section will fill in a piece of the puzzle.**

Keywords: Cloud Computing, business Intelligence, platform as a services.

I. INTRODUCTION

In the wake of the economic slowdown, organizations are increasingly looking for ways to do more with the same resources; articulate differently - to make every penny, input and contribution count. In such situations, technologies like Cloud Computing and Business Intelligence (BI) are becoming increasingly important in gaining and maintaining a competitive edge. [5]These technologies when combined enable a variety of new analytic data management projects and business possibilities. Cloud computing will change the economics of BI by making available the hardware, networking, security and software needed to create data marts and data warehouses on demand with a pay-as-user-go approach to usage and licensing. More and more businesses are turning to analytic applications to provide critical business insights. [8]Whether focused on achieving higher Return On Investment (ROI), better understanding of the competitive landscape, improving product and service quality, Business Intelligence is one of the few technologies that can equip organizations to more effectively prepare for tomorrow today. It's no wonder the BI platform is expected to grow by 7.9% through 2012. [11] Cloud computing is characterized by ability to consume resources as required in an elastic manner and scaling the consumption arbitrarily as required. The advent of infrastructure as a service implies that Computational power is available on demand and on a pay-as-user-go basis with similar characteristics applying to storage of data as well. This enables a layer of services sitting on top of this infrastructure to decouple the delivery aspect of the services from the core business oriented aspects involved in these services. Related to this, is the fact that storage of the underlying data is also decoupled and segregated from the services. [4]

II. BUSINESS INTELLIGENCE

Business intelligence (BI) has been referred to as the process of making better decisions through the use of people, processes, data and related tools and methodologies. The roots of business intelligence are found in relational databases, data warehouses and data marts that help organizing historical information in the hands of business analysts to generate reporting that informs executives and senior departmental managers of strategic and tactical trends and opportunities. [8] In recent years, business intelligence has also come to rely on near real-time operational data found in systems including enterprise resource planning (ERP), customer relationship management (CRM), supply chain, marketing and other databases. “Operational” BI is meant to provision many more functions in the organization with role-specific dashboards and scorecards and is increasingly tied to the topics of performance management and business process management. Inherent to any form of BI is the notion of data quality, consistent and dependable data and the processes involved in its creation and maintenance. [12]Business Intelligence involves intelligent reporting on top of existing data which helps in prompt and actionable decision making. These decision making might involve "geography based investment decision for a multinational company" or even a "buy decision for a product by the consumer". BI has evolved over time but the key components still continue to hold true. It is still necessary to be able to aggregate the factual data from various data sources and doing involved transformations. This data then either needs to be stored in a data mart or warehouse to enable reporting and analysis on top or it could then be further aggregated into metrics which are then reported. Nevertheless the ability to perform BI involves key aspects related to data management and computationally

978-1-61284-486-2/11/$26.00 ©2011 IEEE

expensive analytics or reporting. [11]*Cloud computing is transforming the economics of BI and opens up the opportunity for smaller enterprises to compete using the insights that BI provides. Cloud-based analytics will impact BI by:*

Accelerating BI technology adoption: the cloud becoming the default platform for evaluating new software.

Easier evaluation: the cloud enables software companies to make new technology available to evaluators on a self-services basis, avoiding the need to download and set up free software downloads. [16]

Increased short-term ad-hoc analysis: avoiding data marts spawned as a result of new business conditions or events. Where short term needs [weeks or months] for BI is required, cloud services are ideal. A data mart can create in a few hours or days, used for the necessary period, and then the cloud cluster cancelled, leaving behind no redundant hardware or software licenses. The cloud makes short term projects very economical.[11]

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. [17]

Growth considerations

As data volumes grow, for analytic cloud projects to succeed they will require a database architecture that is designed to function efficiently in elastic, hosted computing environments like the cloud. At a minimum, such databases must include the following architectural features:

- "Scale-out" shared nothing architecture to handle changing analytic workloads as elastically as the cloud.
- Aggressive data compression - to keep storage costs low.
- Automatic grid replication and failover- to provide high availability in the cloud.[11]

III. THE CLOUD COMPUTING TSUNAMI (EFFICIENCY AND COST CONTROL IN THE IT INDUSTRY)

Cloud computing has become one of the hottest buzzwords in the IT area. Many companies and institutions are rushing to define clouds and provide cloud solutions in various ways. [1]Cloud computing is Internet-based computing, whereby shared resources, software, and information are provided to computers and other devices on demand, like the electricity grid. Due to the fact it involves the existence of data centers that are able to provide services; the cloud can be seen as a unique access points for all requests coming from the world wide spread clients. [19]

Fig. 1 Cloud computing [1]

The cloud computing wave is the most dramatic change we have observed in the computing industry since the wave of the Internet. Cloud computing will significantly change data centers and IT organizations as well as the infrastructure and software vendors' business models. In fact, the cloud computing wave is not just a wave – it is more like a tsunami. What is causing this cloud tsunami? [21]We start with listing 4 of the Gartner top 10 IT predictions for the next three to five years for cloud computing, software as service (SaaS), data center power/cooling efficiency and open source software. All of these predictions indicate that data center efficiency and cost containment will transform *the IT industry over the recent & next years. Key predictions for the data center for the next 5 years:* [20]

- In 2011, early technology adopters will forgo capital expenditures and instead purchase 40 percent of their IT infrastructure as a service.
- By 2012, at least one-third of business application software spending will be as service subscription instead of as product license.
- In 2009, more than one-third of IT organizations will have one or more environmental criteria in their top six buying criteria for IT-related goods. Initially, the motivation will come from the wish to contain costs. Enterprise data centers are struggling to keep pace with the increasing power requirements of their infrastructures.
- By 2012, 80 percent of all commercial software will include elements of open source technology.

These predictions clearly indicate that data center efficiency and cost are forming massive waves for increasing data center efficiency with virtualization, better server utilization, cost reduction with more efficient power/cooling and the leveraging of open source software. [18]

The Cloud Impact on Data Centers:

A mix of private and public clouds will become the norm.

Many organizations and lines of business will bypass IT to secure cloud-based infrastructure and SaaS applications.

- Virtualization and cloud-based infrastructure will become the norm.
- Power and cooling efficiency and green data centers will become critical and the norm.
- A new breed of cloud computing skills will become common in data center operations.
- Private cloud technology, such as cloud storage, will find its way to the IT organizations.
- Service oriented architectures (SOAs) will drive the IT infrastructure and application architecture. [11]

The Cloud Impact on IT Organizations:

- We will see a transformation from programming to service integration and customization.
- With the cloud and SaaS usage-based pricing, IT budgets will transform from CAPEX to more OPEX, opening the door for immediate IT investments.[16]

The Cloud Impact on Cloud Computing Vendors:

- Significant market growth and momentum will fuel hyper growth.
- Cloud infrastructure utilization and efficiency will become critical to success.
- Power and cooling costs will become enormously important factors to profitability.
- Cloud infrastructure and SaaS vendors will become the new giants of the industry where the IT operations shop for infrastructure and SaaS applications.
- Merger and acquisition frenzy will become the norm for hyper growth.[17,18,19]

The Cloud Impact on SaaS Vendors:

- There will be hyper growth in the number of SaaS applications and vendors.
- Venture spending will grow significantly.
- New requirements and standards for APIs, reporting, security and service-level agreements (SLAs) will emerge.
- SaaS vendors will become the main source of applications.[1]

The Cloud Impact on Infrastructure Vendors:

- Server, storage and networking customer influence will decrease.
- Server, storage and networking vendors will be selling to cloud vendors.
- Infrastructure vendors will be fighting for mind-share with both cloud and SaaS vendors.
- Infrastructure vendors will lose contact with many enterprise customers as they flock to cloud infrastructure and SaaS.
- Merger and acquisition frenzy will become the norm for survival.
- Infrastructure vendors will experience a dramatic change of business model. [15]

The Cloud Impact on Application Software Vendors:

- Application software vendors will have to adopt the SaaS model to survive.
- They will lose business to SaaS companies.
- Software licensing will dramatically change.
- Merger and acquisition frenzy will become the norm for survival.[7,8]

IV.THE BENEFITS OF CLOUD COMPUTING FOR BUSINESS INTELLIGENCE

The convergence of the Cloud, Open Source, and specialized analytic databases has brought a new world of possibilities to businesses. No longer are dedicated hardware and software required to meet ad hoc or unpredictable BI requirements. The combination of these technologies running in the cloud will enable organizations to: [2]

- Create analytical sandboxes on demand to optimize response to sudden market changes.
- Offload expensive processing from overburdened data warehouses and avoid costly upgrades.
- Prototype new proof-of-concept BI applications without purchasing and implementing new data center infrastructure.
- Run full business intelligence and data warehousing for existing cloud applications entirely in the cloud.[12]

V. CLOUD BUSINESS INTELLIGENCE: STORMY FUTURE

Much has been said and written about cloud computing. The cost and scale advantages it can offer public and private computing infrastructures are astounding - providing genuine hope for more efficient and far-reaching, next-generation computing. In fact, cloud BI holds great potential to substantially disrupt the business intelligence market because of its perfect storm of low-cost, scalability, and flexibility. With business intelligence software running in the cloud, it is still possible to make comprehensive integration with back-end systems – both within the user company and in the cloud. [13]Cloud computing environments mean that organizations no longer need to expend capital upfront for hardware and software purchases. Nor do they need to suffer through prolonged in-house implementations. In these two areas, cloud computing and SaaS models both share similar benefits. [1]Cloud BI represents a way for reporting and analysis solutions to be developed, installed, and consumed more easily due to its lower cost and easier deployment. Ideally, a cloud-based business intelligence platform makes use of infrastructure-as-a-service (IaaS), complements and extends today's platform-as-a-service (PAAS), utilizes an on-demand, virtualized, elastic software and hardware environment, and delivers application-level functionality as a service (commonly referred to as software-as-a-service). [2]

Cloud-based BI platform is used to solve one of three primary customer needs:

1. As a **horizontal BI tool** to deliver standalone, internally facing reporting and analysis applications probably using a traditional relational database (or data mart) as the primary source data system.
2. As an application framework or pre-built reporting and analysis template for systems integrators to use for

assembling customer-specific solutions more quickly. These solutions are probably function or domain specific and contain reusable components and application logic (but are assembled uniquely for each customer).[3]

3. As a development platform that enables embeddable, externally-facing applications that solve a function-specific data analysis problem (for example, CRM analytics, financial analytics, or supply chain analytics). In this case, an ISV (or an enterprise IT team with appropriate skills) would probably use the BI platform to deliver reporting and analytics as a well-defined and well-featured layer within its larger application. The result is an analytic application that solves a customer problem with minimal customization and that is ideally delivered using a software-as-a-service architecture on top of a cloud infrastructure.[4]

Also cloud computing represents several benefits to the BI such as:

- *On-Demand:* Immediately available with no infrastructure to deploy.
- *Elastic:* Can scale up or down quickly with changing requirements.
- *Affordable:* No large upfront costs, pay as user go.
- *Flexibility:* to scale computing resources with few barriers.
- Geographic scalability. [7]
- Deploying BI in the cloud can help programs become more flexible, scalable and agile.
- It can be challenging to configure databases and BI tools to run in the cloud.[6]
- *Cheap Processing Power*. The parallelization of cloud computing makes Relational Online Analytical Processing. (ROLAP)-based analytics possible, since queries can be spread across multiple CPUs simultaneously.
- *Elastic Scale*. BI is very sensitive to unpredictable and high peak loads. Users don't have to build for this, and can elastically scale to meet demand when it happens. [7]
- *Massive Multi-Tenancy*. Users run a single instance of their BI platform across 1000s of customers, meaning the marginal cost to provision, service and upgrade each user is extremely low.
- *Service-Orientation*. Given the transient nature of hardware nodes. [5]

Additionally, Cloud computing provides key enablers for BI in these specific areas:

- Computational resources can be consumed for heavy calculations that could be involved in predictive analytics as an example.
- Extremely heavy data loads could be stored for cheaper prices in storage resources in the cloud. [7]
- Reporting and visualization are naturally fit for offering in a software as a service (SaaS) model, this would enable newer Consumption behavior for these specific BI components, also given that web is the most common delivery method for reports. It makes it seamless to add new users as well allow user.
- As more and more applications and data sets move to the cloud, BI services need to adapt to look at Cloud as the data source. [14]

VI. THE PROPOSED MODEL

The future will be very bright for the use of BI in the cloud, both because of the advantages that underpin this new computing paradigm as well as the explosion of digital data that grows each day. *"BI in the Cloud" architecture is only going to be feasible when most of user's source data lives in the cloud already, possibly in something like SQL Server Data Services or Amazon Simple DB or Google BigTable; or possibly in a hosted app. like Salesforce.com.* Cloud BI is the new way to do Business Intelligence instead of implementing expensive and complex software on-site, the BI software runs in the Cloud. It is accessible via any web browser in a so-called software-as-a-service model. There is no need to install software, or to buy any hardware. And when users are computing needs grow, the system will automatically assign more resources. This elastic scale is what makes Cloud BI so powerful users pay for what they use as opposed to always paying to provision for peak load. Starting at the back, the first objection raised to a purely 'BI in the cloud' architecture is that user has got to upload his data to it somehow. Users can use the tools and applications they're familiar with to work from anywhere. With the choice and flexibility of cloud computing, users' businesses can deploy services on-premises, in the cloud, or a blend of both. And, our solutions are all built on a unified productivity platform that's not only cost-effective, but gives user the agility to respond as business needs evolve.

Component of the proposed model (Business Intelligence in the Cloud

The proposed model represents a new environment atmosphere for the business intelligence to make *the ability to shorten BI implementation windows*, *reduced cost for BI programs*, *Ability to add environments for testing, proof-of-concepts and upgrades* .This environment represented as follows:

1- Cloud computing

The cloud provides a virtually unlimited pool of computing power, storage and memory. However, these resources are delivered in discrete modules. Each node consists of "standard" units of processing power, storage space and memory. While the amounts may vary (by service provider, price point, etc.) and they may increase over time, the Cloud's pool of resources is a large grid of interchangeable, industry-standard computing resources. Achieving true scalability requires a database architecture that can fully maximize this pool of resources.

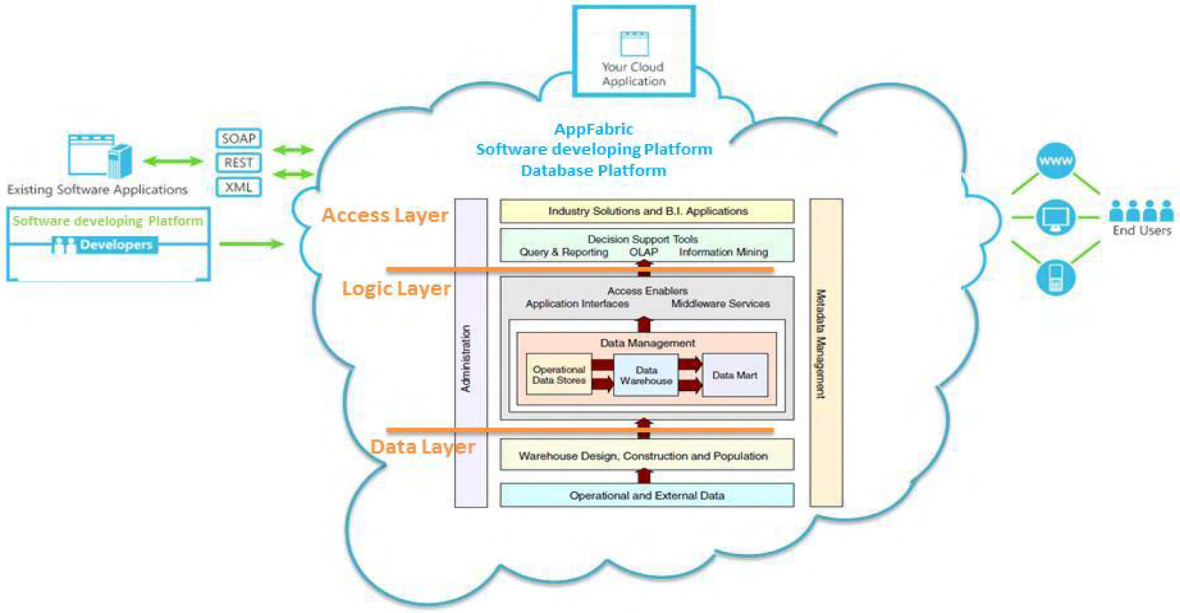

Fig. 2 Business Intelligence Proposed Model

A shared nothing, massively parallel database architecture is particularly designed to take advantage of multiple units of computing resources. In the proposed model we use Microsoft windows azure and we describe *How cloud computing can be affected on the business intelligence?*

2. AppFabric:

Helps connect applications and services in the cloud or on-premise, for example applications running on Windows Azure, Windows Server and a number of other platforms including Java, Ruby, PHP and others.

3. Platform as services:

Platform as services is a cloud services operating system that serves as the development, service hosting and service management environment. It is a flexible cloud–computing offering that lets user focus on solving business problems and addressing customer needs. No need to invest upfront on expensive infrastructure. Pay only for what user use, scale up when he need capacity and pull it back when he don't. Users handle all the patches and maintenance all in a secure environment. Cloud computing as a platform supports multiple languages and integrates with existing on-premises environment. In addition, it supports popular standards, protocols and languages.

3.1 Database Platform:

Business intelligence can be moved to cloud using Platform as services. It is a cloud-based relational database service built on database technologies. It provides a highly available, scalable, multi-tenant database service hosted in the cloud. Platform as a services helps to ease provisioning and deployment of multiple databases. Users do not have to install setup and patch or manage any software. High availability and fault tolerance is built-in and no physical administration is Database Platform of their existing on-premises databases. Database Platform delivers scale to meet the needs of the entire organization and provides IT with flexibility to respond quickly to the evolving needs of the business. Organizations can bring great new experiences and empowerment to their end users on a familiar IT infrastructure that's more manageable and cost effective. Database Platform includes features to help user manage critical data assets company-wide and across diverse systems, helping to ensure integrity of information. Mostly master data is maintained into the permanent staging database and synchronized with delta from extracts and this piece of data is not that huge too. With a proper design and dissecting permanent staging area into two parts, by moving master tables to the cloud and delta records to a temporary staging area, the intermediate need of a staging server can be eliminated. Database Platform is a fit for this, as we just

need to store the staging data and queries are not that complex. And just for storing of this staging data, we do not need an enterprise level database and access to this data is not that frequent too.

2. BI infrastructure(consists of three layers)

The Data Layer: The Data Layer is responsible for storing structured and unstructured data for management support. Regarding structured data, the central component is the data warehouse (DWH). A DWH is commonly defined as a "subject-oriented, integrated, time-variant, and non volatile collection of data in support of management's decision-making process". Many current realizations of DWHs are based on so called Core DWHs .Core DWHs are usually not used as a direct source for analysis systems, but rather distribute data to individual Data Marts. Data Marts keep excerpts of application specific data. More recently; there has been a shift towards DWH infrastructures that are integrated with operational systems. This is usually achieved by the introduction of an Operational Data Store (ODS) that is designed to keep real time data on a transactional level for time critical tasks. ODS/DWH architectures allow to build Closed-loop and Active Data Warehousing solutions. To feed the various data storages, ETL (Extract-Transform-Load) tools are needed. An ETL tool supports the extraction and transformation of data from heterogeneous source systems. The transformation includes filtering out syntactical and semantic errors, harmonizing data from different sources, as well as aggregating and enriching it. For the storage and administration of unstructured data, Content Management Systems (CMS) and Document Management Systems (DMS) are inserted into the data layer.

The Logic Layer: The Logic Layer provides functionality to analyze structured data or unstructured content and supports the distribution of relevant knowledge among different users. The most salient tools in BI environments are reporting, data mining, and OLAP tools: Reporting tools present quantitative data in a report-oriented format that might include numbers, charts, or business graphics. OLAP denotes a concept for interactive and multidimensional analysis of aggregated quantitative business facts. Data mining tools support the identification of hidden patterns in large volumes of structured data based on statistical methods like association analysis, classification, or clustering. Data mining and similar model based tools are also referred to by the term Advanced Analytics.

The Access Layer: The Access Layer allows the user to conveniently use all relevant functions of the Logic Layer in an integrated fashion within the confines of defined user roles and user rights.

VII. CONCLOSION AND FUTURE WORK

The development of business Intelligence field cannot ignore the cloud computing trends. There are many benefits from using the cloud computing for business intelligence. It influences the way business intelligence software projects are managed which it provide a virtually unlimited pool of computing power, storage space and memory for the business intelligence infrastructure, so our proposed model represents a new environment atmosphere for the business intelligence that help in shortening BI implementation windows, reduced ion of cost for BI programs, enabling to add environments for testing, proof-of-concepts and upgrades. Business Intelligence in the cloud has been developed in order to enhance the efficiency and productivity of business intelligence and increase the performance of BI software. In the future we are aiming to develop business intelligence by using web 3.0 technologies. In the future we are aiming to develop business intelligence by using web 3.0 technologies.

REFERENCES

[1] Pocatilu, P., F. Alecu, et al. "Measuring the Efficiency of Cloud Computing for E-learning Systems", Romania /January 2010.

[2] Baars, Henning and Kemper, et al, *"Business Intelligence in the Cloud?"*, (2010). *PACIS 2010 Proceedings.*

[3] Withee, Microsoft Business Intelligence for Dummies, Wiley Publishing, Inc.

[4] *Relan, p.andSibbingz,"Business intelligence for the cloud software stack helps gaming sites quickly adapt to evolving customer needs"*, 2009.

[5] *Dialogic Corporation, "Introduction to cloud computing"*, white paper.

[6] *J.Cios, K., W. Pedrycz, et al, "Data Mining A Knowledge Discovery Approach"*, 2007, Springer.

[7] www.wikipedia.com

[8] www.microsoft.com

[9] http://www.information-management.com

[10] www.thebusinessintelligenceguide.com

[11] http://www.ctoedge.com

[12] www.thebusinessintelligenceguide.com

[13] http://www.slideshare.net

[14] http://technet.microsoft.com

[15] http/www.salesforce.com

[16] http://aws.amazon.com, Amazon Elastic Compute Cloud (Amazon EC2).

[17] http://www.gc-solutions.net, Cloud Computing & How It Impacts e-Learning. June 24th, 2010 by Gupta, M.

[18] http://code.google.com, Google App Engine

[19] http://content.dell.com, Dell and Cloud Computing.

[20] https://www.microsoft.com, Windows Azure Platform.

[21] http://www.ibm.com.