Advances in Computing 2017, 7(3): 74-79
DOI: 10.5923/j.ac.20170703.02

The Cloud Promise (Moving Data to Cloud)

Subash Thota

Data Architect, CRGT, USA

Abstract Business Intelligence (BI) & Data Analytics deals with integrated approaches to management support. Currently, there are constraints to data management for business intelligence these limitations 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 article Cloud computing will be used as a possible remedy as Cloud computing also enables organizations to analyze terabytes of data faster and economically than ever before. When we define data, information and analytics as services, we see traditional methodology, which is mainly time and cost driven, which do not work well. Organizations need to consider the value of service level and quality in addition to the cost and duration of delivered services. CLOUD enables scale, scope and speed economies. Many believe that Cloud will reshape the entire industry as a revolution. In this paper, we aim to pinpoint the challenges and issues of Cloud computing. This study highlights that organizations face various challenges using BI cloud. The research objectives of this study are a conceptualization of the BI cloud issue and an investigation of some benefits and risks of BI cloud. The results of this research can be used by IT and business leaders as they plan and develop BI cloud in their organizations.

Keywords Cloud Computing, On-Premise, Business Intelligence, Cloud Strategy

1. Introduction

We live in the golden age of distributed Data management. Public cloud platforms now offer practically unlimited compute and storage resources on demand. A few years back, in the 1980's point-of-sale scanner data changed the balance of power between consumer-packaged goods manufacturers. At that point in the late 1990s, web clicks turned into new knowledge currency, empowering online merchants to gain significant competitive advantage. Then the "big data" era comes with new challenges. New sources of social media, mobile and sensor or machine-generated data hold the potential to rewire an organizations value creation process.

Information is pouring in from every conceivable direction - more than 5 billion individuals are calling, texting, tweeting and browsing on mobile phones around the world. Facebook handles more than 250 million photo uploads and the interactions of 900 million active users with more than 900 million objects. Wal-Mart handles more than a million client transactions every hour and imports into databases estimated to contain around 2.5 petabytes of information. The hopeful vision of significant data is that organizations will have the capacity to harvest and harness each byte of data and use it to make the best and meaningful decisions. Big data technologies not only support the ability to collect large amounts but more importantly, the capacity to understand and take advantage of its full value.

To better understand the concept of Cloud Computing, below are few main features:

- Abstract computing and IT service-oriented approach
- Virtual, dynamic, scalable and massive infrastructure
- Shared, configurable, flexible, dynamic resources
- Accessible via the internet from any device
- Platform with minimal management or self-management
- Utilization model based on self-service
- Charging based on consumption

A recent study shows that 70% believe that it would lead to increased business flexibility, 71% of the organizations consider Cloud Computing a realistic technological option, 62% think that it would speed up response to market conditions, and 65% feel that it would lead to greater focus on the most important aspects of the business.

* Corresponding author:
thota.subash@gmail.com (Subash Thota)
Published online at http://journal.sapub.org/ac

The cloud computing paradigm enables rapid on-demand provisioning of server resources to users, with minimal management efforts. The on-demand and elastic nature of resource provisioning make a cloud platform attractive for the execution of various applications, especially computation-intensive ones. More and more data-intensive applications, e.g., Facebook, Twitter, and big data analytics applications are relying on the cloud for processing and analyzing their petabyte-scale data sets. Perhaps more than any other application, data processing is the perfect match for cloud computing with its immense scale and on-demand access to computing and storage capacity.

Data as a Service (DaaS) makes distributed compute and storage technology available instantaneously, scalable and on-demand with minimal configuration effort. DaaS heavily utilizes cloud computing. Over last ten years, the storage and analysis of data have taken place in massive data warehouses. Increasingly though, companies are moving their data into the cloud, making the traditional warehouse method obsolete.

Many believe making a move is a big undertaking, so what are the advantages of doing so?

- Keeping data in the cloud also allows for use across multiple regions. Data in the cloud can be accessed from anywhere, making cloud storage a better solution if you have a business that spans borders.
- Flexibility – Leverage the enormous computing power available on the web, scale up and scale down based on changing requirements.
- Attaining and keeping up a large data warehouse can often be a significant drain on resources, particularly IT budget as you must purchase equipment and ensure the security of the warehouse. With cloud storage, the responsibility is shifted from the company to the cloud service provider – making it a more cost-effective pricing system, pay per use model.
- The speed of Implementation and Deployment – Immediate availability of environment without any dependency on the long stretches associated with infrastructure procurement, application deployment. Drastically reduces the BI implementation time window.
- On-request Availability – Support mobile and remote clients, program based access to control everything from the cloud stage to database management to the data warehouse layer to the analytics stage.

Infrastructure as a Service (IaaS), e.g., virtual machines, storage, networks, or servers, is the most fundamental building block and incorporates anything (actual or virtual) in a data center. One level up exists the Platform as a Service (PaaS) which includes regular software like web and applications like database servers, or Hadoop ecosystem. Next level up is the Software as a Service (SaaS) which are general services like web email, customer relationship management systems or content. At last, beyond SaaS are usually domain or business-specific applications. Hadoop or an alternative distributed compute and storage technology at the platform level naturally builds the core of a DaaS. Consequently, any DaaS solution includes the PaaS layer and potentially SaaS and IaaS. This leaves us with three possible combinations for DaaS:

- PaaS only – focus on Hadoop.
- IaaS and PaaS – focus on Hadoop and optimized infrastructure for performance.
- PaaS and SaaS – focus on Hadoop and features for productivity and exchangeable infrastructure IaaS and PaaS and SaaS – focus on complete vertical integration for features and performance.

The term "Network as a Service" (NaaS) is also part of the nomenclature of cloud computing, NaaS is a cloud service where the capability provided to the cloud service user is to use between inter-cloud network connectivity services and network/transport connectivity services. NaaS includes the improvement of asset allocations by considering network and computing resources. Conventional NaaS services include flexible and extended VPN and bandwidth on demand. NaaS concept materialization also includes the provision of a virtual network service by the owners of the network infrastructure to a third party.

It appears that enterprises are taking slow steps to move data into the cloud. It has been observed that few organizations are using the public cloud frequently for applications which are based on SaaS, but the process of migration of data to the cloud is quite slow despite the cost, elasticity and scalability benefits.

Few challenges of moving data into the cloud are described below:

- Data Security - Security concerns including privacy, availability of the data and integrity, continues to be one of the top most concern for utilizing the Cloud. For few organizations, the concerns over security may be a barrier. However, as more and more organizations move to the Cloud, it is expected that the Cloud vendors will provide a more secure environment than what exists at customer sites today.
- On-premise Integration – A phased approach is always recommended, and there will be a co-existence model until the cloud BI is mature and stable.
- Jumping from experimentation to production is challenging. Implementation of new Cloud environment leads to hurdles with complex data and system integrated work.
- Security and privacy are major cloud challenges as per surveys in the past few years, but over-time experts are trying to strengthen the Cloud securities.
- No standards are determined for cloud services. Some providers offer flat rate while some offers pay-as-you-go service.
- It's been emphasized that moving data into the cloud is terrible. Several discussions happen on building large petabyte-sized data sets into the cloud. It's a fact that such large volumes of data are moved in physical media

like some big servers or other modes of massive transfers.

- Performance - Depends on the size and performance of data warehouses in the Cloud, Significant performance impact if the data exists at a client site, especially when processing and returning large amounts of data.
- Pricing models – Lack of standardized pricing models makes it difficult for customers to select the right one.

Cloud BI Options

The following options list the basic models by which BI can be deployed on the cloud:

Private cloud: The cloud framework is provisioned for exclusive use by a single organization involving multiple consumers (e.g., business units). It may be managed, owned, and operated by the organization, a third party, or some combination of them, and it may exist ON/OFF the premises. When implemented right, it will improve business, but every step in the project raises several security concerns that must be addressed before hand to prevent serious vulnerabilities.

Community cloud: A community cloud is a collaborative effort in which infrastructure is shared between few associations from a specific community with common concerns (security, compliance, jurisdiction), regardless of whether managed or hosted internally or externally or by a third-party. The expenses are spread over fewer users than a public cloud, so some of the cost saving potentials of cloud computing are realized.

Public cloud: The cloud application is provisioned for open use by the people in general. It may be owned, managed, and operated by academic, business, or government or any organization or any combination. It exists on the cloud provider premises. Cloud service providers like Microsoft, Amazon AWS, operate the infrastructure only via the Internet (direct connectivity is not offered).

Hybrid cloud: The cloud framework is a composition of two or more distinct cloud infrastructures (private, public or community) that remain unique and is bound together by proprietary technology that enables portability (e.g., cloud bursting for load balancing between clouds). Such composition expands deployment options for cloud services, allowing IT organizations to use public cloud computing resources to meet temporary needs. This capacity empowers cross breed mists to utilize cloud blasting for scaling crosswise over mists.

Cloud blasting empowers server farms to make an in-house IT framework that backings normal workloads, and utilize cloud assets from open or private mists, amid spikes in preparing requests. Hybrid clouds lack the flexibility, security, and certainty of in-house applications. It gives the adaptability of in-house applications with the adaptation to internal failure and versatility of cloud-based administrations.

Cloud BI Integration Strategy

This underlying stage consolidates recognizing and defining the business objectives, assessing the fiscal perspective, setting up the working gathering, the due dates and the advantages for utilization and support. For the appraisal of the Cloud BI framework, it is essential to perceive resources used as a piece of the everyday operations and in setting up the operation base. Planning the Cloud BI arrangement begins with portraying a methodology and by setting up the major segments that will be the starting stages in sketching out future models. The route to the achievement of the Cloud BI strategy is getting the correct blend between the cloud and inward appraisal for abusing the two game plans. Changes to the cloud may be proficient a tiny bit at a time, and the item gathering models will match with a period because most associations are utilizing hybrid models of Cloud Computing, keeping up enter components of their in-house framework.

Determination of Cloud BI supplier must be done given the latest sources (supplier, counseling firms, the most current articles), whatever is left of the wellsprings of data from market of BI in Cloud Computing. Inside this stage happens the investigation and testing of Cloud BI arrangements, the choice of competitor arrangements, distinguishing proof of real changes and acquiring protection from the chose Cloud BI suppliers. The significant choices, are vital investigations for establishing choices of movement/joining/keeping considering business factors, basic examinations in view of elusive money saving advantage investigations of approach choices and the chose arrangement considering the surveys made.

The way toward choosing the Cloud BI supplier is iterative and depends on the consequences of the specialized and financial investigations. Each cycle recognizes and disposes of the stages of arrangements that are unsatisfactory or inaccessible, designs that are excessively costly or arrangements with an unsuitable level of security. The quantity of emphases relies upon the quantity of accessible applicant arrangements. Considering the last offer of the supplier, the group will quantify the effect on the association of the chose arrangement, evaluating expenses and advantages to deciding the chance of incorporation and the key components under a transaction. At long last, considering the offer and the model of an agreement, the legally binding components are arranged: value, benefit conditions, installment, commitments of the gatherings, and the agreement is shut. In choosing a BI arrangement, associations must consider the most up to date slants on the BI and Cloud Computing market, the present and future needs and the open door for coordination. To be fruitful, the determination of a Cloud BI arrangement must be accomplished equitably by a group of experts in view of good criteria that have been dissected and weighted by the requirements of the association.

Usage of the arrangement might be performed in iterative stages, through constant transmission of information, administration, and procedures toward the cloud, with conceivable comes back from the cloud to operations, facilitated inside. Likewise, usage includes setting an adaptable program of hazard administration, testing the arrangement's execution and the administration of execution.

Information relocation must be accomplished by keeping the ideal harmony between information exactness, the speed of movement, time of non-operation and insignificant expenses. Use of a Cloud BI arrangement includes changes in the control of procedures, for example, information handling, advancement, receipt of data, methods for capacity, documenting and sharing information. Movement of basic applications and foundation to the Cloud and keeping up the business action includes critical HR for the administration of delicate information and of utilization amid relocation.

Toward the finish of the arrangement's execution happens the preparation of clients in working the new frameworks. An administration displays that ought to incorporate strategies on security, the administration of utilization and foundation, hazard administration and ceaseless assessment of the Cloud BI arrangement. Among the most critical are: checking URLs, observing framework assets, observing log records, occasion administration, multi-customer capacities, assessment of agreements with suppliers.

Threats & Opportunities

The Cloud offers several great points, however: infrastructure flexibility, faster deployment of applications and data, cost control, an adaptation of Cloud resources to real needs, improved productivity, etc. The early 2010's Cloud market is overwhelmed by programming and administrations in SaaS mode and IaaS (foundation), particularly the private Cloud. PaaS and people in general Cloud are further back. A few impediments to the Cloud remain. Among them, are dependability, accessibility of administrations and information, security, multifaceted nature, costs, controls and legal issues, execution, relocation, inversion, the absence of models, and constrained customization.

IT governance

The presentation of distributed computing requires a fitting IT administration model to guarantee a secured registering condition and to conform to all applicable hierarchical data innovation approaches. All information considered organizations need an arrangement of capacities that are fundamental when adequately executing and overseeing cloud administrations, including request administration, relationship management, information security administration, application lifecycle administration, hazard and consistent administration. The danger lies with the explosion of companies joining the growth in cloud computing by becoming providers.

Legal

As with other changes in cloud computing, certain legal issues arise with cloud computing, including trademark infringement, sharing of proprietary data resources, and security concerns. One imperative however not regularly said the issue with distributed computing is the issue of who is "under lock and key" of the information. On the off chance that a cloud organization is the holder of the information, the owner has certain legitimate rights. If the cloud organization is the "caretaker" of the information, at that point an alternate arrangement of rights would apply. The following issue in the legalities of distributed computing is the issue of legitimate responsibility for information. Many Terms of Service understandings are noiseless on the topic of possession.

Open Source / Open Standards

Open source software has provided the foundation for many cloud computing implementations, prominent examples being the Hadoop framework. Most cloud suppliers uncover APIs that are commonly very much recorded (regularly under a Creative Commons permit) yet additionally interesting to their execution and in this way not interoperable. A few merchants have embraced others' APIs, and there are a few obvious guidelines being worked on, aiming to convey interoperability and transportability. OpenStack, founded in 2010 by Rackspace and NASA, and now administered by the OpenStack Foundation. OpenStack supporters include Intel, SUSE Linux, Red Hat, IBM, and Rackspace.

Privacy

Privacy advocates have condemned the cloud to demonstrate for facilitating organizations' more noteworthy simplicity can control—and in this manner, can screen voluntarily—correspondence between have organization and end client, and access client information (with or without consent). A cloud specialist organization (CSP) can confound information security because of the degree of virtualization (virtual machines) and distributed storage used to execute cloud benefit. CSP operations, client or occupant information may not stay on a similar framework, or in a similar server farm or even inside a similar supplier's cloud; this can prompt lawful worries over jurisdiction.

Compliance

To comply with regulations including FISMA, clients may need to receive group or half and half arrangement modes that are regularly costlier and may offer confined advantages. This is how Google can "manage and meet additional government policy requirements beyond FISMA," and Rackspace Cloud can claim PCI compliance. Many providers also obtain an SAS 70 Type II audit, be that as it may, this has been condemned because the hand-picked set of objectives and models dictated by the evaluator and the auditee are frequently not unveiled and can change generally. Suppliers commonly make this data accessible on ask for, under a non-revelation understanding.

A huge number of laws and controls have constrained requirements onto many organizations that gather, produce or store information. These strategies may direct a comprehensive exhibit of information stockpiling arrangements, for example, to what extent data must be held, the procedure utilized for erasing information, and even certain recuperation designs. The following are a few cases of laws or directions. The Health Insurance Portability and Accountability Act (HIPAA) requires an alternate course of

action that incorporates, information reinforcements, information recuperation, and information access amid crises.

In a virtualized Cloud Computing condition, clients may never know precisely where their information is put away. Truth be told, information might be put away over different server farms to enhance unwavering quality, increment execution, and give redundancies. This geographic dispersion may make it harder to ascertain legal jurisdiction if disputes arise.

Vendor lock-in

Since distributed computing is still moderately new, norms are yet being created. Many cloud stages and administrations are exclusive, implying that they are based on the guidelines, apparatuses, and conventions created by a seller for its cloud advertising. This can make moving off an exclusive cloud stage restrictively convoluted and costly.

Three types of vendor lock-in can happen with cloud computing:

- Platform lock-in: oud administrations tend to be based on one of a few conceivable virtualization stages, for instance, VMWare or Xen. Relocating from a cloud supplier utilizing one stage to a cloud supplier utilizing an alternate stage could be exceptionally muddled.
- Data lock-in: Since the cloud computing is still new, measures of ownership, i.e., who owns the data once it resides on a cloud stage, are not yet created, which could make it convoluted if distributed computing clients ever choose to move information off a cloud vendor's stage.
- Tools lock-in: Tools built to handle a cloud environment are not compatible with different kinds of both virtual and physical infrastructure, these tools will only be able to manage data or apps that live in the vendor's cloud environment. Heterogeneous distributed computing is characterized as a sort of cloud condition that averts merchant secure and lines up with big business server farms that are working half and half cloud models. The nonattendance of merchant secure let’s clouds head select his or her decision of hypervisors for undertakings, or to convey virtualized frameworks to different ventures without the need to consider the kind of hypervisor in the other endeavor.

2. Sustainability

Even though distributed computing is regularly thought to be a type of green registering, no distributed investigation substantiates this supposition. Referring to the servers' consequences for the ecological impacts of distributed computing, in regions where atmosphere favors regular cooling and sustainable power is promptly accessible, the natural impacts will be more direct. (Similar remains constant for "customary" server farms.) Thus, nations with positive conditions, for example, Finland, Sweden, and Switzerland are attempting to draw in distributed computing server farms. Vitality effectiveness in distributed computing can come about because of vitality mindful booking and server solidification.

3. Conclusions

We have entered an era of Big Data, and an ODS and a Core DWH will surely stay within the confines of the home enterprise for quite a while. Cloud computing promises significant benefits. Cloud BI has been developed to enhance the efficiency, productivity, and performance of business intelligence of BI application. It helps in quick BI implementations, reduction of cost BI applications. Cloud facilitates testing and upgrading of BI programs.

Despite these certain advantages, there are various risks and vulnerabilities during cloud BI using – be it on the side of the tools that still lack some necessary prerequisites or data protection or on the edge of the often-demanding data requirements. It also seems that many factors must be considered to reasonably assess the use of BI clouds (including the size of the organization, its nature, and its strategic objectives). As Cloud BI is not in its maturity yet, ongoing research activities should focus on further structuring, detailing, and evaluating possible scenarios. Research is also required on the contracting aspect for virtualized and outsourced BI infrastructures. Furthermore, there is a necessity for action and design-oriented research to gather real-life experiences and to come to adequate plans. The relevance of such research is evident: The potential benefits of focusing resources on conceptual and organizational issues rather than using them for building and operating hardware and software tools plus the sourcing and flexibility advantages of Cloud BI are only hard to ignore. A completely virtualized BI infrastructure that is composed as a mashup of internet services is admittedly long-term vision than midterm reality – but it still needs to be watched out for.

REFERENCES

[1] Li, H., Spence, C., Armstrong, R., Godfrey, R., Schneider, R., Smith, J., White, J. (2010), Intel Cloud Computing Taxonomy and Ecosystem Analysis, Printed in USA 0210/KC/KC/PDF, 1-4.

[2] Linthicum, D. S. (2009), Moving to Cloud Computing Step-by-Step, The Linthicum Group.

[3] Malcolm, D. (2009), The five defining characteristics of cloud computing, ZDNet.

[4] McKendrick, J. (2007), BI, Delivered from the Cloud, Ebizq Net, The Insider's Guide to Business and IT Agility.

[5] Menken, I. (2009), Cloud Computing - The Complete Cornerstone Guide to Cloud Computing Best Practices: Concepts, Terms, and Techniques for Successfully Planning,

Implementing and Managing Enterprise IT Cloud Computing Technology, Publisher Emereo Pty Ltd, 172 pages, ISBN-10: 1921573007, ISBN-13: 978-1921573002.

[6] Avanade (2009), 2009 Global Survey of Cloud Computing.

[7] Thota, S., 2017. Big Data Quality. *Encyclopedia of Big Data*, pp.1-5. https://link.springer.com/referenceworkentry/10.1007/978-3-319-32001-4_240-1.

[8] Birst (2010), Why Cloud BI? The 9 Substantial Benefits of Software-as-a service Business Intelligence, Birst, Inc.

[9] Ghilic-Micu, B., Stoica, M., Mircea, M. (2008), A framework for measuring the impact of BI solution, The 9th WSEAS Int. Conf. on Mathematics And Computers In Business And Economics, Bucharest, Romania, ISBN 978-960- 6766-76-3, ISSN 1790-5109, 68-73.

[10] Hinchcliffe, D. (2009), Eight ways that cloud computing will change business, ZDNet.

[11] Sheelvant, R. (2009), ten things to know about cloud computing strategy.

[12] Verizon (2009), Start Packing. You're Moving to the Cloud – and We Can Help, Verizon Business.

[13] Wayne, W. E. (2009), Implementing BI in the Cloud, The Data Warehousing Institute.

[14] Marinela Mircea, Bogdan Ghilic-Micu, Marian Stoica, Combining Business Intelligence wit Cloud Computing to Delivery Agility in Actual Economy.

[15] Winans, T., Brown, J.S (2009), Cloud Computing, 1-39.

[16] Berkowitz, J. (2009), Cloud Computing (Part 1): Advantages, Types and Challenges, CRM Mastery Weblog.