Advantages & Disadvantages of Using IBM Big Data Analytics On Cloud

Discussing around the advantages & disadvantages would be just a list. In order to discuss and study in details, we need to try to understand the strategy of a company providing the service, have an overview of the major commonly used products, analyze the documentations and free resources offered.

The term Big Data, which is often used today even by the lay press, is the storage and analysis of large amounts of data from different sources, with the aim of generating an economic benefit for an organization.

This article namely Advantages & Disadvantages of Using IBM Big Data Analytics On Cloud is like a case study which deals with the big data analytics tools from one of the largest IT corporations in the world – IBM. The goal is to determine the importance of incorporating IBM's big data analytics tools in the company's strategy, what big data offering from them already exists and how their offers can be compared with the competition products. From our end, a comprehensive research has been carried out, in particular through publications from IBM on strategy,

products from other companies as well as studying the external sources which contain serious and objective information about the topic under discussion.

Analysis of Strategy of IBM For Advantages & Disadvantages of Using IBM Big Data Analytics On Cloud

In order to generate competitive advantages and thus business benefits from the acquired knowledge, nowadays it is essential for the enterprises and organizations to draw insights from the enormous data available. IBM is focused on the solutions that enable growth, and the workforce of employees dynamically adapt to ever-changing requirements. In order to achieve these goals, the following principles must be considered :

1. Innovation through cloud computing, business analytics, acquisitions and other strategic initiatives
2. Enterprise-wide automation and integration of business processes
3. Introduction of the workplace of the future on the basis of flexible systems, which enable mobility, i.e. the integration of privately used terminals, in

business life.

Another strategic goal of IBM is radical simplification. Therefore, their main focus apparently is to continuously simplify, standardize, automate, or completely eliminate tasks and processes which may prevent integration or cause unnecessary complexities. Other important goals in their strategies are the significant improvement in cycle times and the reduction of transaction costs throughout the company through slimmer, more flexible processes. As management and control system, they offer dashboards which visualize a real-time view of the company's status – from finance to infrastructure.

IBM has been one of the world's largest donors from the business world, providing financial resources and technologies, as well as the know-how of the employees for non-profit purposes (which includes Apache Foundation, Free Software Foundation) primarily for projects in the field of education and open source software development. Social commitment is an important part of IBM's corporate strategy. IBM is well positioned for data analysis because they have more advanced softwares. However, this was also largely bought by taking over significantly smaller companies with appropriate know-how.

Tools Currently Offered By IBM Big Data Analytics On Cloud

IBM offers their customers a complete Big Data offering, ranging from storage, management, to analysis of large amounts of data. IBM's core products, BigInsight and

InfoSphere, enable IBM to simplify the use of Hadoop, enabling mid-size enterprises to benefit from the new technology without the need for multiple Big Data experts.

The range of IBM Big Data solutions is deliberately designed in such a way that not all components have to be introduced, but rather, can be delivered to the customer according to individual requirements. This is made possible by the open-source approach, which they very closely follow. This matter is particularly interesting for the companies which need only limited functionalities and do not want to bear the costs for a complete implementation. In the field of predictive analytics, IBM also offers a real-time evaluation of functionality that most competitors can not afford.

With Watson Foundations, IBM offers their customers the ability to outsource large parts of their Big Data platform, which can be attractive to many customers due to the often lack of know-how. In order to satisfy the customers and to highlight the advantages of its offer against the competition, IBM has built up an extensive Big Data internet portals :

1   IBM Big Data Hub
2   IBM Developer Works Data and analytics articles & tutorials
3   IBM Big Data on YouTube
4   IBM Data Science

It is worthy to mention that we found many GitHub projects (both official and community created), one example :

1   BigInsights on Hadoop (Github repo)

We noticed StackOverflow support by the users as well.

IBM BigSQL, a multipurpose product has forever free to test server instance for the developers. Previously we have discussed about that free IBM BigSQL with SSH access and we have also shown example of server log analysis with Apache Pig:

Advantages Of the Tools Offered By IBM Big Data Analytics On Cloud

As we can clearly see that IBM in reality selling their hardware and networking resources mainly running open source big data softwares. It is quite easy for a system administrator to measure the performance and benchmark to compare on-premise installation of Hadoop and IBM hosted Hadoop on cloud.

IBM's strategy supports potential customers in the right choice of products. With its infrastructure and software portfolio, IBM is able to select the most necessary components and to keep them running without implementing a complete ecosystem. This highlights IBM's SaaS capabilities. This is also supported by the goal of IBM to automate all processes, to dispense with complex systems and thus make it as simple as possible for customers.

The IBM competition is similar. Extra over using on-premise Hadoop to the customers is the availability of IBM's experienced skilled employees. As like any cloud product, the selection of cloud of course due to some

limitations (which is commonly cost) and we observed the similar advantages :

1. **Economical benefit** – Elimination of CAPEX model, low operating cost
2. **Speed** – It is faster to deploy a project without the need of setup
3. **Distributedness**
4. **Reliability through redundancy**
5. **Scalability**
6. **Least burden of server sysadmin works**
7. **Least one-time investment**
8. **Least vendor lock-in**
9. **Least downtime unlike normal server maintenance**

IBM Watson is a specialized product of scientific interest. We are not aware about ready open source product that can replace IBM Watson, but we guess that many cognitive architectures may be tailored to function basic works like IBM Watson.

Disadvantages Of the Tools Offered By IBM Big Data Analytics On Cloud

In true sense, while using a cloud service, it is impractical to head to head compare matters like physical access to on-premise. It can be seen that although IBM wants to become an innovative company, this market segment has only recently discovered itself. IBM was able to secure a place in Big Data market

through its many years of experience, infrastructure and financial strength.

1 **Communication problems** are possible resulting in poorer performance
2 Ambivalence between **security** and ease of access to resources
3 **Higher cost** for too bigger projects
4 **Lesser reach of product offerings**, features to the client resulting in wrong product selection

Terminologies

Strategy is the description of a new and unique position that the company is striving to achieve to cut down on their consequences and needs for the company as precisely as necessary for a clear understanding among the stakeholders, thus enabling successful implementation.

The terminology Predictive Analytics is used in the context of data processing and is a reliable forecasting set of tools which allows the data to predict future developments in a personalized and structured fashion. Predictive Analytics serves a company as well-founded decision-making aid and enables the company to build a competitive edge through innovative ideas. During the decision-making phase of a company, the data obtained shall be used to take and derive measures for the occurrence of opportunities and risks.

The SWOT analysis is an analysis method and an abbreviation for the terms "strengths, weaknesses,

opportunities and threats".

Open Source software refers to the computer softwares in which the source code is available to the general public for use or modification.

A NoSQL system is a non-relational, distributed data storage system that often does not meet the ACID criteria (atomic, consistency, isolation, and durability).

Free software, also known as software libre refers to the computer softwares "distributed under terms which allows the software users to run the software for any purpose as well as to study, change, and distribute the software and any adapted versions".

The Apache Software Foundation is a non-profit corporation to support Apache software projects, with license compatible with Free software and Open Source.