

Chapter 91

A Survey on Comparison of Performance Analysis on a Cloud-Based Big Data Framework

Krishan Tuli

Chandigarh University, India

Amanpreet Kaur

Chandigarh University, India

Meenakshi Sharma

Galgotias University, India

ABSTRACT

Cloud computing is offering various IT services to many users in the work on the basis of pay-as-you-use model. As the data is increasing day by day, there is a huge requirement for cloud applications that manage such a huge amount of data. Basically, a best solution for analyzing such amounts of data and handles a large dataset. Various companies are providing such framesets for particular applications. A cloud framework is the accrument of different components which is similar to the development tools, various middleware for particular applications and various other database management services that are needed for cloud computing deployment, development and managing the various applications of the cloud. This results in an effective model for scaling such a huge amount of data in dynamically allocated recourses along with solving their complex problems. This article is about the survey on the performance of the big data framework based on a cloud from various endeavors which assists ventures to pick a suitable framework for their work and get a desired outcome.

DOI: 10.4018/978-1-6684-3662-2.ch091

1. INTRODUCTION

1.1 Getting Together with Cloud and Data Analytics: Flawless Alliance

As we all know that, cloud is working & delivering their services on Pay as per use, on demanded and provide scalable services. Data Analysis mainly focuses on the five v's of big data. They are the characteristics of big data.

- First one is the, volume, which means the total amount of data.
- Variety is called the unusual set-up of data or we can say is the different formats of data that can be used in Big Data or data analytics. For example, unstructured videos and pictures, unformatted numbers etcetera.
- Velocity is defined as the speed on which datasets is increasing.
- Furthermore, there is veracity on which the accuracy and trustworthiness of structured or unstructured data and information.
- Value is helpfulness of data.

Huge data over the network can be managed efficiently by Cloud Computing with the help of virtualization technique, consequently, adding the big data availability, accessibility & Scalability (Khan, Shakil, & Alam, 2017; Skourletopoulos, Mavromoustakis, Mastorakis, Batalla, Dobre, Panagiotakis, & Pallis, 2016). Furthermore, Distributed computing also brings high-class numerical trappings for ingenious dispensation and studies of large amount of data over a provision named as Huge Records as a Provision (Khan, Shakil, & Alam, 2017; Zhang, Zhou, Li, & Gao, 2017). Afterward, in cooperation with huge amount of data and cloud unite collected to carry an importance to initiatives by attractive the nimbleness, bounciness, convenience and the comfort of meting out of distributed created large information, and, by dropping its price of possession and application difficulty of large information resolutions (Khan, Shakil, & Alam, 2017). A cloud framework is the best solution for solving this problem & managing the data at sustainable cost. A cloud framework is the accrument of different components which is similar to the development tools, various middleware for particular application & various other database management services that is needed for cloud computing deployment, development & managing the various applications of cloud which results in an effective model for scaling such a huge data at a dynamically allocated recourses along with solving their complex problems.

1.2 Enterprise Based Data Analytics – Distributed

The Social Networking sites and applications that are being used by many people and become their hobby to use such sites, for example: Tweeter, Snapchat, Instagram, WhatsApp Messenger, Facebook & a variety of other online venture application are the enterprise period that broad cast a wide-ranging of data (Exabyte& petabyte data) that is on day to day basis. Such a large amount of data which may be spawning during messages, satellite images, social networking sites, Electronic mail and many more can be structured and unstructured depending on their availability &accessibility and is known to be Big Data. (Skourletopoulos et al., 2017; Math, 2017; Pokorny & Stantic, 2016). As a result, various ventures are in front of a demanding job of tracking and managing such a huge sum of data.

9 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:

www.igi-global.com/chapter/a-survey-on-comparison-of-performance-analysis-on-a-cloud-based-big-data-framework/291070

Related Content

Big Data Applications in Vaccinology

Joseph E. Kasten (2021). *International Journal of Big Data and Analytics in Healthcare* (pp. 59-80).

www.irma-international.org/article/big-data-applications-in-vaccinology/276927

Development of a New Means to Improve the Performance of Self-Organizing Maps

Vijaya Prabhagar Murugesanand Punniyamoorthy M. (2022). *International Journal of Data Analytics* (pp. 1-16).

www.irma-international.org/article/development-of-a-new-means-to-improve-the-performance-of-self-organizing-maps/307065

SEO Through the E-Customer Journey

Ana Fernández López (2023). *Big Data Marketing Strategies for Superior Customer Experience* (pp. 165-193).

www.irma-international.org/chapter/seo-through-the-e-customer-journey/322196

Visualization of High-Level Associations from Twitter Data

Luca Caglieroand Naeem A. Mahoto (2014). *Packaging Digital Information for Enhanced Learning and Analysis: Data Visualization, Spatialization, and Multidimensionality* (pp. 164-183).

www.irma-international.org/chapter/visualization-of-high-level-associations-from-twitter-data/80217

Solving Solid Transportation Problems With Multi-Choice Cost and Stochastic Supply and Demand

Sankar Kumar Royand Deshabrata Roy Mahapatra (2018). *Intelligent Transportation and Planning: Breakthroughs in Research and Practice* (pp. 137-170).

www.irma-international.org/chapter/solving-solid-transportation-problems-with-multi-choice-cost-and-stochastic-supply-and-demand/197131