

Chapter 3

Emerging Trends and Techniques in Cloud-Based Data Analytics

Shaila S. G.

Dayananda Sagar University, India

Monish L.

Dayananda Sagar University, India

Lavanya S.

Dayananda Sagar University, India

Sowmya H. D.

Dayananda Sagar University, India

Divya K.

Dayananda Sagar University, India

ABSTRACT

The new trending technologies such as big data and cloud computing are in line with social media applications due to their fast growth and usage. The big data characteristic makes data management challenging. The term big data refers to an immense collection of both organised and unorganised data from various sources, and nowadays, cloud computing supports in storing and processing such a huge data. Analytics are done on huge data that helps decision makers to take decisions. However, merging two conflicting design principles brings a challenge, but it has its own advantage in business and various fields. Big data analytics in the cloud places rigorous demands on networks, storage, and servers. The chapter discusses the importance of cloud platform for big data, importance of analytics in cloud and gives detail insight about the trends and techniques adopted for cloud analytics.

DOI: 10.4018/978-1-7998-3111-2.ch003

1. Why Cloud Platform for Big Data?

Big data refers to storage of bulk data whereas cloud computing deals with storage and computing. Big data in computing environment can be managed by using distributed storage. Traditional storage method is not sufficient for storing big data. Hence, cloud evolved to provide the solution. The expansion of big data storage is done by cloud through virtual machines and makes big data more accessible. Initially, big data and cloud are used in various ways. Many investors decided to join the two techniques to make a maximum business outcome. Both the technologies aimed to reduce the investment cost and increase the income of the organizations. Microsoft, Google, amazon are successful in using big data in cloud platform. The world runs online. People surf data and leave data online. For example: people use internet in various ways like social media such as Facebook, twitter etc., for ecommerce websites like flipkart, amazon, etc, to get any valuable information they can use Google search engine and so on and leave data online. A new survey found that 42 percent of Americans use the internet several times a day. Companies gather all this data and utilize accordingly to provide better products to satisfy customer requirements and increase their business turn over (Singh and Reddy, 2014). The grouping of these two techniques led to the drastic changes and improvements. It gave huge advantage to analysts and changed the decision-making process for companies, and their outcome is improved.

Cloud technology is a fast-growing field as it offers services to the consumers on a pay-as-you-go model. Instead of storing a data in personal computer or local servers the data is stored in cloud which is easy to manage and provide security. Cloud offers the features such as pay as you-go model, storing, processing, etc. Another feature is that cloud supports large network access in which user from any location can access and upload information to the cloud using their devices. These capabilities are accessible over the network. Availability of the cloud makes it extensible based on requirements. Based on user necessities extra cloud storage (small amount) is provided for their usage. The most important feature of any cloud platform is security. It is one of the best features of cloud platform. The data is safeguarded even if one of the server goes down because cloud creates snapshot of data that is stored. The data that is stored in cloud is neither hacked nor exploited by the intruder, because data is stored within the storage devices with high security. Another feature in cloud is it support data analytics. Cloud offers three primary service models which consist of Platform as a Service (PaaS), Software as a Service (SaaS), and Infrastructure as a Service (IaaS). Platform as a Service uses public, private and hybrid clouds. PaaS does not include hosting, but it provides open source software. The Cloud contributor provide resources like object storage, runtime, queuing, databases, etc. Customers are able to develop, run and manage applications by using this platform. Ex: Google App Engine, Herokus, etc. SaaS is known as on-demand software or web-based software or hosted software. This service provides all the necessary settings and infrastructure in place. There are two types of SaaS that is, Vertical SaaS and Horizontal SaaS. Ex: Salesforce Customer Relationship Management (CRM). IaaS is high-level APIs and is provided through the internet. It offers resources such as virtual-machine, disk-image library, IP addresses and virtual local area networks. This service provider provides whole structure along with the related tasks such as Amazon Elastic Compute Cloud (EC2).

The benefit of using cloud platform is that user need not have to develop the infrastructure from the scratch for storage. Apart from this if companies needs to do some analysis, they can create the virtual copies for analytics. Analytics speed up the process and helps in achieving goals. Cloud simplifies the connectivity within an organization which gives employee access to relevant analytics. The recovery solutions and backups are also implemented within cloud. Agility is yet another advantage because

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/emerging-trends-and-techniques-in-cloud-based-data-analytics/269555

Related Content

Energy-Saving QoS Resource Management of Virtualized Networked Data Centers for Big Data Stream Computing

Nicola Cordeschi, Mohammad Shojafar, Danilo Amendola and Enzo Baccarelli (2015). *Emerging Research in Cloud Distributed Computing Systems* (pp. 122-155).

www.irma-international.org/chapter/energy-saving-qos-resource-management-of-virtualized-networked-data-centers-for-big-data-stream-computing/130270

Governance and Risk Management in the Cloud with Cloud Controls Matrix V3 and ISO/IEC 38500:2008

Abhik Chaudhuri (2015). *Handbook of Research on Security Considerations in Cloud Computing* (pp. 80-101).

www.irma-international.org/chapter/governance-and-risk-management-in-the-cloud-with-cloud-controls-matrix-v3-and-isoiec-385002008/134288

Green Cloud Computing: Data Center Case Study

Ahmed Abdul Hassan Al-Fatlawi and Seifedine Kadry (2014). *Handbook of Research on Cloud Infrastructures for Big Data Analytics* (pp. 508-533).

www.irma-international.org/chapter/green-cloud-computing/103228

Recent Advances in Edge Computing Paradigms: Taxonomy Benchmarks and Standards for Unconventional Computing

Sana Sodanapalli, Hewan Shrestha, Chandramohan Dhasarathan, Puviyarasi T. and Sam Goundar (2021). *International Journal of Fog Computing* (pp. 37-51).

www.irma-international.org/article/recent-advances-in-edge-computing-paradigms/284863

From Cloud Computing to Fog Computing: Platforms for the Internet of Things (IoT)

Sanjay P. Ahuja and Niharika Deval (2018). *International Journal of Fog Computing* (pp. 1-14).

www.irma-international.org/article/from-cloud-computing-to-fog-computing/198409