Chapter 65 Analytics-as-a-Service (AaaS): An Elucidation to SOA

Chitresh Verma Amity University, India

Rajiv Pandey Amity University, India

ABSTRACT

Big Data Analytics is a major branch of data science where the huge amount raw data is processed to get insight for relevant business processes. Integration of big data, its analytics along with Service Oriented Architecture (SOA) is need of the hour, such integration shall render reusability and scalability to various business processes. This chapter explains the concept of Big Data and Big Data Analytics at its implementation level. The Chapter further describes Hadoop and its technologies which are one of the popular frameworks for Big Data Analytics and envisage integrating SOA with relevant case studies. The chapter demonstrates the SOA integration with Big Data through, two case studies of two different scenarios are incorporated that integrates real world implementation with theory and enables better understanding of the industrial level processes and practices.

BIG DATA: AN INTRODUCTION

Big Data as a terminology is mistaking as it is not small or big in term of data, but size in terms of volume as well as type of the data (structured/unstructured) in system. The Big Data is normally defined as the data set which is beyond the ability of traditional system to process. (Zikopoulos et al., 2011)

Evolution of Big Data and Beyond

Figure 1 the big data landscape envisages a huge collection of Technologies, Architectures and concepts. The evolution of Big Data can be traced backward to dot com period of late 1990. The record of many years as well as the rate of generation of the data has reached new high in the process of evolution. The Big Data is data which is generated by the various sources primarily the social network, extending to

DOI: 10.4018/978-1-5225-7501-6.ch065

Analytics-as-a-Service (AaaS)

Figure 1. Big data landscape Source: https://www.wikipedia.com, Wikipedia, 2016)



Internet of things and high end Information and analysis system like black box of airplane, DNA and forensic analysis and stock markets.

The term Big Data was coined by "Gartner Inc." in 2007. In 2012, Gartner defined "Big data is high volume, high velocity, and/or high variety information assets that require new forms of processing to enable enhanced decision making, insight discovery and process optimization". They suggest multiple V's to define the Big Data and these V's are discussed in the next section.

V's of Big Data

The Vs are used for defining the Big Data. (Sagiroglu, S. et.al. 2013) The Vs of Big Data can define in terms Volume, Value, Velocity, Veracity and Variety and shown in figure 2. (Verma, C. et. al., 2016) The different Vs are described below.

Volume

The volume measures the data in the terms of size. Due to high rate of generation with parallelism has led to many fold increase in the amount of the data.

Value

The value is cost of the data in term of its worth to the users. For example, the data about discount offer on certain product to prospective customer may have high value and same data may no value for uninterested person. 18 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/analytics-as-a-service-aaas/217886

Related Content

Service-Oriented Solution Framework for Internet Banking

Tony Chao Shanand Winnie Wei Hua (2006). *International Journal of Web Services Research (pp. 29-48)*. www.irma-international.org/article/service-oriented-solution-framework-internet/3073

Learning Workflow Models from Event Logs Using Co-clustering

Xumin Liuand Chen Ding (2013). International Journal of Web Services Research (pp. 42-59). www.irma-international.org/article/learning-workflow-models-from-event-logs-using-co-clustering/100661

SDSCCM: Secure Distributed System Communication for Cloud-Based Manufacturing

Danish Javeed, Tianhan Gao, Muhammad Shahid Saeed, Rafi Ullah Khanand Zeeshan Jamil (2023). *Protecting User Privacy in Web Search Utilization (pp. 200-214).* www.irma-international.org/chapter/sdsccm/322592

A Statistical Analysis for the Accessibility of Electronic Data Delivery System of the Central Bank of the Turkish Republic

Yakup Ari (2022). *App and Website Accessibility Developments and Compliance Strategies (pp. 38-57).* www.irma-international.org/chapter/a-statistical-analysis-for-the-accessibility-of-electronic-data-delivery-system-of-thecentral-bank-of-the-turkish-republic/287253

A Services Classification Method Based on Heterogeneous Information Networks and Generative Adversarial Networks

Xiang Xie, Jianxun Liu, Buqing Cao, Mi Peng, Guosheng Kang, Yiping Wenand Kenneth K. Fletcher (2023). *International Journal of Web Services Research (pp. 1-17).*

www.irma-international.org/article/a-services-classification-method-based-on-heterogeneous-information-networks-andgenerative-adversarial-networks/319960