

# Chapter 83

## Evolution of Cloud in Big Data With Hadoop on Docker Platform

**Meenu Gupta**

*Ansal University, India*

**Neha Singla**

*Punjabi University, India*

### ABSTRACT

*Data can be anything but from a large data base extraction of useful information is known as data mining. Cloud computing is a term which represent a collection of huge amount of data. Cloud computing can be correlated with data mining and Big Data Hadoop. Big data is high volume, high velocity, and/or high variety information asset that require new form of processing to enable enhanced decision making, insight discovery and process optimization. Data growth, speed and complexity are being accompanied by deployment of smart sensors and devices that transmit data commonly called the Internet of Things, multimedia and by other sources of semi-structured and structured data. Big Data is defined as the core element of nearly every digital transformation today.*

### INTRODUCTION

Big data is a term that indicates a very large amount of data which floods and controls a business on a daily basis. Moreover, what really matters are the manipulations of data in industry specific environment. Big data can be analyzed for insights that lead to better decisions and strategic business moves. The data has to be surveyed regularly, analyzed before it undergoes any business action. The general advantages of big data are uncovered truths, predict product and consumer trends, reveal product reliability, brand loyalty, manage personalized value chains and discover real accountability. IT industry is arranging big data processing on regular basis with the good relation to Cloud Based IT solutions (Bollier, 2010). Cloud-based Big Data solutions are hosted on Infrastructure as a Service (IaaS), delivered as Platform as a Service (PaaS), or as Big Data applications (and data services) via Software as a Service (SaaS)

DOI: 10.4018/978-1-5225-7501-6.ch083

manifestations. Each must follow some of the Service Level Agreements (SLAs) for the business development. Cost effectiveness is a must to provide a better service to the industry. *For* example, applications and better customer experiences are often powered by smart devices and enable the ability to respond the moment customer act. Smart products being sold can capture an entire environmental context. New analytical techniques and models are being developed by business analysts and data scientists to uncover the value provided by this data (Bhosale & Gadekar, 2014).

## **Goals of Big Data**

- The enterprises should invest in acquiring both tools and skills as Big Data analytics strictly depend on analytical skills and analytics tools.
- The Big Data strategy must involve an evaluation of the decision-making processes of the organization as well as an evaluation on the groups and types of decision makers.
- To discover new metrics, operational indicators and new analytic techniques, to look at new and existing data in a different way. This, generally, requires setting up a separate Big Data team with research purpose.
- The Big Data is required to support concrete business needs and provides new reliable information to decision makers.
- The most suitable technology can, only, meet all the Big Data requirements due to presence of different workloads, data types, and user types. For example, Hadoop could be the best choice for a large-scale Web log analysis but is not suitable for a real-time streaming at all. Multiple Big Data technologies must coexist and address use cases for which they are optimized.

In this, we are mainly focusing on Hadoop, working of Big Data on Docker platform and its relation with Cloud Computing.

## **MAIN FOCUS OF THE CHAPTER**

While the term “big data” is quite new in the IT field, the act of storing large amounts of information for eventual analysis is old aged.

Big data represents a prototype shift in the technologies and techniques for storing, analysing and leveraging information assets. That can be characterized by 7 V's (Khan M. Ali-ud-din et al., 2014) (Figure 1):

- Volume
- Velocity
- Variability
- Variety
- Value
- Veracity
- Visualization

20 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/evolution-of-cloud-in-big-data-with-hadoop-on-docker-platform/217904](http://www.igi-global.com/chapter/evolution-of-cloud-in-big-data-with-hadoop-on-docker-platform/217904)

## Related Content

---

### Estimating the Privacy Protection Capability of a Web Service Provider

George O.M. Yee (2009). *International Journal of Web Services Research* (pp. 20-41).

[www.irma-international.org/article/estimating-privacy-protection-capability-web/4102](http://www.irma-international.org/article/estimating-privacy-protection-capability-web/4102)

### Analysis of Grievances in the Banking Sector through Big Data

Arushi Jain and Vishal Bhatnagar (2019). *Web Services: Concepts, Methodologies, Tools, and Applications* (pp. 2144-2160).

[www.irma-international.org/chapter/analysis-of-grievances-in-the-banking-sector-through-big-data/217935](http://www.irma-international.org/chapter/analysis-of-grievances-in-the-banking-sector-through-big-data/217935)

### Big Data and Data Analytics Research: From Metaphors to Value Space for Collective Wisdom in Human Decision Making and Smart Machines

Miltiadis D. Lytras, Vijay Raghavan and Ernesto Damiani (2019). *Web Services: Concepts, Methodologies, Tools, and Applications* (pp. 2161-2171).

[www.irma-international.org/chapter/big-data-and-data-analytics-research/217936](http://www.irma-international.org/chapter/big-data-and-data-analytics-research/217936)

### Big Data Techniques for Supporting Official Statistics: The Use of Web Scraping for Collecting Price Data

Antonino Virgillito and Federico Polidoro (2019). *Web Services: Concepts, Methodologies, Tools, and Applications* (pp. 728-744).

[www.irma-international.org/chapter/big-data-techniques-for-supporting-official-statistics/217860](http://www.irma-international.org/chapter/big-data-techniques-for-supporting-official-statistics/217860)

### WSBen: A Web Services Discovery and Composition Benchmark Toolkit

Seog-Chan Oh and Dongwon Lee (2009). *International Journal of Web Services Research* (pp. 1-19).

[www.irma-international.org/article/wsben-web-services-discovery-composition/3131](http://www.irma-international.org/article/wsben-web-services-discovery-composition/3131)