Chapter 4
The Compute Infrastructures for Big Data Analytics

Pethuru Raj
IBM India Pvt Ltd, India

ABSTRACT

The implications of the digitization process among a bevy of trends are definitely many and memorable. One is the abnormal growth in data generation, gathering, and storage due to a steady increase in the number of data sources, structures, scopes, sizes, and speeds. In this chapter, the author shows some of the impactful developments brewing in the IT space, how the tremendous amount of data getting produced and processed all over the world impacts the IT and business domains, how next-generation IT infrastructures are accordingly getting refactored, remedied, and readied for the impending big data-induced challenges, how likely the move of the big data analytics discipline towards fulfilling the digital universe requirements of extracting and extrapolating actionable insights for the knowledge-parched is, and finally, the establishment and sustenance of the dreamt smarter planet.

INTRODUCTION

The big data era is steadily setting in to firmly settle amongst us. Enterprises in its long and arduous journey are increasingly becoming data-driven in order to systematically zoom ahead of their competitors by smartly leveraging their streaming and accumulated data assets. These days every decision, small or big, is being taken based on data by individuals, innovators and institutions. The much-complicated and enterprise-wide strategy-making process solely depends on data. The persistence on the slogan "more with less" is bound to see a neat and nice reality with the sparkling and scintillating big data idea. There are concerted efforts underway in order to ensure

DOI: 10.4018/978-1-4666-5864-6.ch004

and enable data availability, accessibility, accuracy, integrity, and confidentiality. The utility and usability of data are going up significantly while robust and resilient mechanisms for data capture, loading, storage, transformation, and transition data manipulation, analysis, filtering, cleansing, polishing, mining, and processing for knowledge discovery; information visualization and dissemination tasks are fast-emerging and maturing. On the positive side, the technologies and tools enabling the transition from data to information and to knowledge are really competent and captivating. The noteworthy advancements in the big data discipline are being taken note of by business executives, decision-makers and leaders in a positive manner. There is a close synergy among university professors and industry professionals in identifying inhibiting issues, perpetual drawbacks, limitations and barriers in order to come out with appropriate technology-sponsored solutions to speed up the adoption(Ahuja & Moore (2013)).

Heads of all kinds of business horizontals and verticals are purposefully planning and prototyping big data projects in order to fully understand the tactics, strategic advantages, any hidden risks, business, technical and user benefits, etc. towards framing and formulating a flexible and futuristic big data strategy. There is a greater awakening and articulation in embracing and encapsulating standards-compliant big data platforms, integrated processes, design patterns, evaluation metrics and best practices in order to reap all the originally envisaged benefits of the big data paradigm. There are a number of generic as well as specific big data applications and services emanating and evolving consistently based on the big data ideas from the development community.

The principal application of big data is nonetheless the big data analytics, which grandiosely promises extracting and exposing actionable and pragmatic insights from data heap for speedier decision-enablement and efficient actuation. The subject of data analytics in the big data era is a serious and strategic step which is not to be

taken lightly by growing companies in order to be distinctive and dramatic in their offerings, operations and outputs. There are a multitude of reports from leading market analysis and research groups to substantiate the importance of the big data idea in propping up and propelling industry segments in their everyday activities. The much-anticipated business automation, augmentation and acceleration are bound to get a strong improvement and impetus with the maturity of big data products from highly accomplished and acclaimed product vendors.

In this chapter, I would like to highlight the infrastructural requirements for next-generation big data analytics. Especially the series of long-lasting infrastructural optimizations being represented by the raging cloud idea has a deeper impact on the big data analytics. This chapter dedicatedly details about the server infrastructures on public, private and even hybrid cloud environments.

THE BIG DATA COMPUTING DISTINCTIONS

As we all know, the volume of data getting generated globally has been growing at a phenomenal scale and pace. Apart from the volume, the growing data variety further adds to the complexity of data management, governance, processing, mining, and analysis. Generally data are being produced by diverse, distributed and decentralised machines (sensors, devices, etc.) as well as by men in their everyday knowledge activities. Thereby data generation, capture and accumulation have been a continuously running process and it is expected that the data growth is not to decelerate anytime soon. The uninhibited growth in local, metro and wide area networks (the Internet, intranets and extranets) and communication (wireline as well as wireless) infrastructures contribute for the unparalleled swelling of data. The last mile connectivity is another trend catapulting data creation. With the emergence of social sites as a powerful and

Related Content

Resource Allocation With Multiagent Trading Over the Edge Services

Yee-Ming Chenand Chung-Hung Hsieh (2022). *International Journal of Fog Computing (pp. 1-11).*

www.irma-international.org/article/resource-allocation-with-multiagent-trading-over-the-edge-services/309138

Fog Computing Qos Review and Open Challenges

R. Babu, K. Jayashreeand R. Abirami (2018). *International Journal of Fog Computing (pp. 109-118).*

www.irma-international.org/article/fog-computing-qos-review-and-open-challenges/210568

Multi-Layer Token Based Authentication Through Honey Password in Fog Computing

Praveen Kumar Rayani, Bharath Bhushanand Vaishali Ravindra Thakare (2018). *International Journal of Fog Computing (pp. 50-62).*

www.irma-international.org/article/multi-layer-token-based-authentication-through-honey-password-in-fog-computing/198412

Big Data and Its Visualization With Fog Computing

Richard S. Segalland Gao Niu (2018). *International Journal of Fog Computing (pp. 51-82).*

www.irma-international.org/article/big-data-and-its-visualization-with-fog-computing/210566

Industrial Patterns on Cloud

Sreekrishnan Venkateswaran (2017). *Handbook of Research on End-to-End Cloud Computing Architecture Design (pp. 73-103).*

www.irma-international.org/chapter/industrial-patterns-on-cloud/168149