



International Journal of Emerging Technologies in Computational and Applied Sciences(IJETCAS)

(Open Access, Double Blind Peer-reviewed, Refereed and Indexed Journal)

www.iasir.net

Big Data and Cloud Computing in SDN

Ahmed Dheyaa Basha

Faculty of Computer Science & Information Technology
University of Sumer, Thi-Qar- ALQul, aa 64003, Iraq

Abstract: Data has become the most powerful and powerful application in all applications such as Smart Applications, IT industry, and IT business. Massive data such as the word used to identify datasets that exceed the size of the typical database software tools on storage, management and analysis. Therefore, it has the ability to make very fast or faster progress in many scientific fields and improve the profitability of many companies. Massive data presents many difficulties in good execution, including scalability, storage, and noise accumulation, falsification, and measurement errors. These challenges are characterized and studied and require a new mathematical and statistical model. This paper presents the general and background characteristics of the huge data and related technologies, issues and challenges with a focus on the distinctive features of the massive data, and discusses some ways to deal with big or massive data

Keywords: Big Data, technologies, issues and challenges, SDN

I. Introduction

MapreduceHadoop defines massive data as "data sets that are difficult to capture, difficult or difficult to measure, or manipulated across the same computers in a simple module." McKinsey and Company, the public consulting agency, has announced huge data that are ways of innovation, competition and productivity. Massive data means data sets that are difficult to manipulate quickly, like running on a simple database model. The first indication is that the sizes of data sets that are consistent with the large data standard change may grow over time or with technological advances; second, the data sets that correspond to the large data standard differ in different applications. Overall, large data today is measured around several types of TB to several PBs, starting with McKinsey and the company, and the size of the data set is not the only criterion for large data. The next two key features include the growing range of data to be increased and the ways in which they are to be operated more aggressively or through traditional database techniques.

Massive or big data has been defined as an analyst in META. At present, Gartner identified the challenges and opportunities that massive data brought with the 3Vs model, such as, Velocity, and Variety. This unit was not originally used for massive data, but Gartner and many other companies, including IBM and some research departments at Microsoft, still use the "3V" model to describe the massive data over the last 10 years [2-3]. In the "3Vs" model, the size, with the generation and collection of data blocks, means that the volume of data increases dramatically; speed means the timing of large data, specifically data collection and analysis, must be made quickly and in a timely manner to maximize the commercial value of large data; Diversity into different data types, which include structures, semi-structured and unstructured data such as audio, video, web page and text, as well as traditional structured data. IDC data is defined in 2011 "Massive data technologies described as a new model of architecture techniques, used to obtain the economic extraction value of very large amounts of a wide range of data, by enabling capture and / detection and / or high-speed analysis". With this definition, the characteristics of large data can be summed up as four types, such as size, and velocity (rapid generation), value (large data value but very low density) [4-5]. This idea is arranged because it is important, and necessary for massive data requirements such as detecting hidden large values, points to the most important problem in large data of this definition, namely how to discover values from large-scale data sets and different types, rapid generation. As Jay Parekh, deputy chief engineer at Facebook, said, "You can only have a set of data that contains large data if you do not use the data collected." big data identified by the National Institute of Standards and Technology as "large data should explain data requirements at speed Data, quantity, and massive representation of data limits the ability to use traditional relational methods for effective analysis or data that can be effectively addressed using significant horizontal approximation techniques," which focuses on the technological aspect of large data. Effective needs to be developed and used to analyze and process massive data. There are many important discussions from both the industry and academia on the definition of large or big data [6-7-8].

II. Big data Applications

Large data can be used in different types of uses such as IT applications, Internet-based applications, health care applications, and scientific research applications. In addition, all corporate application requirements used in the Google, Amazon, Facebook, Microsoft, and Oracle applications have started to develop large data projects. A large

resources at the system level; large data at the top level is supported by cloud computing, the availability of database-like rolls and efficient data processing capability. EMC president Kissinger pointed out that the application of massive data must be based on cloud computing. The rapid growth of application requirements and cloud computing developed by virtualization technologies is leading to the development of large data. Cloud computing, as well as adding more account flexibility and fast manipulation of large data, serve as service insights. Cloud computing privileges include enhanced data development, and both complement each other [13]. In the Internet object model, the interaction between Internet objects and massive data includes expanded interconnection of advanced smart devices for network sensors embodied in multiple devices, applications and machines in the real world.

Data access requirements in the Internet include transport facilities, appliances, utilities and mobile devices, which are different between the vast data produced by Internet objects, which are low frequency, low data retrieval and low bandwidth data. Use devices that communicate with each other because of different and different types of data collected, most of which are classic in size, their diversity, non-structural features, noise and high frequency. Now, many IoT standards data are proposed to facilitate and simplify application programmers and service provider functions as explained in Fig. 2.



Figure 2. Description and enumeration of the data acquisition equipment in IoT

IV. Architecture of data

A data center is a data storage platform focused on big data model. It is also subject to large responsibilities, likes, manages data, formats data, obtains data, and takes advantage of data values and functions. Data centers relate to "data" mainly with the word "center". It includes huge amounts of large data and applications that may be data management applicable to application requirements, life cycle mode, which are executable, make a good location and a good source. The emergence of large data brings voice to development opportunities and the great challenges of data centers. So the huge data is the emerging model, the massive growth of infrastructure and related software in the data center. The physical data center network is the basis for massive data support, and is also the basic infrastructure urgently required. [17]. Massive data needs to be supplied to D.C to provide behind-the-scenes providers. The large data model has stricter requirements for storage, processing capabilities and network transmission capacity. In the huge data center, the data center must care about hardware facilities and also give power and soft capabilities, such as manipulating, manipulating, merging, analyzing, affixing, and complying with hardware from massive data. DC aims to help business people analyze current data that is then used to help detect IT problems and discover solutions for large data, so large data opens more avenues to the data center.

V. Software-Defined Networking (SDN)

In order to increase the amount of data for various applications over the Internet, SDN becomes one of their protocols in SDN is an open flow also called SDN enabled tool. First, the term SDN was formulated to represent ideas and work around Open Flow at Stanford University. SDN according to the original definition, this applies to the control network this model. Control plane separated from the data level. Many SDN protocol standards were developed in real applications. At present, SDN is not limited to any application, but is a generic term for the platform [19]. The Open Networks Foundation (ONF) defines and transmits SDN as "in the network structure, the control and data systems are separated, the network intelligence and the state are logically centralized, and the basic network infrastructure is extracted from the applications." A non-profit consortium dedicated to the development and consolidation of SDN.

VI. Processing SDN in Cloud Datacenters

It is critical to efficiently map and manage cloud data center resources to meet service level agreements for large data applications because large data is typically processed in cloud data centers, and as a result resource

requirements for large data applications dynamically change in large data. The Service Level Agreement for the Application of Massive Data is the negotiated agreement between the services of the data provider and its users. It defines the huge data feature, including service level objectives, QoS, and penalties within implementation to be met by large data service providers. Some common service level agreements for the performance of large data applications are response time, processing time, large amount of defects, defects, security, negotiation time, overtime, cost, amount of data being processed, data emitted, quality of service, Provide data .The technological direction of cloud computing is to push data center service providers to create more flexible services for better performance and better security. With SDN-enabled data centers, and requirements on networks that are provided as a service and completed by SDN. From the infrastructure perspective, the cloud platform is built on the basis of data centers, which are called cloud data centers. Cloud service providers can host this infrastructure for massive data applications. Software-based networks based on the cloud data center are widely studied for large data applications. Specifically, the SDN-based OpenFlow network is proposed with stand-alone built-in keys and crosspoint input (CICQ) to schedule packets for different bigdata applications. In this approach, the controller maintains the bandwidth table for the different types of large data applications and sends them to the CICQ keys. The switches then prioritize the package scheduling based on the bandwidth management table of the console.

VII. Delivery SDN and Big Data Applications

Given the large size, data presentation in large data applications is a major challenge. Samadi suggests. Et al. (2014) on an SDN network based optical network to accelerate it, providing delivery traffic to huge applications. Massive data applications have led to increased traffic volume characterized by diverse communication patterns that indicate a set of unicast and multicast broadcasts in all types of casting. The hybrids include the optical and electrical approach provided by Samadi et al., Which utilizes the physical optics of the layer to accelerate movement Traffic for each pattern. The application-based control level corresponds to the SDN, allowing flexible and dynamic runtime for photonic devices to support complex traffic patterns [21].To deal with explosion data in massive data applications, Sadasivarao et al. (2013). The local transport structure that supports SDN and which seamlessly interferes with SDN deployment. In this programmable programming structure, a simple transfer node is stripped to a programmable virtual switch that enhances the open flow control protocol. Prototype demonstration of large data applications, programmability and flexibility can greatly help to provide data for large data applications brought by SDN [22].

VIII. Layer Design big data in SDN

Massive data can benefit other layers of SDN as well as traffic architecture, and networks are divided into traditionally different layers, which are used in adjacent layers. In traditional class design, direct contact between non-adjacent layers is prohibited. The speed of recent developments in layered layered designs is that noncontiguous layers share information during the operating period, resulting in new algorithms and significantly improved performance in network systems.

The information in this article can be more than others. Fortunately, large data can be designed across layers in the SDN. The SDN logical central console is achieved by a comprehensive network view, which can be accessed from a high-quality database .Can happen in large data technology and web data management. Performance, cross-layer design in the SDN will be challenging.

IX. Security Attacks and Big data in SDN

When the SDN is separated into multi-target layers, possible malicious attacks can be launched on these three layers of the SDN structure. Jan and others (2015) describe attacks on SDN networks according to potential targets: the application layer attacks, the control layer attacks, and the infrastructure layer hackers. Methods 2 in application layer attacks. The first is to attack some applications; the others to attack the Northern API (API). The console can be seen as one point of network failure risk, so it is a particularly attractive target for attacks in the SDN structure [25].There are many modules that control plane attacks such as attacking a controller, the Northbound API, or the Southbound API. But the attacks of the basic infrastructure plane have two ways. One attack are Routers and the second method is to attack the API south. Some attackers, such as denial of service, benefit from robotic networks and other high-speed Internet access technologies, and attacks have increased dramatically.

X.Wireless Mobile and Big Data

SDN networks for wired networks and tireless mobile networks, such as cellular networks and networks (eg 4G and 5G networks) and sensor networks, were studied. SDN's wireless mobile networks will bring some features or features that occur in each other's applications, expand service requirements and applications for hybrid modeling, mix and difficulties, and improve continuous and transparent network latencies.

XI. Conclusion

This study simulates the rapid development of wireless mobile networks and cloud computing, it brings together more data from mobile devices and networks. Thus, it must be conducted by studying the role and benefits of

SDN for large data applications in portables devices networks. For instance, Smart Applications, IT industry, and IT business. Thus Massive data it has the ability to make very fast or faster progress in many scientific fields and improve the profitability of many companies especially in cloud computing fields. This study showed the general and background characteristics of the big data and related technologies, moreover. The challenges with a focus on the distinctive features of the big data, and discusses multi-processing in this scope.

XII. References

- [1] P. Zikopoulos and C. Eaton, Understanding big data: Analytics for enterprise class hadoop and streaming data: McGraw-Hill Osborne Media, 2011.
- [2] J. Manyika, M. Chui, B. Brown, J. Bughin, R. Dobbs, C. Roxburgh, et al., "Big data: The next frontier for innovation, competition, and productivity," 2011.
- [3] H. Chen, R. H. Chiang, and V. C. Storey, "Business intelligence and analytics: from big data to big impact," MIS quarterly, pp. 1165-1188, 2012.
- [4] R. Margolis, L. Derr, M. Dunn, M. Huerta, J. Larkin, J. Sheehan, et al., "The National Institutes of Health's Big Data to Knowledge (BD2K) initiative: capitalizing on biomedical big data," Journal of the American Medical Informatics Association, vol. 21, pp. 957-958, 2014.
- [5] P. B. Goes, "Editor's comments: big data and IS research," Mis Quarterly, vol. 38, pp. iii-viii, 2014.
- [6] C. Lynch, "Big data: How do your data grow?," Nature, vol. 455, p. 28, 2008.
- [7] D. Agrawal, S. Das, and A. El Abbadi, "Big data and cloud computing: current state and future opportunities," in Proceedings of the 14th International Conference on Extending Database Technology, 2011, pp. 530-533.
- [8] R. L. Krutz and R. D. Vines, Cloud security: A comprehensive guide to secure cloud computing: Wiley Publishing, 2010.
- [9] A. Fox, R. Griffith, A. Joseph, R. Katz, A. Konwinski, G. Lee, et al., "Above the clouds: A berkeley view of cloud computing," Dept. Electrical Eng. and Comput. Sciences, University of California, Berkeley, Rep. UCB/EECS, vol. 28, p. 2009, 2009.
- [10] M. D. Assunção, R. N. Calheiros, S. Bianchi, M. A. Netto, and R. Buyya, "Big Data computing and clouds: Trends and future directions," Journal of Parallel and Distributed Computing, vol. 79, pp. 3-15, 2015.
- [11] A. Fernández, S. del Río, V. López, A. Bawakid, M. J. del Jesus, J. M. Benítez, et al., "Big Data with Cloud Computing: an insight on the computing environment, MapReduce, and programming frameworks," Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, vol. 4, pp. 380-409, 2014.
- [12] A. Beloglazov, J. Abawajy, and R. Buyya, "Energy-aware resource allocation heuristics for efficient management of data centers for cloud computing," Future generation computer systems, vol. 28, pp. 755-768, 2012.
- [13] A. Weiss, "Computing in the clouds," networker, vol. 11, pp. 16-25, 2007.
- [14] D. Agrawal, S. Das, and A. El Abbadi, "Big data and cloud computing: new wine or just new bottles?," Proceedings of the VLDB Endowment, vol. 3, pp. 1647-1648, 2010.
- [15] D. Agrawal, A. El Abbadi, S. Antony, and S. Das, "Data management challenges in cloud computing infrastructures," in International Workshop on Databases in Networked Information Systems, 2010, pp. 1-10.

Acknowledgments

I would think organizations in collection some facts regarding their works and regards to University of Sumer – department of Computer science & Information Technology, to help in publishes my paper.