

A Review on Performance Analysis and Improvement of Internet of Things Application

Ajay Kumar Bharti¹, Rashmi Negi^{2*}, Deepak Kumar Verma³

^{1,2}Department of Computer Science, Maharishi University of Information Technology, Lucknow, India

³Department of Computer Science, JNPG College, University of Lucknow, Lucknow, India

*Corresponding Author: rnsweetrashmi@gmail.com

DOI: <https://doi.org/10.26438/ijcse/v7i2.367371> | Available online at: www.ijcseonline.org

Accepted: 13/Feb/2019, Published: 28/Feb/2019

Abstract— Big data and IoT is two different challenging terms in IT industry. Currently, IoT is an emerging technology and is being used various application for development and research. Big Data Stream Computing (BDSC) is an emerging feature for dealing a real-time data streams and providing faster decisions. BDSC is being used in much of the real time IOT applications. The main objective of the work is to review and measure the performance analysis of real-time IoT application data processing using BDSC platform.

Keywords— IOT, Big Data, BDSC

I. INTRODUCTION

The expression "Internet of Things" which is too in the blink of an eye surely understood as IoT is authored from the two words i.e. the principal word is "Web" and the second word is "Things". The Internet is a worldwide arrangement of interconnected PC organizes that utilization the standard Internet convention suite (TCP/IP) to serve billions of clients around the world. It is a system of systems that comprises of millions of private, open, scholarly, business, and government systems, of nearby to worldwide degree, that are connected by an expansive cluster of electronic, remote and optical systems administration innovations. Today more than 100 nations are connected into trades of information, news and sentiments through Internet. News and suppositions through Internet. While going to the Things that can be any question or individual which can be recognizable by this present reality. Ordinary items incorporate not just electronic gadgets we experience and utilize day by day and innovatively propelled items, for example, gear and contraptions, in any case, "things" that we don't do ordinarily consider as electronic by any stretch of the imagination, for example, nourishment, dress; and furniture; materials, parts and gear, stock and concentrated things; points of interest, landmarks and show-stoppers what not the variety of trade, culture and modernity. These systems produce colossal volumes of information that stream to PCs for investigation. At the point when articles can both sense the environment and impart, they get to be devices for comprehension multifaceted nature and reacting to it quickly.

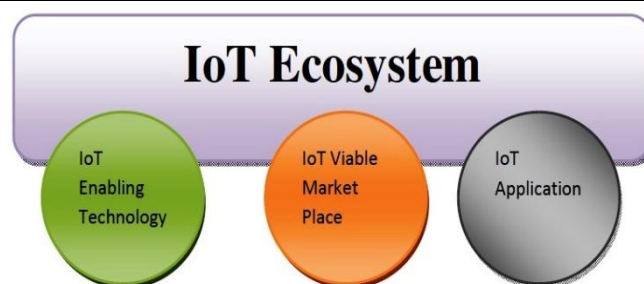


Fig. 1

As per the proposals of the International Telecommunication Union (ITU), the system, Architecture of Internet of Things comprises of:

- (a) The Sensing Layer
- (b) The Access Layer
- (c) The Network Layer
- (d) The Middleware Layer
- (e) The Application Layers

These resemble the Open Systems Interconnection (OSI) reference display in system and information correspondence.

II. BIG DATA TECHNOLOGY

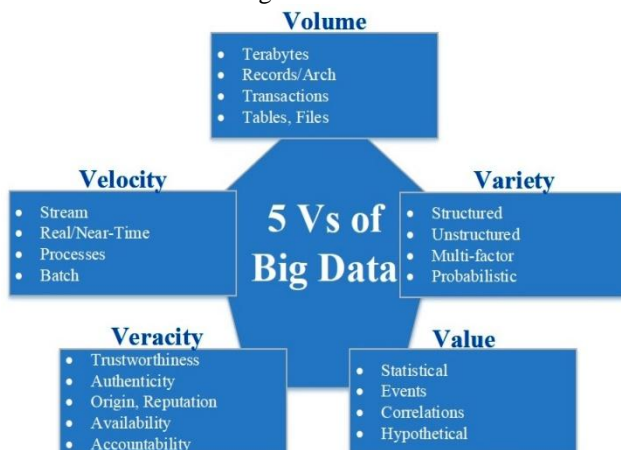
Big Data referred from Mohanty et al (2015) is everywhere massively increasing. In a big data environment, real-time streaming data analytics playing a major role. In general aspect, data processing policies are three types:

- a) Batch processing,
- b) Interactive or online processing, and
- c) Real-Time processing referred from Liu et al (2014).

These policies are handled by the two different computing environment in Big data: (1) Batch Computing referred from Li et al (2015) and (2) Stream Computing referred from Sun et al (2014). In Batch computing, performing the operations on stored offline data dealing with the Hadoop which consists of Map Reduce referred from Bhandarkar (2010) and HDFS. Hadoop is suitable for batch tasks, but it is not capable of handling real-time and online data. Stream processing is a significant feature of performing analytics on real-time online fast streaming data. Both Storm and S4 referred from Neumeyer and Robbins (2010) both the techniques play a vital role in the Stream computing referred from Muthu krishnan (2011) environment. The “big data” term remains problematic to recognize since it can mean so many dissimilar people to access dissimilar things. Find the variation will be differentiation are seen at big data finished to a technology lens, contrasted with an industry lens or business lens. Big Data is technology where operating the huge datasets. In big data technology having four Phases with two computing techniques 5V's Separately (Volume, Velocity, Variety, Veracity and Value) In Figure Shows the overview of the big data phases and their characteristics.

Volume

The volume discusses to the massive quantity of data produced each second. Just think of altogether the photos, video clips, emails, sensor data and Twitter posts that share and produce every second. In this system are not discarding terabytes, but brontobytes or zettabytes of data. On individual Facebook, itself sends 10 billion posts per day, click the upload 350 million new pictures and like button 4.5 billion times each and every day. Taken all the generated information in the world among the start time of the year 2015, it is the equal volume of data now produce each minute! Progressively this makes datasets too big to analyse and store using classical database technology. Through big data technology, nowadays use and stock these datasets with the help of circulated systems, where data fragments to be connected by networks, brought together and stored in different locations through software.



Velocity

Velocity discusses with the phase at which data flows as of sources like machines, networks, business processes and human association with things like cell phones, online networking locales, and so forth. The stream of data is enormous and ceaseless. This continuous information can help organizations and specialists settle on significant choices that give key upper hands and on the off chance that process can deal with the speed. Experts recommend that data analytics can help manage issues like volume and speed.

Variety

Variety discourses to the typical categories the data can nowadays utilize. Before, concentrated on organized information that perfectly fits into social databases or tables, for example, business process information (for example, region wise or product sales). In world's information, as per the statistics, 80% of the current data is unstructured and in this way, can't easily put into relational databases or tables. Uninhabited strategy, the reason of videos, photos, or online updates of the social media. With big data technology, now connect differed types of data including messages, sensor data, social media conversations, photos, voice or video recordings and brought them composed with more traditional form structured data.

Veracity

Veracity discourses to the trustworthiness or messiness of the data. With various big data forms, accuracy and quality are less controllable, for example, abbreviations, Twitter posts with hash tags, colloquial and typos speech. Big data analytics technology now a days allows us to work with these types of datasets. The volumes frequently makeup for the lack of accuracy or quality.

Value

Value discourses to this approach are capable of turning data into towards value. It is important that businesses make a situation for any challenge to leverage and collect big data. It is easy to drop into the embark and buzz trap on big data enterprises without a clear kind of the business value it will take.

III. RELATED WORK

Arkady Zaslavsky et.al, 2013 discussed emerging Internet of Things (IoT) architecture, Internet of Things (IoT) will contain billions of gadgets that can detect, convey, process and conceivably incite. Information streams originating from these gadgets will challenge the conventional ways to deal with information administration and add to the rising worldview of enormous information.

Feng Chen et.al, 2015, have given an efficient approach to survey information mining in learning view, Data mining

includes finding novel, fascinating, and conceivably helpful examples from vast information sets and applying calculations to the extraction of shrouded data. Numerous different terms are utilized for information mining, for instance, learning revelation (mining) in databases, learning extraction, information/design examination, information prehistoric studies, information digging, also, data collecting.

SomayyaMadakamet. Al, 2015, provided an overview of Internet of Things, architectures, one of the popular expressions in the Information Technology is Internet of Things (IoT). What's to come is Internet of Things, which will change this present reality objects into keen virtual items.

Samer M. Barakat , 2016, The late move up to IPV6 web convention which extended the quantity of accessible IP addresses shows that the Internet is moving into a major development as far as the quantity of things and articles planned to be associated over the system, empowering it to wind up distinctly an Internet of Things (IoT). This association might empower the articles and things to share data among them and therefore open the entryway for new web, versatile and remote applications.

Lina Yao and Quan Z. Sheng et.al, 2015 proposed web based management of the Internet of Things. The Internet of Things (IoT) framework introduced here consistently coordinates virtual what's more, physical universes to proficiently oversee things of intrigue (TOIs), where administrations and assets offered by things effectively can be checked, envisioned, what's more, accumulated for esteem included administrations by clients.

Dave Evans, Cisco's chief futurist and chief technologist for Cisco IBSG, 2011 discussed future steps of IoTs as similar as with numerous new ideas, IoTs underlying foundations can be followed back to the Massachusetts Institute of Innovation (MIT), from work at the Auto-ID Center. Established in 1999, this gathering was working in the field of arranged radio recurrence recognizable proof (RFID) and rising detecting advancements.

Jeffrey Cohen et.al, 2009 discussed different analysis practices in Big Data. As enormous information securing and capacity turns out to be progressively adorable, a wide assortment of endeavors is utilizing analysts to take part in complex information investigation. In this paper we highlight the rising routine of Magnetic, Agile, and Deep information investigation as a radical take off from customary Enterprise Data Warehouses and Business Intelligence. We introduce our plan rationality, methods and experience giving this investigation to one of the world's biggest promoting systems at workshop parallel database framework. We depict database plan strategies that bolster the dexterous working style of examiners in these settings. We introduce data

parallel calculations for refined factual techniques, with an attention on thickness strategies.

Hong ZHOU et. al , 2016, In this paper author have discussed IoT applications in the field of agriculture, Coordinating at the web and flow improvement state of the web of things and in view of the accessible innovation investigation of the web of things, the paper makes examination and research on the web of things regarding mechanical levels and frameworks.

Dmitry Namiot, 2015, this paper displays the study of the current indoor situating advancements and frameworks. We examine Wi-Fi what's more, Bluetooth based situating frameworks, light-based and attractive field frameworks. It is our first endeavor to group indoor situating procedures. We attempt to highlight the points of interest and hindrances of different indoor situating advancements and frameworks, think about their exactness, pertinence, working standards and execution parameters.

FaheemZafariet. al, 2015, Smaller scale area is the way toward finding any substance with high precision (conceivably in centimeters), while geo fencing is the way toward making a virtual fence around a supposed Purpose of Interest (PoI). In this paper, we display an understanding into different smaller scale area empowering advancements and administrations. We additionally talk about how these can quicken the fuse of Web of Things (IoT) in shrewd structures.

Vishwajeet H. Bhide, 2014. In the IoT, things are relied upon to end up distinctly dynamic members where they can collaborate and impart among themselves by trading information and data detected about the earth. For that they respond self-ruling to this present reality occasions and furnish administrations with or without direct human mediation.

IV. ANALYSIS OF BDSC ORIENTED IOT APPLICATIONS

In BDSC is become the most efficient and fastest way to get valuable information from what is trendy now, organizations allowing to respond fast when problems look or to notice new trends serving to their performance improvement. BDSC is required to manage the information currently produced at an ever-increasing frequency referred from Sumit Ganguly (2009) from such applications as click streams or log records in blogging, twitter posts, and web exploring.

In detail, all data created can be measured as information gushing or as a preview of information spilling. There are about difficulties that professionals and specialists need to manage in an accompanying couple of years, for example, high adaptability, high consistency, high throughput, high

load adjusting, high adaptation to internal failure et cetera. Those difficulties emerge from the way of the data stream, i.e., arrive information streams at rapid and must be done beneath exceptionally strict limitations of time and space alluded from Albert (2013) Demirkan (2013).

High Fault Tolerance: In-Stream Computing, can permit a framework towards proceeding with appropriately working in case of the disappointment of some of its instruments. Adaptation to internal failure is most looked for after in life-basic frameworks or high-accessibility. In BDSC conditions, it is diverse to achieve high adaptation to non-critical failure, as spilling information is continuous and unending, and encourage altogether, the vast majority of the information are unusable. To achieve high adaptation to internal failure in BDSC situations, a high adaptation to internal failure procedure great and versatile is required. As an adaptation to internal failure make accessible additional assets that permit a request to stay at work after a constituent disappointment denied of interference.

High Throughput: In-Stream Computing, will grow to spill information registering office by running different autonomous outlines of an undertaking topology diagram (DSG) on different processors for the comparable period. In BDSC conditions, it is not at all like finish high throughput, as it is difficult to pick how to order the requiring reiteration sub-graph in information stream diagram, to pick the no. of imitations, and to choose the cut of the information stream to dispense to every copy. To accomplish high throughput in BDSC conditions, great different event procedures are replication is required. Commonly, stacking of the information stream in all instances of all hubs in an information stream graph are proportional is a decent idea, as the figuring ability of all registering hubs is efficient, and it is anything but difficult to accomplish high throughput states.

High Scalability: In-Stream Computing, can enlarge toward support is developing for spilling information and meet the clients QoS, or psychologist to enhance asset use and encourage falling data stream. In BDSC system, it is distinctive to attain high adaptability, as the variety of constant gushing information alluded from Liu et al (2014) is sudden. The critical is that the product changes alongside the information stream graph change, delivers alongside the enhanced methodology, or therapists alongside their use are diminished. This implies versatile projects take up deficient assets and space for littler information requiring, yet can become proficient as more challenges are sited on the continuous information stream alluded from Michael et al (2005). To achieve high adaptability in BDSC conditions, a prudent versatile arrangement engineering, a righteous information stream figuring mode are fundamental and a temperate agent asset assignment methodology.

High Consistency: exceptionally dependable stream processing can enhance framework effectiveness and extend framework security. In BDSC conditions, it is not at all like achieve high unwavering quality, as it is difficult to pick which hubs would be wise to predictable, and which information are required. To finish high unwavering quality in BDSC conditions, an idealistic framework structure is compulsory. Every now and again, the ace slave structure is a decent select, as all information is the primary hub, and it is anything but difficult to achieve exceptionally consistency states.

High Load Balancing: In-Stream Computing, can make an information stream registering framework self-versatile to dodge stack shedding and the varieties of gushing information. In BDSC situations, it is not at all like accomplish high load adjusting, as it is agonizing to assets commitment that spreads top burdens for all day, every day. For the most part, information stream registering groupings utilize shedding of the workload when the limit surpasses their preparing. These administrations an exchange off between affirming that all approaching gushing information are prepared and conveyed a response in the low-idleness state.

Table: Classification of Real-Time BDSC Engines

Property	Storm	Samza	S4
Data Scaling and Partition	Handled by user configuration and coding	Based on the topic partitioning/ message keys	Based on key value pairs
Mobility of Data streams	Pull based, no Blocking operations	Pull Based, data stored at the message broker file storage	Push based
Data Querying	Trident	None	None
HA & Message Processing Securities	Highly available with rollback recovery using an upstream backup. At least once processing.	Highly available with rollback recovery. Data lost when broker failure happens	Highly available with Gap Recovery
Imperfection of data stream handling	User has to implement	Can use stored data at the brokers for	None

		such cases	
Data Storage	None	Data stored at the brokers and can use these for later Processing	None
Non-Deterministic or Deterministic	Doesn't specify	Doesn't specify	Doesn't specify

V. CONCLUSION AND FUTURE SCOPE

The main objective of this paper is Performance evaluation and Improvement of IoT Application Data Processing Using BDSC Platform. We have analyzed different levels of the BDSC Challenges, Real-Time BDSC Techniques, Real-Time BDSC Engines, Real-Time BDSC Platforms, Heterogeneous Data Stream Computing, Resource Scheduling Strategies, Energy-Aware Data Stream Scheduling, Performance Aware Data Stream Scheduling. It shows the gap between the existing BDSC platforms in performance. Our future research plan is to develop an Evolving BDSC platform with the features of high fault tolerance, highly scalability, high throughput and high consistency, for organizing such a system used for real BDSC environment.

REFERENCES

- [1] SWDS Li, J., Bao, Z. and Z. Li, "Modeling Demand Response Capability by Internet Data Centers Processing Batch Computing Jobs", IEEE Trans. on Smart Grid, Vol. 6, No. 2, pp. 737-747, 2015.
- [2] Liu, X., N. Iftikhar and X. Xie, "Survey of Real-Time Processing Systems for Big Data", 18th Int. Database Engineering and Applications Symposium, New York, pp. 356-361, USA, 2014.
- [3] Liu, X., N. Iftikhar and X. Xie, "Survey of Real-Time Processing Systems for Big Data", 18th Int. Database Engineering and Applications Symposium, New York, pp. 356-361, USA, 2014.
- [4] Shao, H., L. Rao, Z. Wang, X. Liu, Z. Wang and K. Ren., "Optimal Load Balancing and Energy Cost Management for Internet Data Centers in Deregulated Electricity Markets", IEEE Trans. Parall. Distr. Syst., Vol. 25, No. 10, pp. 2659-2669, 2014.
- [5] Sharifi, M., S. Shahrivari and H. Salimi, "PASTA: A Power-aware Solution to Scheduling of Precedence-constrained Tasks on Heterogeneous Computing Resources", J. Computing, Vol. 95, No. 1, pp. 67-88, 2013.
- [6] Singh, K. and R. Kaur, "Hadoop: Addressing Challenges of Big Data", 2014 IEEE Int. Advance Computing Conf., Navi Mumbai, pp. 686-689, India, 2014.
- [7] Sinnen, O., A. Tok and M. Khan, "Contention-aware Scheduling with Task Duplication", J. Parall. Distr. Comput., Vol. 71, No. 1, pp. 77-86, 2011.
- [8] Sun, D., G. Fu, X. Liu and H. Zhang, "Optimizing Data Stream Graph for Big Data Stream Computing in Cloud Datacenter

- Environments*", Int. J. of Advancements in Computing Technology, Vol. 6, No. 5, pp. 53-65, 2014.
- [9] K. Parimala, G. Rajkumar, A. Ruba, S. Vijayalakshmi, "Challenges and Opportunities with Big Data", International Journal of Scientific Research in Computer Science and Engineering, Vol.5, Issue.5, pp.16- 20, 2017
- [10] Sun, D., G. Zhang, S. Yang, Zheng W., S. U.Khan and K. Li, "Re-stream: Realtime and Energy-efficient Resource Scheduling in Big Data Stream Computing Environments", Information Sciences, No. 319, pp. 92-112, 2015.
- [11] Mantripatjit Kaur, Anjum Mohd Aslam, "Big Data Analytics on IOT: Challenges, Open Research Issues and Tools", International Journal of Scientific Research in Computer Science and Engineering, Vol.6, Issue.3, pp.81-85, 2018
- [12] V.K. Gujare, P. Malviya, "Big Data Clustering Using Data Mining Technique", International Journal of Scientific Research in Computer Science and Engineering, Vol.5, Issue.2, pp.9-13, 2017.
- [13] Shilpa Manjit Kaur, "BIG Data and Methodology- A review", International Journal of Advanced Research in Computer Science and Software Engineering, Volume 3, Issue 10, October 2013.

Authors Profile

Dr. Ajay Kumar Bharti, MCA, Ph.D., is working as Professor, Computer Science & Engineering at Maharishi University, Lucknow(U.P.) India. His research interest is e-Governance, Service Oriented Architecture and Knowledge Based System. He has published number of research papers in reputed national/international journals and conferences.



Ms. Rashmi Negi, is pursuing M.Tech (CSE) in Department of Computer Science & Engineering at Maharishi University of Information Technology, Lucknow (U.P.) India. Her research interest is IOT and Cloud Computing



Dr. Deepak Kumar Verma, MCA, Ph.D., is working as Assistant Professor, Department of Computer Science, JNPG College, University of Lucknow, Lucknow(U.P.) India. His research interests are IOT, Artificial intelligence and Software Engineering. He has published number of research papers in reputed national/international journals and conferences.

