ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 5, Special Issue 4, February 2018

AMALGAMATION OF BIG DATA AND CLOUD IoT

*** Ms. S. Geetha Rani [1], Mr. V. Sathish[2]**

[1]*Assistant Professor, Department of Computer Science, Thiruthangal Nadar College, Chennai.*

✉ sgeetharani.mca@gmail.com ✆ +91 9566017456

[2]*I M Sc., (CS), Department of Computer Science, Thiruthangal Nadar College, Chennai.*

✉ sathishkumaran170797@gmail.com ✆ +91 9962228410

Abstract – *Big data storage and processing are considered as one of the main applications for cloud computing systems. Furthermore, the development of the Internet of Things (IoT) paradigm has advanced the research on Machine to Machine (M2M) communications. The growing popularity and increasing usage of Internet of Things (IoT) have raised deep concerns regarding the IoT data. IoT is also being characterized as ubiquitous sensing, where the data is continuously generated, especially from sensing systems, which eventually becomes voluminous. Big data, Cloud computing and Internet of Things (IoT), three very different technologies, are already in part of our life. Their massive adoption and use is expected to increase further, making them important components of the Future Internet. In this paper we focus attention on the combination of big data and cloudIoT. To identify research challenges and opportunities on big data analytics to support cloud based information systems.*

KeyWords: Bigdata, CloudIoT, Privacy, Integrity, Security, 5V's, SaaS, PaaS, IaaS,

1. INTRODUCTION

Internet of Things (IoT) provides the sensors that can sense in real-time even while moving. These sensors will produce big data, that is high volume, high variety of data in maximum velocity. So these collected data need to be analysed in order to extract knowledge. Predictive analysis techniques will do the job. A flexible infrastructure, such as the cloud, needs to be used to perform such techniques over big data. In other terms, the cloud binds to the Internet of Things. Cloud computing is expected to play a significant role in IoT paradigm. It provides the promise of big data implementation to small and medium sized businesses.

Big Data processing is performed through a programming paradigm known as MapReduce. Data collection and analysis are essential to any

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 5, Special Issue 4, February 2018

information system. The scale, complexity, capacity of an information system relates closely to available data and the means to process the data.

Amalgamation of Big data and CloudIoT is as shown in the following fig: 1

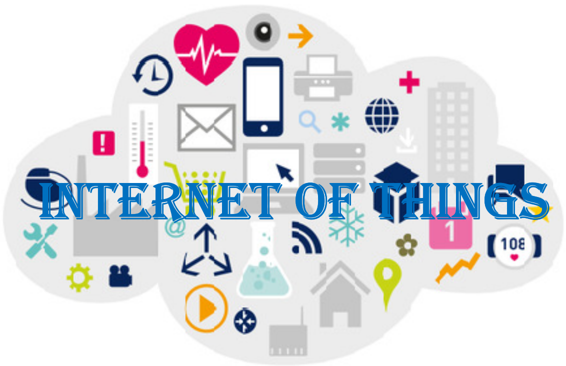

Fig1: Internet of Things

IoT is commonly characterized by factual world and small things with limited storage and processing capacity, and significant issues regarding reliability, performance, security, and privacy. On the other hand, Cloud computing has virtually infinite capabilities in terms of storage and processing power, is a much more advanced technology, and has most of the IoT issues at least partially solved. Thus, two complementary technologies Cloud and IoT fused together (CloudIoT) is expected to disrupt both current and future Internet. The availability of virtually unlimited storage and processing capabilities at low cost enabled the realization of a new computing model, in which virtualized resources can be rented in an on-demand fashion, being provided as general utilities. Huge companies (like Amazon, Google, Flipkart etc.) extensively adopted this paradigm for delivering services over the Internet, gaining both economic and technical benefits. The integration of Cloud and IoT, a promising topic for both research and industry, witnessed by the more recent and rapidly increasing trend dealing with Cloud and IoT together (Fig:2)

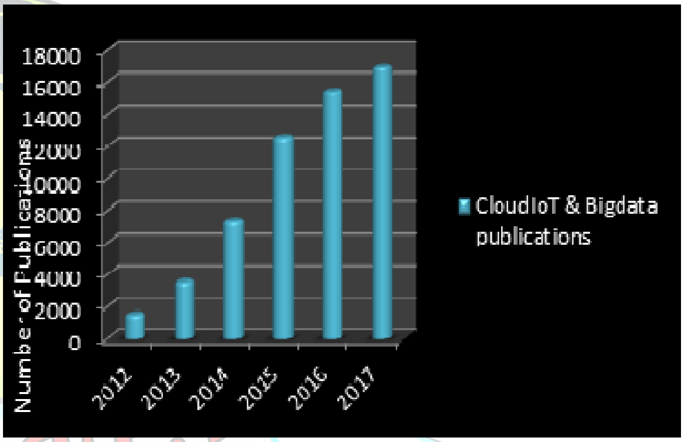

Fig: 2 According to Google scholar trends Internet & Research trends about CloudIoI and Big data

An enormous number of application opportunities are emerging in different scenarios, such as cities [smart lighting, air quality], smart metering [water flow, water leakages), emergency [radiation levels, fire exposure, access control], retail [smart payment, item location], industry [process monitoring, safety], agriculture [green houses, smart irrigation], home automation [energy and water use, remote control], and healthcare [patient observation]. [2] discussed about a method, Optimality results are presented

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 5, Special Issue 4, February 2018

for an end-to-end inference approach to correct(i.e., diagnose and repair) probabilistic network faults at minimum expected cost. One motivating application of using this end-to-end inference approach is an externally managed overlay network, where we cannot directly access and monitor nodes that are independently operated by different administrative domains, but instead we must infer failures via end to-end measurements. We show that first checking the node that is most likely faulty or has the least checking cost does not necessarily minimize the expected cost of correcting all faulty nodes. In view of this, we construct a potential function for identifying the candidate nodes, one of which should be first checked by an optimal strategy. Due to the difficulty of finding the best node from the set of candidate nodes, we propose several efficient heuristics that are suitable for correcting fault nodes in large-scale overlay networks. We show that the candidate node with the highest potential is actually the best node in at least 95% of time, and that checking first the candidate nodes can reduce the cost of correcting faulty nodes as compared to checking first the most likely faulty nodes.

2. PROBLEM ANALYSIS AND SYSTEM DESIGN

2.1 Architecture of BDCloudIoT

IoT can benefit from the resources and virtually unlimited capabilities of Cloud to compensate its technological constraints (e.g., storage, processing, and communication). To mention a few examples, Cloud can offer an effective solution for IoT service management as well as for implementing applications and services that exploit things or data produced by them. On the other hand, Cloud can benefit from IoT by lengthening its scope to deal with real world things in a more distributed and dynamic manner, and for distributing new services in a large number of truthful life scenarios.

The following fig 3 shows the Architecture of Big data and CloudIoT.

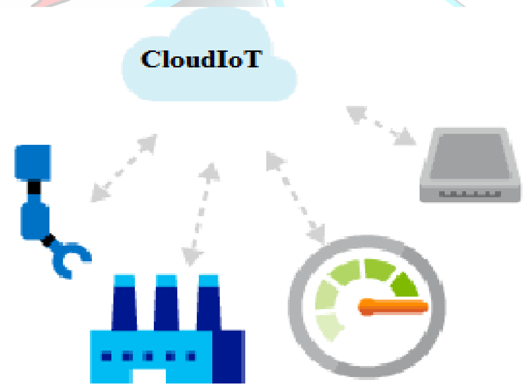

Fig 3: Architecture of Bigdata and CloudIoT

2.2 Five V's of Big data:

Big data is defined based on some of its characteristics. It doesn't mean the size. There are five characteristics that can be used to define big

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 5, Special Issue 4, February 2018

data, as also known as 5V's.There are many properties associated with big data. The prominent aspects are volume, variety, velocity, variability and value.

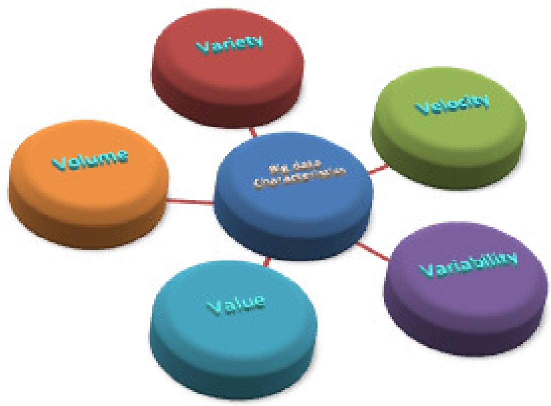

Fig: 5 V's of Big data characteristics

- **Volume:** It relates to size of the data such as TeraBytes, PetaBytes, ZettaBytes etc., Many factors contribute for the increase in volume like storage of data, live streaming etc.,
- **Variety:** It means the type of data. In addition, difference sources will produce big data such as sensors, devices, actuators, social networks, the web, mobile phones, healthcare monitors, etc.,
- **Velocity:** It means how repeatedly the data is generated (e.g. every millisecond, second, minute, hour, day, week, month, year). Velocity is the speed at which the files are created and processed.
- **Variability:** It describes the amount of variance used in summaries kept within the data bank and refers how they are spread out or closely grouped within the data set.
- **Value:** All enterprises and e-commerce systems are strong in improving the customer relationship by providing value added services.

2.3 Technologies around Big Data, CloudIoT

The three technologies are more favoured to be used commonly across in big data, cloud and IoT applications. Popular key-value based data storage products are Dynamo aws.amazon.com/dynamodb), Redis (redis.io), Riak (basho.com/Riak), Amazon SimpleDB (aws.amazon.com/simpledb), and Windows Azure Table Storage (windowsazure.com). Popular document store technology based products are CouchDB (couchdb.apache.org), and MongoDB (mongodb.org). Popular wide column based products are Apache HBase (hbase.apache.org), Hadoop (hadoop.apache.org), and Cassandra (cassandra.apache.org). However, the one getting the most attention is Hadoop as it based on popular distributed processing friendly MapReduce technology. However, there is no single perfect data management solution for the cloud to manage big data. In addition, Future-Grid (http://FutureGrid.org) and Open- Stack provide software stack definitions for cloud data centers. The techniques in big data need to be developed to tackle the challenges such as extracting, transformation, integrating, sorting and manipulating data. Basic techniques of meaningful data extraction from big data consists five steps define, search, transform, entity resolution, answer

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 5, Special Issue 4, February 2018

the query. Arduino as sensor platform, Google App Engine (appengine.google.com) as cloud platform, and Google Data store as the database to store sensor data has been employed in the experiments. Connecting and managing sensor via cloud is a critical milestone in the process of sharing sensors and sensor data in Sensing as a service model.

2.4 Service model in cloudIoT:

IoT envision that sensors to be attached everywhere. In such an environment, owner still be able publish their sensor data and earn a fee or discount as a return. To visualize in this concept in real world let's consider a scenario. The Objective of IoT is to provide utility based services such as Sensing as a service (SaaS), Location as a Service (LaaS), and Traceability as a Service (TaaS).

Sensing as a service (SaaS) at personal level where homogeneous processes could be applied at private and public organization levels. Owners of sensors (could be a person, private organization, public organization, government, etc.,) will be able to publish their sensor data and get a return on investment. Privacy and security need to be handled appropriately.

Cloud computing consists of three main layers or model, Infrastructure as a Service (IaaS), Platform as a Service (PaaS), and Software as a Service (SaaS). In addition to the above main layer, some other layers are also introduced and discussed in literature such as Database as a Service (DBaaS), Data as a Service (DaaS), Ethernet as a Service (EaaS), Network as a Service (NaaS), Identity and Policy Management as a Service (IPMaaS), and Sensing as a service (SaaS). In general, all these models are called XaaS, which means 'X' can be virtually anything.

3. CHALLENGES IN CloudIoT:

We have discussed how incorporating Cloud and IoT provides several benefits and encourage the birth or improvement of a number of interesting applications. At the same time, we have seen that the complex CloudIoT situation imposes several challenges for each application that remains currently receiving attention by the research community. We then focus on other important frequent challenging topics strictly related to CloudIoT.

3.1 Security and privacy:

When critical IoT applications move towards the Cloud, concerns arise due to the lack of, e.g., trust in the service provider, knowledge about service level agreements (SLAs), and knowledge about the physical location of data. Multi-tenancy can also compromise security and lead to sensitive information leakage. Moreover, public key cryptography cannot be applied at all layers due to the computing power constraints imposed by the things. These are examples of topics that are currently under investigation in

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 5, Special Issue 4, February 2018

order to tackle the big challenge of security and privacy in CloudIoT.

3.2 Heterogeneity:

A big challenge in CloudIoT is related to the wide heterogeneity of devices, operating systems, platforms, and services available and possibly used for new or improved applications. To deal with data heterogeneity, we must address the following issues properly: How to improve quality (accuracy) of data in real time; how to unify data representation and processing models to accommodate heterogeneous or new types of data; how to improve intelligent data interpretation and semantic interoperability; how to implement inter-situation analysis and prediction; how to implement knowledge creation and reasoning; and how to conduct short-term and long-term storage.

3.3 Smart Energy and Smart Grid:

IoT and Cloud can be effectively merged to provide intelligent management of energy distribution and consumption in both local and wide area heterogeneous environments. The combination of Cloud and IoT increases the concerns about security and privacy issues for Smart Grid software deployment for utilities. These concerns should be adequately addressed to realize the full potential of such application: consumers should gain more confidence in sharing data to help improving and optimizing services offered

3.4 Smart City:

IoT can yield a common middleware for future oriented Smart City services, obtaining information from different heterogeneous sensing infrastructures, retrieving all kinds of geo location and IoT technologies through RFID sensors and geo tagging, and exposing information in a uniform way. A number of recently proposed solutions suggest using Cloud architectures to enable the discovery, connection, and integration of sensors and actuators, thus creating platforms able to establish and support pervasive connectivity and real time applications for smart cities.

4. CONCLUSION:

IoT will be one of the important sources of Cloud and Big data. Cloud will enable to store it for long time and to perform complex analyses on it. The amalgamation of Cloud Computing, Big data and Internet of Things represents the next big leap ahead in the future internet. The new applications arising from this combination, we call it as CloudIoT. In this paper, we surveyed the literature in order to recognize the balancing aspects of Cloud and IoT and the main drivers for incorporating them into a unique environment.

5. BIBILIOGRAPHY:

1. Rupali A. Meshram, Komal R. Hole, Rutuja A. Gulhane, Pranita P. Deshmukh, Yugandhara A. Thakare, and Meghna A. Deshmukh (APRIL 2017), *" Internet of Things: Recent Applications and Challenges",* International Journal of

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 5, Special Issue 4, February 2018

Engineering Science and Computing, Volume 7 Issue 4, Pages 10679 to 1068,

2. Christo Ananth, Mona, Kamali, Kausalya, Muthulakshmi, P.Arthy, "Efficient Cost Correction of Faulty Overlay nodes", International Journal of Advanced Research in Management, Architecture, Technology and Engineering (IJARMATE), Volume 1,Issue 1, August 2015,pp:26-28.

3. Qi Jing , Athanasios V. Vasilakos , Jiafu Wan , Jingwei Lu, Dechao Qiu,(2014 JUNE), "Security of the Internet of Things: perspectives and challenges", Journal of Wireless Networks, Volume 20 Issue 8,Pages 2481-2501, Springer-Verlag New York, Inc.