Convergence of Internet of things, Cloud Computing, Big Data and Security

Hitisha Damani
Department of Information Technology
Dwarkadas J. Sanghvi College of Engineering,
Mumbai-400056, Maharashtra-India

Rajavi Mehta
Department of Information Technology
Dwarkadas J. Sanghvi College of Engineering,
Mumbai-400056, Maharashtra-India

Neha Katre
Department of Information Technology
Dwarkadas J. Sanghvi College of Engineering,
Mumbai-400056, Maharashtra-India

ABSTRACT
There are many technologies that are developed and used independently and they are efficient in their performance too. But if these technologies are merged, then it may increase the overall efficiency of the merged technologies and provide benefit to mankind. This paper describes three such technologies which are: Internet of Things, Big Data and Cloud Computing and later attempts to hypothetically combine them as one and try to cover loopholes by stating ways and methods for secured movement and storage of data from one phase to another.

General Terms
Convergence, encryption, attack

Keywords
Internet of things, Cloud Computing, Big Data and Cyber Security

1. INTRODUCTION
Big data, the Internet of Things (IoT) and cloud computing are distinct technological categories that have developed independently over a large time period, but in this process of developing, they are increasingly becoming interdependent**,** and thus, are displaying their ability to revolutionize the fields of health, technology and other economic sectors.

Cloud computing is an information technology paradigm that enables ubiquitous access to shared pools of configurable system resources and higher-level services that can be rapidly provisioned with minimal management effort, often over the Internet. Cloud computing is more like a combination of various technologies rather than being a single discrete technology [1]. In other words, cloud computing makes all the resources available to the user over the Internet and within no time. Though cloud computing has been in trend since recent times, it is no new concept. The first cloud was developed about ten years ago. Even while one is sending mails, using Google

drives or surfing on network, he is unknowingly at the base level using cloud. The idea of cloud computing is now widespread because of the fundamental advantages of its architecture like cost saving, high performance, shared resources, ease of use and rapid elasticity. And, now after these ten years it is gradually merging with Internet of Things.

Internet is a phenomenon that is spread over the whole world. Anything from a small object to a whole city can be connected to each other via the Internet. Every corner of the globe is connected to each other via the Internet. An upcoming advancement in this field is the **Internet of Things** which is to connect any object via the internet and get information from its surrounding through sensors and actuators.

In this era of digitalization, every possible action is getting digitized and therefore a huge amount of data is getting generated. One needs to store this data to make it useful by making it accessible to the users for future use. Big Data analytics is used in such situations.

Big Data is a collection of huge amount of data that is continuously generated which may be structured, semi structured or unstructured. Data can be as much as in Terabytes or Petabytes or even exabyte which are collected over a period of time [4].

Big Data is described by the 3Vs [5] as:

Volume: describes the huge amount of volume of data which is collected. Data can be of any format, be it text, video, image etc. Therefore, as data increases in size, different methods and techniques are required to store it.

Variety: It is the cast variety or the different kinds of data types used by different sources as data can be stored as a video or image or even a simple text format.

Velocity: It is the speed or the velocity at which the collected data is analyzed to be able to retrieve the required data. It also determines the velocity at which data is changing as data of yesterday is no longer called recent data in these times of digitalization.

Any technology ever developed always requires security. The user wants his/her data or information to be secure and safe. **Computer security** is the process of securing your computer, software and hardware, against threats, misuse, and disruption of data. There are two parameters to computer security: - cyber and physical security. While cybersecurity deals with the technical threats to the data like hacking, malware, etc., physical security deals with security against fire, breakage, natural disasters, etc. The three important components of physical security are access control, surveillance and testing. For control against attacks to surveillance, physical locations should be monitored using surveillance cameras and other systems, such as intrusion detection sensors, heat sensors and smoke detectors should be installed. We can reinforce disaster recovery policies and procedures that must be tested on a regular basis to ensure safety [7].

Having looked at the basics of Cloud computing, IoT and Big Data, we will now look at the architectures of each of these technologies. Section II describes the architecture of Cloud Computing. Section III throws light on the architecture of Internet of Things and Section IV discusses the architecture of Bigdata. Section

V discusses the role of security in these 3 technologies and Section VI discusses the convergence of these technologies. Applications are discussed in section VII and the conclusion is presented in section VIII.

2. ARCHITECTURE OF CLOUD COMPUTING

The cost required to open a new firm with a large amount of hardware and software resources is extremely high. Each time new software or hardware is made; all installations have to be done again. Here, is where the cloud comes to rescue. Cloud computing architecture has the primary advantage of resource sharing. Figure 1 shows the architecture of Cloud Computing.

Cloud Computing architecture comprises of many loosely coupled cloud components. The following are the four most important and broadly classified components of a cloud:-

1. front end platform,
2. back end platforms ,
3. a cloud based delivery, and
4. a network.

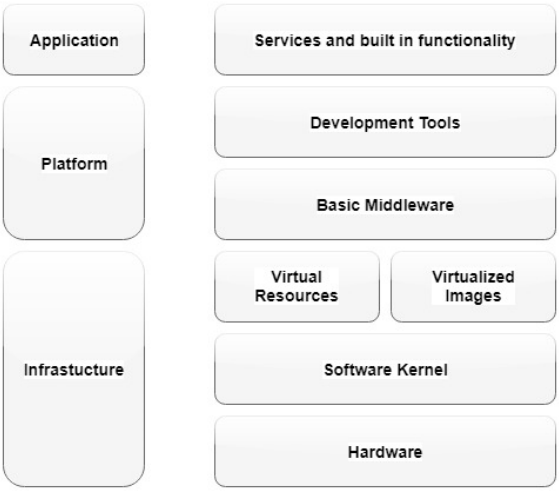

Figure 1: Architecture of Cloud Computing

The front end refers to the client part of cloud computing system. These clients can be classified as servers, fat clients, thin clients, zero clients, tablets and mobile devices. These client platforms interact with the cloud data storage via an application middleware, web browser, or a virtual session [1]. The front end also includes interfaces and applications, like Web Browser, that are required to access the cloud computing platforms [8].

The back end of the cloud architecture refers to the cloud itself. It consists of the resources that are required to provide cloud computing services to the user. These resources include huge data storage, virtual machines, security mechanism, deployment models etc. The main purpose of the back end section is providing built-in security mechanism, managing protocols and traffic control.

The various cloud based services have their own cloud architectures which can be uniquely classified as:

1. Cloud Delivery Models:-
   - Software as a Service (SaaS)
   - Platform as a Service (PaaS)
   - Infrastructure as a Service (IaaS)
2. Cloud Deployment Models
   - Private Cloud
   - Public Cloud
   - Community Cloud
   - Hybrid Cloud[9].

3. ARCHITECTURE OF INTERNET OF THINGS

How about getting a notification from your oven when your food is heated and is ready to be taken out. In such situations, IoT is used. Basically any device/ machine/ equipment that has an on and off switch can be connected to the internet via sensors is known as Internet of Things. Here internet means the network to which the object is connected and things means literally anything with a switch. Connection can be through sensors, wireless connection and to anything which can assigned an IP address through which data can be transferred. By connecting to the internet that object is connected to other objects too that are connected in the network, hence, there is object-object connection too [11].

Another application for IoT is Smart cities which can help us reduce waste and improve efficiency for things such as energy use which will helping us to understand and improve how we work and live.

According to most of the researchers, IoT consists of the following layers as shown in Figure 2:

1. Perception layer: it is the physical layer consisting of sensors to sense and gather information from the surroundings of the object. It also contains the network connectivity to transfer data to the network layer from the physical layer.
2. Network layer: It is responsible for connecting one object to other smart objects, network devices, and servers, it is responsible for routing the data coming from the sensors to the next layer which is the **Management Service Layer**. Its features are also used for storing and processing sensor data. Hence, it is needed for it to have a big storage capacity to store all the data.
3. Perception layer: it is the physical layer consisting of sensors to sense and gather information from the surroundings of the object. It also contains the network connectivity to transfer data to the network layer from the physical layer.
4. Network layer: It is responsible for connecting one object to other smart objects, network devices, and servers; it is responsible for routing the data coming from the sensors to the next layer which is the **Management Service Layer**. Its features are also used for storing and processing sensor data. Hence, it is needed for it to have a big storage capacity to store all the data.

   There are a few ways for the device to connect back to the larger internet. First is to connect the devices in a peer-to-peer network where each device talks to the other and eventually one device is close enough to the larger network to “tap in”. A second way is a

hub and spoke system where devices talk back to a centralized hub. Finally, devices can be connected directly to the internet via wifi or LTE or a direct connection.

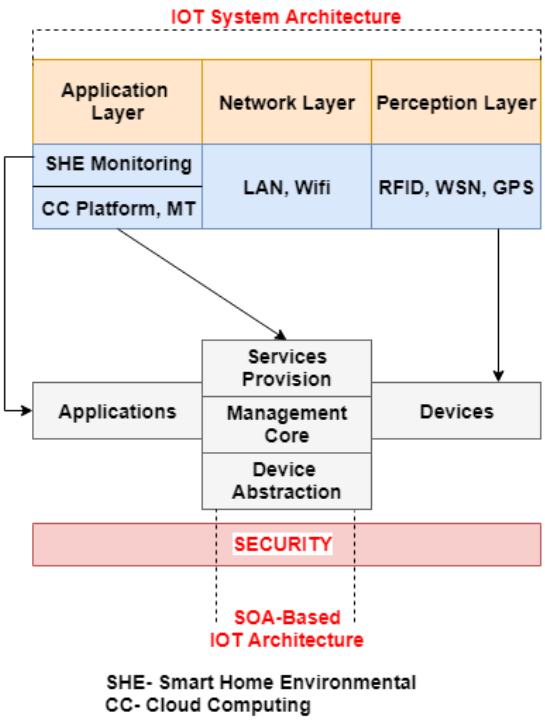

Figure 2: Architecture of IoT

5. Application layer: It is the topmost layer in the IoT architecture. It is the interface between the user and the software.

4. ARCHITECTURE OF BIG DATA

There are four layers in the big data architecture. They are as follows:

1. The source: This defines the source of the data, i.e., from where it is coming. It can be real time data or batch mode. It can come from data warehouses, relational databases, social media, emails, sensors, etc. It can come through sensors or company servers or from third- party data providers.
2. Data massaging and storage layer: This layer receives the data and it can convert it into structured data using analytical tools if necessary and store it. Fast data ingress requires connectors and adapters that can efficiently connect to different storage systems, protocols, and networks [12].
3. Analysis layer: This layer defines the processing of data and using it for business intelligence. The structured data can be processed using technologies like sampling whereas unstructured may need newer and better forms of technologies for processing pertaining to its complexity.
4. Consumption layer: This layer retrieves the data and displays it to the appropriate output layer [13].

5. CONVERGENCE OF INTERNET OF THINGS, BIG DATA, CLOUD COMPUTING AND SECURITY

The demand for these technologies, viz., cloud computing, big data and IoT have increased tremendously in the past decade. These technologies are no longer a privilege, they are a requirement. As the need increases, these technologies are merging into one another as depicted in Figure 3.

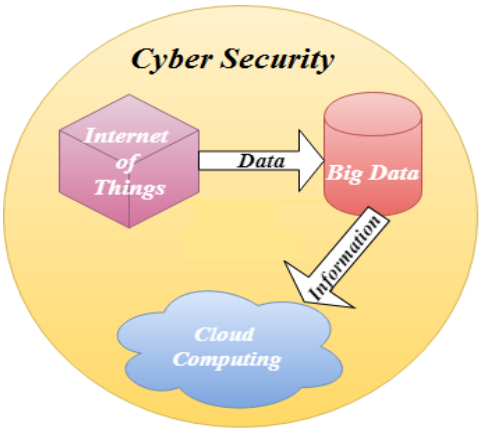

Figure 3: Convergence of IoT, Big Data and Cloud Computing

For example, IoT helps find large amount of data, but this data needs to be stored and used. Thus, cloud computing comes into the picture. A large amount of data is of no use unless it provides useful information. Big Data processes this data and provides useful information and patterns to make data interpretation easier. This information is made available to the users through cloud.

In today's world only about 0.5% [14] of the total data available is being used. All this data can be put to use and made accessible for the consumers via the convergence of these technologies. As the amount of data increases, it means that more people are willing to store their data over the system. A user would always want to protect his/her data from any threats. In the architecture for Convergence of the IoT, Big Data and Cloud Computing technologies, the role of security as a whole is to ensure that the data and information is transferred from IoT to Big Data and then the information generated by Big Data Analysis to Cloud securely and without adulteration by any malicious activities. It must also ensure that the data while in the cloud is encrypted and secure and no unauthorized person should reach it. Information Security ensures that the documents, irrespective of their format, content, size, etc., are secure from unauthorized access. Cyber Security ensures the safety of the hardware, software, and the network from malicious effects which may occur due to the viruses, network access, data and code injection. Many of the threats to the network include denial of service attack, spoofing, packet sniffing, hijacking and clickjacking.

It is always important to go through the flaws of any technology before its benefits. The first step is extracting the data using IoT. Thus, we need to see the security threats to IoT. The major cyber security asset in IoT is the data. The data must be from an authenticated source and it must not be available to an adversary and the channel must be a secure one.

1. **Data Authentication**: - The security of the data must never be compromised. Even if the data is encrypted after collecting it, the hacker can attack it from the source site or the destination site. The access to the data must be made only by

authenticated users. Both the source and destination sites must be authenticated.

2. **Data Encryption**: - The data collected is open to attacks, if not stored securely. The data must be encrypted to store it safely and away from easy access by the attackers. One can use the Secure Socket Layer Protocol (SSL) or any other such protocols to encrypt and secure the data. This will only secure the stored data. The data transmitted through wireless mediums must also be encrypted using public and private key algorithms.
3. **Side-Channel attacks**:- These attacks put less focus on what the information is and more on how it is presented. Such attacks include captcha attacks,timing attack, power-monitoring attack, electromagnetic attack, acoustic cryptanalysis, differential fault analysis, data remanence and software-initiated fault attacks [14].

Now, the next step is storing the data in Big Data and then extracting information from it. Since Big Data stores a large amount of data which is processed to infer useful information, the stored data is an asset to Big Data. The attacker has to go through various layers before getting the result but once the result is obtained, it is a one time struggle compared to the long time gains. Similarly, the organizations face a huge lose in case of an attack. Thus, securing this data is very important. The following measures must be taking

1. **Protecting transaction logs and data: -** While the data is being processed it is transferred through various levels. The IT manager must keep a track of what is being transferred and how.
2. **Validation and filtration of end points inputs: -** Even though in this architecture, the data is taken from IoT and stored in cloud, the authentication of the input and output points must be validated. The data must be sent to and taken from authenticated sources only.
3. **Data Provenance: -** To process the data it is very important to first know the origin and source of the data and thus to have a correct output and information.
4. **Securing and protecting data in real time:-** Since there is is a large amount of data, it is difficult, but important, for the organization to keep a check on this data for every second possible. The attack on the data can occur anywhere and anytime.
5. **Protecting access control methods and encryption: -** The data must be secured by having an efficient access control device and encrypting all the data using cryptography techniques [15].

After receiving the data and processing it to retrieve patterns and information from it, there is no use of storing such a huge data if it is not put to better use. But one cannot share the data with everyone as duplicate (and sometimes outdated too) copies are created and this leads to wastage of storage area. Moreover, getting lots of data which is of no use is waste of money. Thus, it is better to have data as per usage and pay only for relevant data. Hence, Cloud Computing is used to share the data with the users. But this sharing of data depends upon the kind of user and what kind of data is being shared. If the data is confidential to a restricted number of people, only those people should be able to access it. The following are the threats and cares that must be taken:-

1. **Privileged Access**: - The type of cloud defines the restricted access allowed for the information stored in that cloud. The private cloud data must not be available for public use. We can use various method to protect data against access control like biometrics, password authentication, restricted access, etc.
2. **Data Breaches: -** Data loss and breach is a serious threat to cloud computing. Since a large amount of data is stored, the data must be made available to only the authorized users.
3. **Account Or Service Traffic Hijacking: -** Hijacking is a major issue related to any technology. Special care must be taken at this level too that the data is accessed only by the valid users and no hijacking takes place. Strict rules for security must be adopted. For account hijacking, care like biometrics testing must take place. For example, the retina of the user's eye can be scanned. In case of a breach the account must be locked until further verification.
4. **Insecure APIs:-** The user application programming interface, i.e. where the required data is displayed, must be verified and must be accepted if and only if valid, else must be blacklisted. This will ensure information is available only for valid users [16].

6. APPLICATIONS

The basis of all applications of the convergence of Big Data, IoT and Cloud computing is the availability of resources in real time, i.e., get upto date and not use outdated data for any analysis [17].

Some applications are:

1. **Betterment of the Healthcare sector:** Even today, many researches in the healthcare sector is done on the hypothesis-driven research. But after the convergence of these technologies one can perform data-driven researches which can help one to draw to better and realistic conclusions rather than anticipating the results. Also, personal data that is retrieved from sensors can be stored and used to diagnose the diseases remotely and and a better understanding of disease and development of innovative solutions for therapeutics can be easily achieved.
2. **Return Of Investment (ROI):** Companies use data analytics to get information from the raw data to get the best results for their analysis. The cloud helps them to store huge amounts of data and that too at a lesser cost. With increasing data, a safe storage platform, and analysis of such data, the ROI of businesses will become very high.
3. **Industrial internet of things (IIoT) will be benefited**: People will become less dependent on IT sector as most of the functions can be handled with data integration and automation. The BI tools will increasingly become simple and self-sufficient to do basic functions. Analytics as a service will become more common.

7. CONCLUSION

Internet of Thing, Big Data and Cloud Computing are beautiful concepts which if used appropriately and interlinked with one another will benefit the world enormously as a whole provided that the mechanism is

secure. These technologies have developed marvelously over the years but it is safe to say that they are still in the rudimentary stage of development and have a long way to go. Even though the convergence of these technologies is implied, until and unless they are not secured against cyber threats, they are of no use to the mankind. Security at an individual level and also at the complete level will be necessary as security is the heart of these technologies. Therefore, it is very important to secure them since in today's world there is no bigger threat than loss and no bigger offence as breach of personal information.

8. REFERENCES

[1] En.wikipedia.org. (2018). Cloud computing. Available: https://en.wikipedia.org/wiki/Cloud_computing

[2] targroup.uwaterloo.ca. (2018). Available: http://www.stargroup.uwaterloo.ca/~mhamdaqa/publications/Cloud_Computing_Uncovered.pdf

[3] Journal of Electrical and Computer Engineering Volume 2017, Article ID 9324035, 25 pages Available: https://www.hindawi.com/journals/jece/2017/932403/

[4] SearchBusinessAnalytics. (2018). What is big data analytics? - Definition from WhatIs.com. Available: http://searchbusinessanalytics.techtarget.com/definition/big-data-analytics.

[5] Blog.sqlauthority.com. (2018). Available: https://blog.sqlauthority.com/2013/10/02/big-data-what-is-big-data-3-vs-of-big-data-volume-velocity-and-variety-day-2-of-21/

[6] SearchSecurity. (2018). What is cybersecurity? - Definition from WhatIs.com. Available: http://whatis.techtarget.com/definition/cybersecurity

[7] www.tutorialspoint.com. (2018). Cloud Computing Architecture.Available: https://www.tutorialspoint.com/cloud_computing/cloud_computing_architecture.htm

[8] En.wikipedia.org. (2018). Cloud computing architecture. Available: https://en.wikipedia.org/wiki/Cloud_computing_architecture

[9] Krutz, R. and Vines, R. (2010). Cloud security. Indianapolis, Ind.: Wiley Pub. Chapter 3

[10] Bush, B. (2018). The Architecture of the Internet of Things. By.dialexa.com. Available: https://by.dialexa.com/the-architecture-of-the-internet-of-things

[11] Forbes.com.(2018).Available: https://www.forbes.com/sites/jacobmorgan/2014/05/13/simple-explanation-internet-things-that-anyone-can-understand/#1485c7571d09

[12] Pal, S. (2018). How to Design a Big Data Architecture in 6 Easy Steps - Saama. Available: https://www.saama.com/blog/design-big-data-architecture-6-easy-steps/

[13] Datamation.com. (2018). What is Big Data Architecture? - Datamation.Available:https://www.datamation.com/big-data/big-data-architecture.html.

[14] Burn-Murdoch, J. (2018). Study: less than 1% of the world's data is analysed, over 80% is unprotected. the Guardian.Available: https://www.theguardian.com/news/datablog/2012/dec/19/big-data-study-digital-universe-global-volume

[15] Internet Of Things Wiki. (2018). IoT Security-Issues, Challenges and Solutions - Internet Of Things Wiki. Available: https://internetofthingswiki.com/iot-security-issues-challenges-and-solutions/937/

[16] Buttler, P., McNulty, E., Davis, M., McDonald, M., McNulty, E., Davis, M., McDonald, M., McNulty, E., Davis, M. and McDonald, M. (2018). 10 Challenges to Big Data Security and Privacy - Dataconomy. Dataconomy. Available: http://dataconomy.com/2017/07/10-challenges-big-data-security-privacy/

[17] InformationWeek. (2018). 9 Worst Cloud Security Threats - InformationWeek. Available: https://www.informationweek.com/cloud/infrastructure-as-a-service/9-worst-cloud-security-threats/d/d-id/1114085

[18] 9 Worst Cloud Security Threats - InformationWeek", InformationWeek, 2018. Available: https://www.informationweek.com/cloud/infrastructure-as-a-service/9-worst-cloud-security-threats/d/d-id/1114085.