IOT as enablers of M2M communication in industries

[1]Hardik Dilip Thakkar, [2]Srijita Bhattacharjee, [3]Dr Ashok Kanthe

[1]Student, [2]Assistant Professor, [3]HOD
Computer Engineering,
Pillai Hoc College of Engineering and Technology, Rasayni, India

***Abstract*: An In this ever changing world of technology IoT (Internet of Things) and cloud services have become a leading benchmark of innovation and optimizing existing services. Internet of things is itself based on the sole idea of transforming every day to day device into a virtual source of information which can derive huge data sets and can help ourselves to advance as a Human race. This concept has indeed impacted the manufacturing and retail industry. Each day huge industries and their technical support team is devising new ways to exploit Internet of things to optimize their production and to minimize operations cost leading to more profits and better wages to the laborers. But if you consider huge number of machines at every local production center churning out KBs of message per second and then these messages are to be processed and transformed to derive live snapshot to managers and support units to take better decision and troubleshoot, then it is sure that the bandwidth and latency would be some of the exclusive parameters to ensure real time delivery and processing of data. In this paper we would looking at some of the technologies and framework which would help us to better utilize the bandwidth of the network in use and minimize the latency at the same time. The aim is to be comparing the methods and derive conclusive evidence on which method is better than the others.**

***Keywords*: Fragment, Uncertain Traffic Matrix, Storage.**

I. INTRODUCTION

IOT is where information proprietors have an advantageous, dependable, on-request access and it joins the assets, for example, servers, stockpiling and applications over the web. Clients don't have a control of basic equipment framework that is claimed and overseen by the supplier. They get to the administrations or designated assets by utilizing an internet browser. IOT gives reasonable, on demand arrange which can be revealed with basic administration report or gives best administration association. It tends to be provisioned instantly to consolidated assets, for example, stockpiling, applications, administrations, systems and servers. The five noteworthy highlights of a cloud are:

- **On-Demand Relevance:** An information proprietor can naturally set up abilities which can be processed, for example, Stockpiling of system and time of server as required without discussing specifically with each Service Provider.
- **Broad Network Access:** Services are accessible over the system that can be gotten to from anyplace whenever through Ordinary components that are gotten to by different gadgets, for example, tablets, cell phones, workstations or PCs.
- **Asset Pooling:** The cloud specialist organization's estimating assets, for example, transmission capacity of system, memory, and preparing and distributed storage are pooled by utilizing a multi-holder model to serve number of information proprietors. Further, contingent upon information proprietor's interest, different physical and virtual assets are appointed and re-relegated legitimately. Additionally, the feeling of area autonomy is accessible. The information proprietors are ignorant about the given assets total area however they can decide the physical area at an abnormal state, for instance datacenter, state or nation.
- **Fast Flexibility:** Cloud registering abilities can be provisioned legitimately to full fill the interest and load necessities of information proprietors. The capacities regularly appear to be boundless which is accessible for provisioning to the information proprietors. In required amount and whenever these capacities can be doled out.
- **Methodical Service:** By utilizing capacities, at some dimension of reflection, cloud frameworks can screen and enhance the utilization of assets consequently. This ability is adjusted for the sort of administration, for example, transmission capacity, stockpiling, preparing and dynamic client accounts. For both the specialist organization and the purchaser giving straightforwardness asset utilization can be controlled, stamped and watched.

In spite of the fact that distributed computing can aggregate huge information, it neglects to give organize data transfer capacity and spare circle space since at least two customers transfer a similar record. To defeat this issue, a notable information pressure procedure named as reduplication is utilized. In information reduplication, instead of putting away number of information duplicates having same substance, server evacuates copy information duplicates by putting away just a single physical duplicate and alluding other information duplicates to the as of now put away information duplicate.

II. LITERATURE REVIEW

In this section the review of some concepts and learning approaches on Process Fragments with Uncertain Traffic Matrix and the Ensemble learning models from the referred sources are presented.

Shoulu Hou, Shuai Zhao described the issue of business process model fragmentation has received increasing attention in the past few years, especially for the centralized work flow model. Work flow model fragmentation has been investigated to support the distributed execution and reuse of process models which can gradually partition the centralized process model into fragments, and these fragments can migrate to servers to be executed, further fragmented and forwarded. A mechanism for role-based decomposition of several process fragments so that each process fragment can be executed on a different process engine owned by a different participant [1].

A mapping of the data-centric work flow schema to the pub/sub abstraction to achieve distributed and parallel work flow execution and for the distributed execution of work flows based on the fragmentation of high-level process model into multiple small process fragments and described an enactment environment to execute the process fragments.[2]. As client confirmation and access control is critical in the cloud Susmita et al. depicted the different jobs of TPA. In this paper it is accepted that believed outsiders are autonomous specialist organizations and have a specific dimension of trust [3].

For safely evacuating copied information, Halvevi et al. presented Proof of Ownership (POW) [4] which ensures capacity proficiency on the capacity server. Validation by utilizing labels is utilized to accomplish both effectiveness of capacity and information respectability. It is simple for an aggressor to demonstrate to the capacity benefit that it possesses that le who knows the hash mark of a le, subsequently the server enables assailant to download the whole le. Confirmations of-proprietorship (PoWs) enables a customer to productively demonstrate to a server that the customer holds a le, rather than giving some short data about it. To guarantee secure customer side reduplication, a few Merkle hash tree based confirmation of proprietorship conventions are proposed [4].

Keelveedhi et al. structured DupLESS framework in which message-based keys acquired from a key-server by means of an absent PRF convention [5] enables customers to encode their information effectively. It enables customers to store scrambled information and have the administration perform reduplication in the interest of customers and accomplish solid privacy which ensures encryption for reduplicated stockpiling.

Roberto Di Pietro et al. proposed In practical application, there exists the situation that the plan should be prepared beforehand and empirical data are used to guide the follow-up action. It is infeasible to employ random variable or fuzzy variable to describe these uncertain empirical data. Uncertainty theory has been recently applied to solving such problems. For example, in an uncertain traffic network, the most reliable path was proposed aiming to minimize the risk value of a path under a given con dense level our work absorbs the idea of optimal solution, and uses the idea to solve uncertain optimal deployment problem of process fragments. [6].

A private information reduplication convention permits a customer (who holds a private information) to demonstrate to a server (who holds an outline string of the information) that client is the proprietor of that information without conceding additional data to the server. Small Keong Ng et al. presented private reduplication conventions which depended on standard cryptographic suppositions [7]. The security of private information reduplication conventions is described in the reenactment based system with the blend of two-party calculations.

III. PROBLEM STATEMENT

In order to explain the methodologies which we will use we will set out an example to work and derive data from. Consider a pipe/wire manufacturing company having its production center situated at several locations across the globe. Each production center has roughly around 20 machines and they are operated in shifts which produces pipes. Having set the scenario it is clear that the profits of this company is all dependent on how efficient their machines are running and also to troubleshoot them swiftly whenever a problem arises. Also managers at various key regions would like to see which machines are being operated in which shift along with their RPMs and also the employee tag with running the machines. All these data is to be sent to a Web application which will be used by Managers and deliver real time data and preserve the data in a data warehouse which could be used for analyzing the trend and aid the senior managers at making better business decisions in future.

IV. PROPOSED SYSTEM

The goal of this project is to find ways which can optimize the latency caused by enormous data being generated by IoT based devices in an enterprise level. Given this we are comparing two approaches along with the new Azure Storage function which predicts to deliver better performance as compared to the traditional approaches. Finally we are trying to use fragmentation to club data packets and sent it across the centralized server to optimize the bandwidth utilization.

V. WORKFLOW

Now consider a region R1 where there are 20 machines lined by M1, M2 and so on till M20. Each machine will have to send the following key data as part of their IoT message to the processing power of the Web Application.

- Machine Id
- Current RPM
- Shift ID:
- Employee ID:
- Downtime (time instance since the machine is off)

Such message would ideally be of 2KB considering the Meta data and message header of the network associated with it. So it means every second 40KB (2KB * 20 machines) of data is hogged by this production center in their internet connection. Consider this running for 24 hours and it jumps to 93,456,000 KBs (40KB * 24 hours * 60 minutes * 60 seconds). This dataset is only for one production center and if we include other centers as well the data will increase exponentially. In this electronic data everyone knows that the processing speed for various middle end servers will meet the requirement to process such a huge data inflow the only bottleneck is the network with which the machines and internet are connected. In order to derive value from real time it is crucial and necessary to deliver the data instantaneously. So to compare and select the best methodology to transfer data we look at some of the techniques available in market and try to stimulate a real life example as detailed above.

A. Traditional Approach

Traditional approach is nothing but each machine will have its data sent across to the centralized server via its onboard IoT device. This means there won't be any middleman and direct connection is used to transfer the data. Each machine across all regions will do the same and theoretical this would mean that each data should reach the Central server at the same time.

a) Pros

- **Ease of installation**
- **Resilience:** If any one of the machine is down it won't impact the whole network and other machines will keep on running and sending data normally.
- **Redundancy/Backup:** Since data is processed via a single central server we can also create a RAID like structure to duplicate it across multiple regions allowing us to maintain very less number of servers.

b) **Cons**

- No way to optimize the bottleneck which would incur at the central server level.
- Very hard to troubleshoot and trace the request back to the original machine.
- Latency tends to go higher as time progress because of bandwidth congestion.

B. Fragmentation

Instead of every machine sending data to the centralized server, here we setup a mini server at each location which would do some processing and transform the data of the machines and then pass the data to be centralized server for storing and displaying to the user via a Web Application. This allows us to fragment the data at the very initial level and send it through the network in terms of chunks of packets which tend to be more reliable than the previous approach.

Pros

- Latency is reduced and tends to be constant throughout the whole life cycle.
- Back tracking approach can be used to optimize the network in question by gathering previous data and then changing the size of packets are desired to optimize bandwidth utilization.
- Since transformation of data is done locally, the central server can now be a low level server requiring less processing power.

Cons

- Redundancy and backup is now a challenge since a backup server would have to be created at each production center which may put a heavy burden on budgets.
- Maintenance costs and process also increases given the increase in the number of servers.

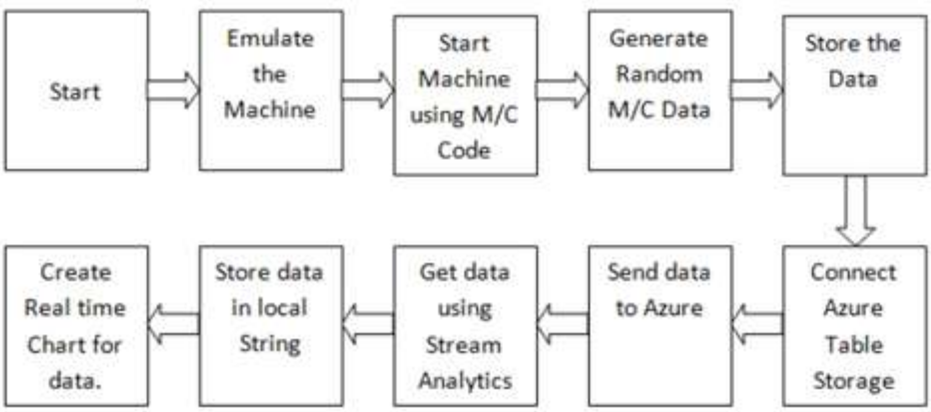

Figure 1. System Architecture

C. Azure Stream Analytics

If we go by the second approach it is very clear that the initial budget to install all the servers and IoT device will cost a fortune and the break-even of the investment will not be immediate, making the investors worried and unwilling to invest in the above architecture. In order to tackle this initial roadblock, Azure one of the leading cloud services offered by Microsoft has came up with Stream analytics as one of their offerings. Using this we can leverage their networks and servers in cloud to transfer and process the real time data and generate meaningful outputs. Here the idea is to have IoT devices on the machines which would directly pass their messages in form of byte streams to Azure Stream analytics server located at various locations across the globe. This in turn will help to reduce the message size by eliminating the need of headers and other unnecessary data load which would have been required for packet transfer. Also since the stream analytics data servers are located at multiple locations it will help us to chose from various server to transfer data depending on their latency. Our assumption is that we can merge this and fragmentation approach to achieve significant less latency time compared to traditional approach.

Pros

- Data is passed as streams, no need to transform or modify data.
- Servers are maintained by Azure, so no extra cost of buying and maintaining the same.
- **Resilience:** This also is taking care by Azure where they have multiple options of backups varying up to backing up of data globally.
- No downtime of servers, since Azure guarantees an SLA of 99.46

Cons

- Less transparency of how data is being routed through the network.
- Azure cost can go higher if not utilized and maintained properly.
- Intense planning and technically skilled labor are required to implement this architecture.

D. Algorithm

a. Use of SHA-512 Algorithm

SHA calculation has more adjusts per byte when contrasted with different calculations. Thus the security given by this calculation is better when contrasted with different calculations. In any case, putting away a SHA bit hash is costly.
The time taken by SHA calculation for creating hash esteem and the quantity of cycles per bytes are productive contrasting with the others. In SHA, the quantity of cycles per byte is to some degree more when contrasted with other hashing capacities, and yet the time required to produce the hashing esteem is a lot littler than others. Along these lines, the SHA hash work is effective and furthermore secure hashing calculation.

b. Steps of SHA Algorithm

- **Step1:** Append padding bits- To bring the message length to 64 bits, message is padded with 1's and as many 0's as necessary.

- **Step2:** Append length- 64 bits are appended at the end of the padded message. This bit indicates the length of the original message holding the binary format of 64 bits.

- **Step3:** Prepare processing functions-SHA requires 80 processing functions.

- **Step4:** Prepare processing constants-SHA requires 80 processing constant words.

- **Step5:** Initialize buffers- SHA requires 160 bits or 5 buffers of words.

VI. RESULT ANALYSIS

Result of this system can be display as differences i.e. comparison between the two graphs in terms of latency. In the upper graphs different machines are running and according to the graphs differ and the blue color shows the graph of the overall machines. As the number of machines increases the latency increases. It works as a directly proportion with respect to time. In the second graphs after fragmentation it clearly shows that the latency has drastically changed as compared to the upper graphs. Hence if the number of machines increases in the future then to the latency of the packets will be less.

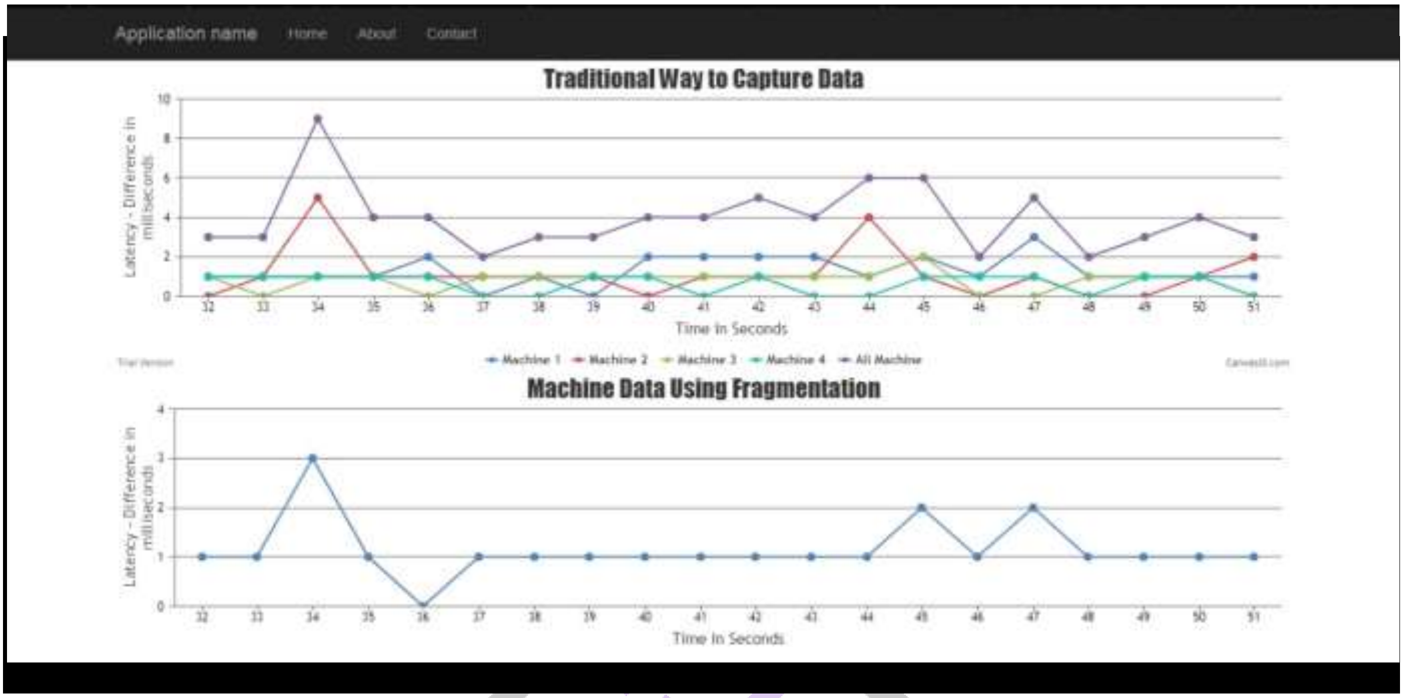

Figure 2. Comparison of packets loss

VII. CONCLUSION

As per the literature and IOT's varied usages in our day to day life. It has been inevitable to see IOT becoming one of the booming technologies that can shape our human life from the very base and make our life much easier as we see now. From wearable smart gadgets to devices keeping production line intact for manufacturing companies, IOT has indeed become the buzzword for this decade and further to come. Every enterprise base companies are utilizing or planning to utilize the potential and market edge which IOT has to offer. Its takes deeper insights on how manufacturing companies can make use of IOT in an efficient way and how IOT could be integrated with legacy systems.

VIII. Acknowledgment

I would like to acknowledge to all who gave guidance, timely advice and support to complete this work.

REFERENCES

[1] Lihong Jiang, Li Da Xui, \An IoT-Oriented Data Storage Framework in Cloud Com-puting Platform", in Proceedings of the IEEE Transactions on Computers, Volume: 10, Issue: 8, Year: 2016, pp. 1551-3203.
[2] S, efki Kolozali, Maria Bermudez-Edo, \A Knowledge-based Approach for Real-Time IoT Data Stream Annotation and Processing, Vol. 5 (3), 2014, pp. 468{475
[3] S. Nair, Anitha K L and R. Kamala,"Trusted Third Party Authentication in Cloud Computing,"in Proceedings of the International Journal of Engineering and Technology (IJERT), Vol. 2 Issue 11,November - 2013,pp. 354-358.
[4] Ralf Tonjes, Muhammad Intizar Ali, \Real Time IoT Stream Processing and Large-scale Data Analytics for Smart City Applications Ralf. ACM, 2011,pp. 635{644
[5] S. Keelveedhi, M. Bellare, and T. Ristenpart, \Dupless: Serveraided encryption for deduplicated storage," in Proceedings of the 22Nd USENIX Conference on Security, ser. SEC'13. Washington, D.C.: USENIX Association, 2013, pp. 179{194.
[6] R. Di Pietro and A. Sorniotti, \Boosting efficiency and security in proof of ownership for deduplication," in Proceedings of the 7th ACM Symposium on Information, Computer and Communications Security,ser. ASIACCS '12. New York, NY, USA: ACM, 2012, pp. 81{82.
[7] W. K. Ng, Y.Wen, and H. Zhu, \Private data deduplication protocols in cloud storage," in Proceedings of the 27th Annual ACM Symposium on Applied Computing, ser. SAC '12. New York, NY, USA: ACM,2012, pp. 441{446.