Proc. of Int. Conf. on Current Trends in Eng., Science and Technology, ICCTEST

"Efficient IoT Data Management for Cloud Environment using Mongo DB"

Anasuya N Jadagerimath[1] and Dr. Prakash. S [2]
[1]Associate Professor,Dept.of CSE,DBIT,Bangalore,India
[2]Professor,Dept.of CSE,DSU,Bangalore,India

***Abstract*—Cloud Computing and Internet of Things (IoT) concepts are currently two of the most popular fields in information technology field, that are expected to change the next era of our life . The convergence of Cloud and IoT will become the important topic of the next decade because of the benefits that IoT has. The Internet of Things (IoT) remains a main topic for standardization and attracts interest from industry, public authorities and end users.**

***Index Terms*— Cloud Computing, IoT, Object Storage, NoSQL, Mongo DB, Datacenter.**

I. INTRODUCTION

The vision of the internet of things has evolved due to a convergence of multiple technologies, ranging from wireless communication to the Internet and from embedded systems to Microelectronics Systems. (MEMS)."Ref[2],that is the traditional fields of embedded systems, wireless sensor networks, system automation and others all contribute to enabling the internet of things. Such a network brings a series of challenges for data storage and processing in a cloud platform. IoT data can be generated quite rapidly, the amount of data can be huge and the types of data can be various. In order to address these potential problems, this paper proposes a databases issues in IOT and identifying aMongo DB as the solution after comparingdifferent NOSQLdatabases.

A. IoT Architecture

As shown in the "Fig 1", there are three main parts in IoT.They are,Different types of Sensors,Network Connectivity and Data Storages.As shown in the figure, Sensors in the IoT devices either communicate directly with the central server for data storage or communicate via gateway devices. The typical wireless technologies used widely are 6LoWPAN, Zigbee, Zwave, RFID, NFC etc. Gateway interfaces with cloud using backbone wireless or wired technologies such as WiFi, Mobile, DSL or Fiber. As shown IoT supports both IPv4 and IPv6 protocols. Due to support of IPv6 which has about 128 bit long IP address length, there are enough addresses available to growing demand of IoT devices. Delay Tolerant Networks is the unique feature of IoT which takes care of large variable delay requirement of IoT based networks compare to traditional computer networks. As shown, IoT service providers offer varied QoS with different pricing and design need for memory, CPU and battery consumption.

Grenze ID: 02.ICCTEST.2017.1.2

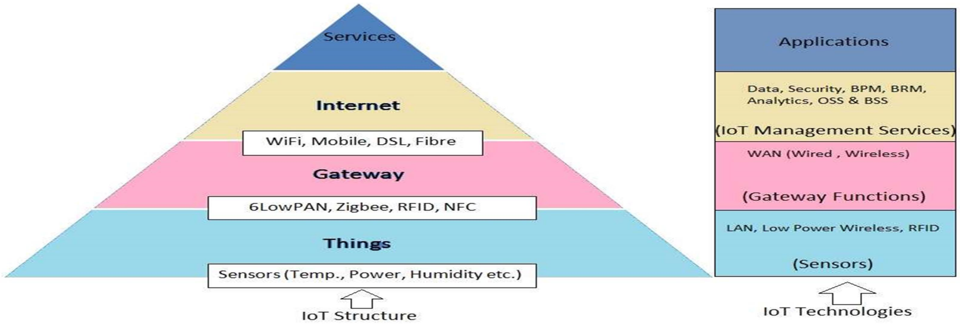

Fig.1 : Basic IOT Architectuire

B. Data Sources in IOT

The "Ref.[24], "Fig 2" shows the details od data collected from different sources.

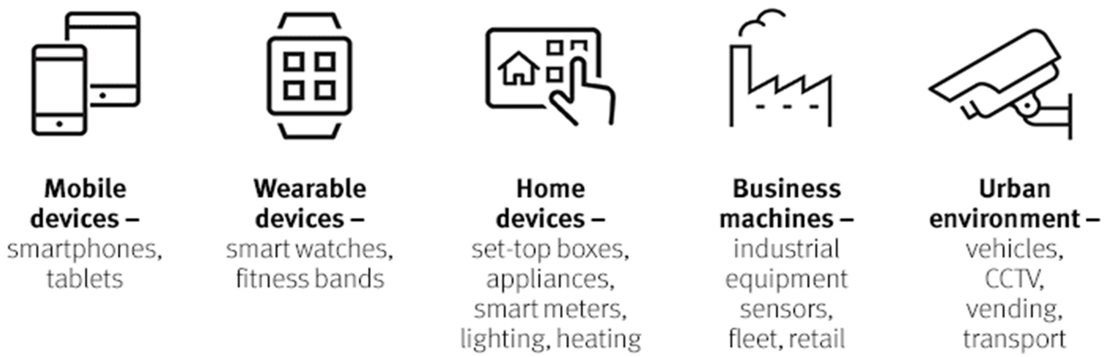

Fig 2: Examples of how different sectors are using IoT include: Connected smart devices in the Internet of Things

C. The Internet of Things is a World of Enterprise Potential

The "Ref. [25],says the global rise of the internet in recent decades, with internet-connected mobile devices and machines producing continuous data, is creating a new world of opportunity for enterprise organizations. The potential of IoT is not just in the devices and the network, but also in the latest 'big data' analytics software. The enterprise opportunities come from being able to combine and analyze IoT data for new valuable insights—and apply these to provide smarter products, innovative solutions and personalized customer experiences. IoT and big data analytics work together to deliver a range of competitive advantages, enabling organizations to,

- Learn consumer preferences and act on observed trends.
- Deliver 'connected' services that are personalized for the individual.
- Accelerate the speed of product delivery
- Reinvent and improve manufacturing process, logistics, and efficiencies.

Of course, the IoT possibilities are not just directly customer-facing—enterprise organizations can also use IoT to transform their internal processes, driving new efficiency, revenue increases and cost savings in areas like operations, supply chain, sales, marketing, maintenance, customer service and more.

- Manufacturing—creating 'smart factories' where IoT elevates automation to a new level, providing unparalleled process control, quality and efficiency.
- Retail—using IoT in everything from learning customer trends, to real-time inventory control with RFID tagging.
- Energy and utilities—deploying 'smart grids' and using IoT for telemetry, metering and general data collection.
- Transport and logistics—using IoT to analyze traffic patterns, increase efficiency and save on fuel costs.

II. Why Datamanagementis the Prominent Step in IOT

The “Fig.3” showas the data growthfrom 2010-15.

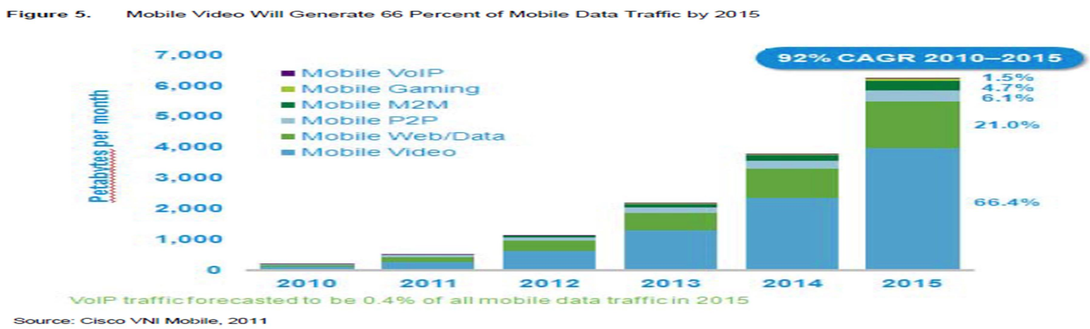

Fig. 3:”Ref.[26]”, Data Growth Rate

According to the survey of Cisco the rate of data growth is shown in the “Fig3”.

III. Database Issues in the Internet of Things

As shown in “Ref [4]”, we set out the areas of challenge for database management in the IoT and point out where developments in traditional databases and other areas may provide the solution.

A. Size, Scale And Indexing

The vast variety of data need to handledIoT. Data will need to be managed via responsible local ownership. Local owners will decide which data and services to make available to the global network. Thus, the IoT may operate on more than one level: private and public. Users may join groups for access to certain privately owned data or May, on the other hand, access data publicly available over the public Internet. There may be differences in quality of data depending on ownership and level of care. Gradually trust and reputation systems will provide information to users on the quality of the data.

B. Query Languages

Current popular query languages in database systems rely on structured data. Structured Query Language (SQL) is the most prominent example. Over the last few years, however, there have been proposals for query languages for semi-structured data, which is more typical of the data held on the Internet. The quantities of data are so vast that it would be unrealistic to expect any sort of uniform structure, except perhaps that of the loosest variety, to be imposed on the IoT. Xtensible Markup Language (XML) offers a means of representing less structured as well as structured data, together with some level of self description.

C. Process Modeling and Transactions

It is likely that most processes will be developed and supplied as services on the IoT. Service Oriented Architecture (SOA) is becoming an important means of supporting interoperability in web-based systems.The central idea is that independent out its offer services in a uniform manner, which other users can then take up. Thus implementation details are hidden from the users of the services. Application processes will typically be made up of a num-ber of lower level transactions. Transactions in turn will be made up of lower level operations or services.Therefore, the question of transaction processing in the IoT arises. In the traditional database systems the matter of concurrent transaction processing has been handled.Through the maintenance of ACID properties through time stamping, locking, and a two-phase commit. ACID properties are atomicity, consistency, isolation, and durability. A transaction must complete in its entirety or not at all, a transaction must leave the database in a consistent State, transactions should not show other transactions, and intermediate results and changes made by a transaction must be permanent. In distributed database systems a two-phase commit is used to preserve consistency. All participating sites must confirm their readiness to commit before the commit command is issued by the coordinating site and written to the database log. It has been recognized that the ACID properties do not web transaction processing well.This is because the individual web services are essentially autonomous and must independently preserve consistency.

D. Homogeneity and Integration

As in “Ref [20],the context of databases the areas of Heterogeneity and integration have been researched since the 1980s, once it was considered useful to achieve interoperability across heterogeneous systems .”Ref [17]” Considering that one might have a personnel system stored at one company in a relational database system, and in another company a similar system might be held in a network database system or even a different relational database system, questions arise as to how to integrate such data. Various solutions have been offered.
As in “Ref.[18]” Some promising solutions suggest the use of a canonical data model, for instance a functional or binary data model. However, it seems that often the solutions offered do not warrant the efforts needed to achieve them.

E. Time Series Aggregation

Time series aggregation is an interesting area, which has been noted as raising challenges in various application domains
It has been recognized that inappropriate time aggregations can give rise to spurious causality.The problem revolves around the ability to select the optimal sampling period for continuous data. Trade-offs include processing time and storage space against accuracy and realistic representation.

F. “Ref.[13]”,Iot Database Requirements.

- Scalability.
- Schema flexibility.
- Distributed Database
- Fault Tolerant.
- High availability.

G. Suitable Database For Iot Is Nosql-Mongodb:

NoSQL database especially key-value, document and column family datyabases which easily accommodates different data types and structures without the need for predefined,fixed schemas. NoSQL databases are googd options when an organization has multiple data types and those data types will likely change over time.
We have categorized the data into the following areas:
1. Radio Frequency Identification address/unique identifiers.
2. Descriptive data, positional and environmental data sensor data
3. Historical data
4. Physics models-that are template for reality
5. State of actuators and command data for control

IV. Discussion On Nosql Database

This section address why NOSQL is used in IoT compare to sql and what are the NOSQL database categories and their performance in IoT.

A. Why NOSQL for IOT?

“Ref.[23]”,Often referred to as NoSQL, non-relational databases feature elasticity and scalability in combination with a capability to store big data and work with cloud computing systems, all of which make them extremely popular. NoSQL data management systems are inherently schema-free (with no obsessive complexity and a flexible data model) and eventually consistent (complying with BASE rather than ACID). They have a simple API, serve huge amounts of data and provide high throughput.

B. Advantages of NOSQL

- Elastic Scaling: NoSQL databases are designed to expand transparently to take advantage of new nodes, and they're usually designed with low-cost commodity hardware in mind.
- Bigdata: Today, the volumes of "big data" that can be handled by NoSQL systems, such as Hadoop, outstrip what can be handled by the biggest RDBMS.

- Economics: NoSQL databases typically use clusters of cheap commodity servers to manage the exploding data and transaction volumes, while RDBMS tends to rely on expensive proprietary servers and storage systems. The result is that the cost per gigabyte or transaction/second for NoSQL can be many times less than the cost for RDBMS, allowing you to store and process more data at a much lower price point.

C. NOSQL Database Categories[14]

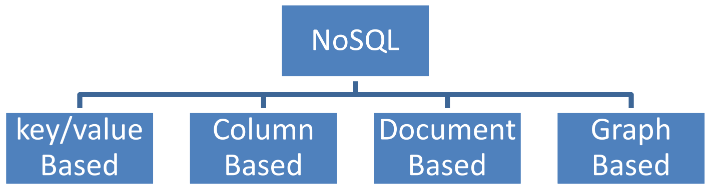

Fig. 4: "Ref.[13]",Categories of NoSQL

D. Based on the NOSQL Categories I Categorized the Data Shown Below: As In "Ref.[23]"

- Cassandra, a column family store.
- HBase (column-oriented, too).
- MongoDB, a document-oriented database.
- Riak, a key-value store.

The detailed comparison is listed in the "Table.1"

TABLE 1: TYPES OF NOSQL DATABASES

	Types of NoSQL database			
	Key/value Based	*Column Based*	*Document Based*	*Graph Based*
Operational /functional models	Redis memCache DB etc	Cassandra Hbaseetc	MongoDB Couch Base etc	Orient Neo4J etc
Data model	Key–Value Store	Column-Oriented Store	Document-Oriented Store	Graph Database
Query language	Any programmi ng Language	Java API REST API	Java ,Xquery	C++,java,phytonetc
Data representation	collection of key-value pairs	Tuple consisting of three elements Unique name,value,time stamp.	Document	Graphs interconnected with representing relations between them

V. WHYMONGODB[27] FOR IOT

Itis a free and openssource cross-platform which is a document oriented database program.MongoDB supports field, range queries, regular expression searches. Queries can return specific fields of documents and also include user-defined Java script functions. Queries can also be configured to return a random sample of results of a given size.
Indexing: Fields in a Mongo DB document can be indexed with primary and secondary indices.
Replication:MongoDB achievedhigh availability with replica sets.
Load balancing: "Ref.[28"] MongoDB scales horizontally using sharding. The user chooses a shard key, which determines how the data in a collection will be distributed.
The data is split into ranges (based on the shard key) and distributed across multiple shards. (A shard is a master with one or more slaves.). Alternatively, the shard key can be hashed to map to a shard – enabling an even data distribution.
File storage: Mongo DB can be used as a file system with load balancing and data replication features over multiple machines for storing files.

Aggregation: Map Reduce can be used for batch processing of data and aggregation operations.
Server-side JavaScript execution: JavaScript can be used in queries, aggregation functions (such as Map Reduce), and sent directly to the database to be executed.
Capped collections: MongoDB supports fixed-size collections called capped collections. This type of collection maintains insertion order and, once the specified size has been reached, behaves like a circular queue.

VI. CONCLUSIONS

The Mongo DB,NoSQL is a good fit for IoT than the RDBMS, because IoT has variety and volume of data which can be handled by Mongo DB efficiently. Mongo DB is thedocument based NoSQL which covers the main requirement of the IoT. The DataAggregation and MapReduce functions of Mongo DB can also be used for BigDataAnalytics.

REFERENCES

[1] https://www.digitalocean.com/community/tutorials/a-comparison-of-nosql-database-management-systems-and-models
[2] Internet of Things (IoT): A Vision, Architectural Elements, and Future Directions JayavardhanaGubbi,aRajkumarBuyya,b* SlavenMarusic,aMarimuthuPalaniswami
[3] www.internetsoceity.org
[4] http://www.cloudamize.com/insights/blog/cloud-comparison-cheat-sheet-aws-vs-google-vs-microsoft-azure/
[5] http://www.datastax.com/nosql-databases/benchmarks-cassandra-vs-mongodb-vs-Hbase
[6] http://www.couchbase.com/nosql-resources/what-is-no-sqlChallenges for Database Management in theInternet of Things Article in IETE Technical Review · September 2009
[7] A.P. Levich (ed.), On the way to understanding the TimePhenomenon: the Constructions of Time in Natural Science,Part I, World Scientific, New York, 1995.
[8] D. Ceballos, and M. Sorrosal, “Time Aggregation Problems in Financial Time Series”, MS’2002 International Conference on Modelling and Simulation in Technical and Social Sciences, Working Article of Institute of Time Nature Exploration, Girona 25-27/6, 2002, pp. 243-52.
[9] E. Newcomer, and G. Lomow, Understanding SOA with Web Services, Addison Wesley, 2005.
[10] https://www.mongodb.com/nosql-explained
[11] http://www.techrepublic.com/blog/10-things/10-things-you-should-know-about-nosql-databases/
[12] A.P. Levich (ed.), On the way to understanding the TimePhenomenon: the Constructions of Time in Natural Science,Part I, World Scientific, New York, 1995.
[13] D. Ceballos, and M. Sorrosal, “Time Aggregation Problems in Financial Time Series”, MS’2002 International Conference on Modelling and Simulation in Technical and Social Sciences, Working Article of Institute of Time Nature Exploration, Girona 25-27/6, 2002, pp. 243-52.
[14] E. Newcomer, and G. Lomow, Understanding SOA with Web Services, Addison Wesley, 2005.
[15] https://www.mongodb.com/nosql-explained
[16] http://www.techrepublic.com/blog/10-things/10-things-you-should-know-about-nosql-databases/
[17] http://internetofthingsagenda.techtarget.com/feature/Find-the-IoT-database-that-best-fits-your-enterprises-needs
[18] http://www.3pillarglobal.com/insights/exploring-the-different-types-of-nosql-databases