RIK International Journal of Science and Technological Research (RIK-IJSTR) *Vol. 7 No. 1, January 2021*
ISSN: 2350-6310 (Print)
ISSN: 2350-543X (Online)

Emerging Database Management Systems: A Review

Mbato, Robinson Nnwobuike
Research Scholar
Department of Computer Science
Ignatius Ajuru University of Education
Rivers State
Nigeria

&

Asagba, Prince Oghenekaro
Visiting Scholar
Department of Computer Science
University of Port-Harcourt
Rivers State
Nigeria

ABSTRACT

With the emergence of new types of computer application such as multimedia, computer aided designs and manufacturing, huge complex data analysis and forecast systems, high intensive data needs sciences in the areas of high energy physics in data storage, data mining, artificial intelligence, etc., with complex data types that are beyond the capabilities of existing conventional databases, it becomes apparent that a new or future database systems or technologies are required to address these needs in a fast changing world. In this paper, we provided an insight on emerging database systems and their contribution to the growing data industry and the drawbacks of the systems. We identified the problems of the existing conventional databases, reviewed extensively the researches carried out by other scholars related to our study, and finally clearly and deeply surveyed some emerging database management systems that have surfaced over the last few years. We adopted the qualitative research method for our review. From our findings, we observed that data redundancy and lack of progressive data integration after initial creation were some of the limitations of the conventional databases which limits it from being suitable for current technologies that require database integrations. This research will be beneficial to database administrators and the research community in the field of database management systems.

Keywords: Emerging Databases, No-SQL Database, Relational Database, SQL.

Introduction

There is an urgent need to deliver very reliable and business critical transactions and business query results more quickly to balance the complexity of managing fast-growing information stores that are generated on daily basis. Timely decisions making and early interventions can only be achieved if there are equivalent responsible systems that can receive, store, process and disseminate such information to processes or procedures that require them for timely decision making. Conventionally, we have the traditional databases created through the process of a relational and hierarchical data models. These developments provided a basis for growth of database technologies. No doubt, these models to some greater extent addressed some aspect of human challenges; these challenges today have assumed new dimensions, making a case to address these limitations with new approach and techniques.

The main driver to this new approach and techniques is speed, cost and scalability. How do we minimize the time it takes to read information stored in the database, potentially resulting in faster query times? Can new data types and standard universal data structure be supported to accommodate or address needs that were not known at the time of data entity/structure creation?, Can there be new architectural framework that can make this possible and at an affordable cost?. This paper will review some of these standards and architectural framework with a view to finding workable solutions for a given type of challenge or requirements.

Statement of the Problem

The types of data stored in the databases were initially relatively simple. However, in the past few years there has been an increasing need for handling new data types in databases, such as temporal data, spatial data, multimedia data, geographic data and many more.

This paper seeks to address the following concerns:

a The Challenges of conventional database systems
b Why we need a new type of database Systems or technology?
c What is the new architectural framework/models that support these new systems or technology?
d How affordable and practicable are these new systems

Aim and Objectives

The aim of this work is to present a clear review of emerging databases, their contribution to the growing data industry and the drawbacks of the system. The specific objectives are:

i. Present a clear understanding of emerging database systems or technology.
ii. Review of the various database models and their limitations
iii. Discuss the framework that supports or governs any type of emerging database design.
iv. Provide guidance to application developers and consumers on the types of emerging database systems that can address or support their current complex database application needs or requirements.

Significance of the Study

This research provides adequate information about the emerging databases which are readily available for utilization today. It clearly states their functionalities and capabilities, which will help developers to make informed decisions about the choice of a DBMS to use for their applications.

Evolution of DBMS

In the early years of computers, performance was a major issue because of the type of processors, computer memory, and computer storage and computer network available at that time. With improvements in sizes, capabilities and performance of these resources, Databases and Database Management Systems also experienced growth in the order of this magnitude [1]. The development of Database technology and associated Database Management Systems can be divided into three eras based on data model or structure. They include:

i. Navigational Database:

In this era, records or objects are found primarily by following references from other objects. The term was popularized by the title of Charles Bachman's 1973 Turing Award paper, The Programmer as a Navigator [2]

ii. **Relational Database**:

This era is based on the relational model of data, as proposed by E.F. Codd [3], and

iii. Post-relational:

This is an era where databases are intended to support an alternative logical data model to the transaction processing oriented relational model [4].

In early 1960, Charles Bachaman from GE (General Electric) used the network model to create the first general purpose database, but by late 1960 of the same year, IBM developed their own Information Management System using the hierarchical data model. In 1970, Edgar Code of IBM developed the relational data model. His work led to several DBMS based on relational model, as well as important theoretical results and ever since, the race to build Database Management Systems began. Today we have over 153 different Database Management Systems built to address different database requirements/needs of organizations or individuals as a result of strong push from the business to get things better and faster. These database management systems and their accompanying architectures have changed or modified to address emerging business challenges.

Conventional Databases and their Limitations: A Case for Emerging Database Systems or Technological Shift

Conventional or traditional database contains persistent data, mostly unaffected by time. The inability of a conventional model to deal with dynamic data that constantly changes makes it difficult to support modern and highly sophisticated operations that are affected with time. An effective system therefore needs to be able to handle time-sensitive queries, return only temporally valid data, and support priority scheduling. By model and design, conventional databases typically consist of many tables, each of which is composed of number of columns. The definition of those tables and columns determine the storage capabilities of database, whereas the relations between the columns define the kinds of facts that can be stored in such a database. Those columns and relations determine the database structure that defines the expression capabilities of database.

The constraints of the conventional databases are further broken into:

1. Once a particular set of data was not covered during the database design and also not included in data model, then such data cannot be stored or represented in the database.
2. Different databases have different data structures, which makes it difficult to integrate data in one database with data from another database or exchange between databases without dedicated data conversion. A database modification or extension requires redesign of the database structure, modification of the software and data conversion, which makes it a relatively complicated and costly exercise.
3. Existence of hardy international standards available or used for the content of the databases, being the data that is entered by its users. This typically means that local conventions are applied to limit the diversity of data that may be entered in those databases. As local conventions usually differ from other local conventions this has as disadvantage that data that are entered in one database cannot be compared or integrated with data in other databases, even if those database structures are the same and even if the application domain of the databases is the same.

Addressing the Limitations of Conventional Database: The Paradigm Shift

Semantic Data Model (SDM): The Gellish Database Approach.

Gellish **[**5**]** is an ontology language for data storage and communication, designed and developed by Andries van Renssen since mid-1990s. It started out as an engineering modeling language ("Generic Engineering Language", giving it the name, "Gellish") but evolved into a universal and extendable conceptual data modeling language with general applications. Ontology languages are formal languages used to construct ontology in computer Science and artificial intelligence [6]. They allow the encoding of knowledge about specific domains and often include reasoning rules that support the processing of that knowledge. In addressing the Semantic limitations which was a major challenge of conventional database, A Gellish method or approach was introduced since it is a subset of a formal language. A Gellish database does not have the semantic limitations that conventional database has. It is flexible and supports standard universal data structure (grammar), which is simple, computer and human interpretable. It consists of one or more database tables, each of which has the same table structure. The fact that those Gellish Database tables are standardized and universally applicable makes a Gellish database application independent. A standardized Gellish database table is universally applicable because it enables the application of following two fundamental principles:

A. The Gellish Principles

i. **Explicit classification of individual things or explicit specialization of classes, with an unlimited number of classes in a dictionary**.

The Gellish database table enables to store any kind of object; because any individual object can be introduced by specification of an explicit classification relation between the object and class, whereas classes can be selected from very large number of classes that are already defined in the Gellish English Dictionary and if the proper class is not available it can be added by specification of a subtype-super type relation with a direct super type of new class. This is fundamentally different from conventional databases that predefine the object types (classes) about which information can be stored by defining a limited number of entity and structure types in a fixed data model

ii. **Explicit classification of relations (facts), by an extensible unlimited number of standardized relation types.**

The Gellish database enables to store any kind of **fact** about any kind of object, because any fact is expressed by a relation, whereas those relations are explicitly classified by relation types that can be selected from the standardized relation types that are defined in the Gellish Dictionary or by relation types that are added to the dictionary as propriety extensions. This is fundamentally different from conventional databases

Related Work

- Shagufta et al [7] presented a literature review of evolving databases. They addressed the evolution of the database from the beginning till date. They also discussed on how several technologies and trends have helped in the evolution process of the database. Finally, a variation in database models since the 90s' till the present century.
- Anisoara et al [8] presented a study on emerging disciplinary research across database management system. It was based on the PIKM 2010, which were the 3rd Ph.D workshops at the International Conference on Information and Knowledge Management (CIKM). The topics addressed included ontology development, data streams, natural language processing, medical databases, green energy, cloud computing, and exploratory search. In addition to core ideas from the workshop, they listed some open research questions in these multidisciplinary areas.
- Singh [9] presented a study on emerging trends in Technologies for big data. They outlined emerging developments in big data which are having remarkable imprint on businesses. Their observation shed light on the trajectory these developments are taking. They also investigated how big data has helped enterprises in different domain, their challenges, opportunities and the future.
- Bhojaraju and Konganurmath [10] proposed a research on database systems, their concept and design. They extensively surveyed the concepts of a database, their applications areas, their limitations and suggested improvements etc. they also carried out a comparative analysis of several database management systems, their advantages and disadvantages. However, they could not trace the evolution of databases and their trends in recent years.

- Hafsa et al [11] presented a literature review on frameworks for querying databases using Natural Language. They carried out a total of 78 reviews of frameworks from 2008 to 2018. These frameworks were then reviewed in order of their supporting language, scheme of their heuristic rule, interoperability support, dataset scope and their overall performance score. Their findings stated that 70% of the work in natural language to database querying has been carried out for SQL, and NoSQL share 15%, 10% and 5% of languages like SPAROL, CYPHER and GREMLIN respectively. They also discovered that most of the frameworks supported only the English language.
- Kristi et al [12] presented a review of history o databases. The main aim of their research was to trace back in time and discover origin of database. Their trace went as far back as the stone ages when nothing digital was in place. They discovered that once the human race could count, database became a concept though it was not formally defined as such.
- Noa et al [13] presented an explorative study on the design needs for the new database era. They discussed the new database technologies and solutions which have emerged in the last decade to meet the demanding requirements of new information systems and applications. They highlighted the tradeoffs which required update or improvements in the traditional database system currently existing. Then they reviewed database design approaches and explored the requirements for these.
- Badia and Lemire [14] presented a study on revisiting database design. They observed that traditional databases are not used in practice. They further argued that if the reverse was the case, there would be a lot of adaptability issues due to the stale design of these databases as compared to the new technologies they interact with in their environment.
- Han et al [15] presented a survey on NoSQL databases. They described the background, basic characteristics, data model of NoSQL. In addition, they also paper classifies NoSQL databases according to the CAP theorem. Finally, the mainstream NoSQL databases were described separately in detail, and some properties to help enterprises to choose NoSQL were extracted and strengthened.

Emerging Database Architectures

New developments in network technologies, artificial intelligence, semantic modeling and object-oriented programming have brought in a lot of inspiration to database designers on how to address the demands in managing complex data types that are beyond the capabilities of existing relational databases. This inspiration has led to the development and research in database topics such as:

i Distributed Databases
ii Deductive
iii Object Orientated Database
iv Active
v Deductive and object-Oriented Databases

Emerging databases have been built based on these technologies. Some examples of Emerging Databases [16]

- i Multimedia Database
- ii Temporal Databases
- iii Geographic Information System Databases
- iv Internet Databases (Database on the World Wide Web)
- v Genome Data Management (human Genes, DNA, etc)

Multimedia Database

Multimedia databases allow users to store and query different types of multimedia information. It hosts one or more primary media files types with following file extensions [17];

- a .txt (documents)
- b .jpg (images)
- c .swf (videos)
- d .mp3 (audio) etc.

These file extensions can further be classified as:

- a Static media (time-independent, i.e. images and handwriting)
- b Dynamic media (time-independent, i.e. video and sound bytes)
- c Dimensional media (i.e. 3D games or computer-aided drafting programs-CAD)

The storage medium of Multimedia Database: Primary media files are stored in binary strings of zeros and ones and are encoded according to file type as defined in the Information Theory Coding template.

There are numerous different types of multimedia databases, including:

1. The Authentication Multimedia Database: This is also known as a Verification Multimedia Database. It encompasses scan able entities like the eye retina and implements 1:1 comparison.
2. The Identification Multimedia Database
   This type involves a data comparison of one-to-many (i.e. passwords and personal identification numbers).
3. Biometrics Multimedia Database
   This type of multimedia database specializes in automatic human verification based on the algorithms of their behavioral or physiological profile. This is a superior technology that Fingerprint Identification and Facial recognition are based.

Difficulties Involved with Multimedia Databases

a. Creating and making multimedia database is costly and very expensive because of the type of technology and huge amount of required bandwidth.
b. Accommodating the vast resources required to utilize artificial intelligence to its fullest.

c. The conventional database made of Binary Large Objects (BLOB) used to store multimedia data does not support content-based searches for multimedia content. This difficulty is as a result of the fact based on the design of the relational database of not been able to recognize the internal structure of a Binary Large Object making it difficult for internal multimedia data components to be retrieved.
d. A relational database is an "everything or nothing" structure- with files retrieved and stored as whole. This inherent feature makes relational database completely inefficient for making multimedia data easily accessible to humans.

Some recommendations to the identified challenges are:
To effectively accommodate multimedia data, a database management system, such as an Object-Oriented Database (OODB) or Object Relational Database Management System (ORDBMS) is implemented as part of an emerging database systems or technology.
Some examples of object Relational Database Management System

a. HP Odaptor
b. UNISQL
c. ODB-II

Tags and Tagging in Multimedia Databases

Non-multimedia data stored in relational databases, multimedia data cannot be easily indexed, retrieved or classified, except by way of social bookmarking and ranking rating, by actual humans. This is made possible by metadata retrieval methods, commonly referred to as tags, and tagging. With this approach, you can search with names of defined objects to get matches [18].

Temporal Database

A temporal database stores data relating to time instances. It offers temporal data types and stores information relating to past, present and future time. These time instances include:

a **Valid time** representing time period during which a fact is true in the real world.
b **Transaction time** representing time period at which a fact was recorded in the database.
c **Decision time** representing time period at which the decision was made about the fact.

A Conventional or traditional database management system is system software for creating and managing databases containing persistent data, mostly unaffected by time. The inability of a conventional database management system to deal with dynamic data that constantly changes makes it difficult to support modern and highly sophisticated operations that are affected with time. This is the reason behind the need of a temporary database. The main goals of a temporary database are to identify an appropriate data type for time, prevent fragmentation of an object description and provide query algebra to deal with temporal data. A temporary database can be used for record the data that is changed with time and it is more convenient, define an object without fragmentation, relation model to describe temporal data, query algebra to deal with temporal data, handle static data, control time dimension etc.

A. History of Temporal Database

With the development of SQL and its attendant use in real-life applications, database users realized that when they added date columns to key fields, some issues arose. For example, if a table has a primary key and some attributes, adding a date to the primary key to track historical changes can lead to creation of more rows than intended. Deletes must also be handled differently when rows are tracked in this way. In 1992, this issue was recognized but standard database theory was not yet up to resolving this issue, and neither was the then-newly formalized standard.

Richard Snodgrass [6] proposed in 1992 that temporal extensions to SQL be developed by the temporal database community. In response to this proposal, a committee was formed to design extensions to the 1992 edition of the SQL standard (ANSI X3.135.-1992 and ISO/IEC 9075:1992); those extensions, known as TSQL2, were developed during 1993 by this committee. In late 1993, Snodgrass presented this work to the group responsible for the American National Standard for Database Language SQL, ANSI Technical Committee X3H2 (now known as NCITS H2). The preliminary language specification appeared in the March 1994 ACM SIGMOD Record. Based on responses to that specification, changes were made to the language, and the definitive version of the TSQL2 Language Specification was published in September 1994.

Internet: Database on the World Wide Web

The Internet revolution of the late 90s have resulted into explosive growth of World Wide Web (WWW) technology and sharply increased direct user access to databases. Organizations converted many of their phone interfaces to databases into Web interfaces and made a variety of services and information available on-line. The transaction requirements of organizations have grown with increasing use of computers and the phenomenal growth in the Web technology. These developments have created many sites with millions of viewers and the increasing amount of data collected from these viewers has produced extremely large databases at many companies [19].

Implementation of World Wide Web databases technology

a. In this technology, a basic client-server kept in publicly accessible shared files are encoded using Hyper Text Markup Language (HTML).
b. A number of tolls enable users to create web pages formatted with HTML tags, freely mixed with multimedia content from graphics to audio and even to video.
c. A page has many interspersed hyperlinks-literally a link that enables user to “browse” or move from one page to another across the internet.
d. This ability has given a tremendous power to end users in searching and navigating related Information-often across different containments, information on the web is organized according to a uniform Resource Locator (URL) something similar to an address that provides the complete pathname of a file.

e The pathname consists of a string of machine and directory names separated slashes and ends in a filename.

The multimedia content is the main database for sharing and distribution and the implementation of emerging database systems or technology is what makes this sharing and distribution possible.

Geographic information systems

A geographic information system (GIS) is a conceptualized framework that provides the ability to capture and analyze spatial and geographic data (topological, geometric or geographic properties). GIS applications (or GIS apps) are computer-based tools that allow the user to create interactive queries (user-created searches), store and edit spatial and non-spatial data, analyze spatial information output, and visually share the results of these operations by presenting them as maps.

There are different components or categories that make up the GIS. They are:

a Cartographic applications
b Geographic applications
c Digital terrain modeling applications

In cartographic and terrain modeling applications, variations in spatial attributes are captured for example, soil characteristics, crop density, and air quality. In geographic objects application, objects of interest are identified from a physical domain- for example, power plants, electoral districts, property parcels, product distribution districts, and city landmarks. These power consumption, voting patterns, property sales volumes, product sales volume and traffic density. The first two categories of GIS applications require a field-based representation, whereas the third category requires an object-based one. The cartographic approach involves special functions that can include the overlapping of layers of maps to combine attribute data that will allow, for example, the measuring or distances in three-dimensional space.

Digital terrain modeling requires a digital representation of parts of earth's surface using land elevations at sample points that are connected to yield surfaces model such as a three-dimensional net (connected observed points as well as visualization). In object-based geographic applications, additional spatial functions are needed to deal with data related to roads, physical pipelines, communication cables, power lines, and such. For example, for a given region, comparable maps can be used for comparison at various points of time to show changes in certain areas.

The functions of a GIS database are:

a **Extensibility:** GISs are required to be extensible to accommodate a variety of constantly evolving applications and corresponding data types. If a standard DBMS is used, it must allow a core set of data types with a provision for defining additional types and methods for those types.

b **Data quality control:** Quality of source data is of paramount importance for providing accurate result to queries. This problem is particularly significant in the GIS content because of the variety of data, sources and measurement techniques involved and the absolute accuracy expected by application users.

c **Visualization:** A crucial function in GIS is related to visualization of the graphical distributes to go with it. Major visualization techniques include contouring through the uses of isoclines, spatial units of lines or arcs of equal attributes values: hill shading an illumination method used for qualitative relief depiction using perspective projection methods from computer graphics. These techniques impose cartographic data and other three-dimensional objects on terrain data providing animated scene such requirements clearly illustrate those standard RDMBSs or ODMBSs do not meet the special need of GIS. It is therefore necessary to design systems that support the vector and raster representations and the spatial functionality as well as the required RDMBS functionality.

B. Geographic Information Systems Data Operation

GIS Applications are conducted through the use of special operators such as the following:

a. Interpolation:

The process derives elevation data for points at which no samples have been taken. It includes computation at single points, computation for a rectangular grid or along a contour, and so forth. Most interpolation methods are based on triangulation that uses the Triangular Irregular Network (TIN) method for interpolating elevations inside the triangle based on those of its vertices [20].

b. Proximity analysis:

Several classes of proximity analysis include computations of "zones of interest" around objects, such as the determination of a buffer around a car on a highway. Shortest path algorithms using 2D or 3D information is an important class of proximity analysis.

c. Raster image processing:

This process can be divided into two categories map algebra, which is used to integrate geographic features on different map layers to produce new maps algebraically, and digital image analysis, which deals with analysis of a digital image for features such as edge detection and object detection. Detecting roads in stalling image of city is an example of the later.

d. Analysis of networks:

Networks occur in GIS as many contexts that must be analyzed and may be subjected to segmentations overlays, and so on. Network overlays refer to a type of spatial join database for example, accident locations to yield, in this case a profile of high-accident roadways.

Some examples of Geographic information systems

a GIS Environmental System Research Institute (ESRI)
b Geo-media (Hexagon Geospatial)
c MapInfo Professional (Pitney Bowes)
d Global Mapper (Blue Marble)

e Manifold GIS (Manifold)
f Small world (General Electric)
g Bentley Map.
h Map Viewer and Surfer (Golden Software)

Genome Data Management

The Genome Database is a public repository of data on human genes and clones. The biologic sciences encompass an enormous variety of information. Environment science gives us a view of how species live and interacting world filled with natural phenomena. Biology and ecology study particular species. Anatomy focuses on the overall structures of an organism, documenting the physical aspects of individual bodies. Traditional medicine and physiology break the organism into systems and the organism as a whole. Histology and cell biology delve into the tissue and cellular levels and provide knowledge about the inner structure and function of the cell. This wealth of information has only recently become a major application of database technology.

Genetic has emerged as an ideal field for the application of information technology. In a broad sense, it can be thought of as the construction of models based on information about genes which can be defined as basic units of heredity-and populations and the seeking out of relationships in that information. The study of genetics can be divided into three branches:

a. Mendelian genetics
b. Molecular genetics, and
c. Population genetics.

Mendelian genetics is the study of the transmission of traits between generations. Molecular genetics are the study of the chemical structure and function of genes while population genetics is the study of how genetic information varies across population of organism.

Characteristics of Biological Data

Biological data exhibits many special characteristics that make management of biological information a challenging problem. Some of these challenges are addressed today using the concept derived from emerging database technology.

Conclusion

Conventional or traditional databases and its associated management system are faced with challenges of meeting up with the complex data/information needs/demands of individuals, organizations, financial and research institutions. This paper discusses solutions provided by emerging database technologies with example databases in addressing these challenges. The kernel of this solution is on Gellish database approach that does not have semantic limitations that conventional or traditional databases have.

References

[1] C. Strozzi, "Nosql-a relational database management system." Lainattu 1998, vol.5, 2014.

[2] M. Iwazume, T. Iwase, K. Tanaka, H. Fujii, M. Hijiya and H. Haraguchi. "Big Data in Memory: Benchmarking Memory Database Using the Distributed Key-Value Store for Machine to Machine Communication. In Proceedings of the 2014 15th IEEE/ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD), Las Vegas, NV, USA, 30 June–2 July 2014; pp. 1–7.

[3] E. F. Codd, O. G. Altrafi and M. O. Ismail. "Relational vs. NoSQL databases: A survey.' *Int. J. Comput. Inf.Technol.* vol. 3, pp. 598–601. 2014.

[4] S. Scherzinger, M. Klettke and U. Störl, "Managing schema evolution in NoSQL data stores." arXiv **2013**, DOI: arXiv:1308.0514.

[5] G. DeCandia, D. Hastorun, M. Jampani, G. Kakulapati, A. Lakshman, A. Pilchin, S. Sivasubramanian, P. Vosshall, and W. D. Vogels, "Amazon's highly available key-value store." *ACM SIGOPS Oper. Syst. Rev*. vol. 41, pp. 205–220. 2007.

[6]. S. Richard, L. Chen and M. Liu. "A survey on NoSQL stores." *ACM Comput. Surv*., pp. 51, 40. 1992.

[7] P. Shagufta, C. Umesh and A. A. Arif. "A Literature Review on Evolving Database." *International Journal of Computer Applications.* Vol.162, No. 9. Mar. 2017.

[8] N. Anisoara, M.S Fabian, and S. V. Aparna. "Emerging Multidisciplinary Research Across Database Management System." European Research Council. Pp. 1-10. 2013.

[9] N. Singh. "Emerging Trends for Big Data." *International Technology Management Review.* Vol. 5, No.4, pp. 202-210. 2015.

[10] G. Bhojaraju, and M.M. Konganurmath. "Database Systems: Concept and Design." Pp. 1-19.

[11] S. D. Hasfa, M. L. Ikramullah, U. D. Moin, M. M. Khalid, and A. C. B. Syed. "Frameworks for Querying Databases using Natural Language: A Literature Review." Pp. 1-5. 2019.

[12] L. B. Kristi, S. Tom, and G. Richa. "History of Databases." *International Journal of Management and Information Systems.* Vol.17, No.1. pp. 29-35. 2013.

[13] R. H. Noa, and S. Arnon, "Exploring the Design Needs for the New Database Era." *Springer International Publishing.* Pp. 276-289. 2018.

[14] A. Badia and D. Lemire. "A call to arms: revisiting database design." ACM SIGMOD Rec. Vol. 40 issue 3, pp. 61–69 (2011).

[15] J. Han, E. Haihong, G. Le and J. Du. "Survey on NoSQL database. In: 2011 6th International Conference on Pervasive Computing and Applications (ICPCA)." pp. 363–366. IEEE, October 2011.

[16] P. O'Neil, E. Cheng, D. Gawlick, E. O'Neil, "The Log-Structured Merge-Tree (LSM-tree). Acta Inform." vol. 33, pp. 351–385. 1996.

[17] F. Chang, J. Dean, S. Ghemawat, W. C. Hsieh, D. A. Wallach, M. Burrows, T. Chandra, A. Fikes and R. E. Gruber. "A distributed storage system for structured data. ACM Trans. Comput. Syst." vol. 26, No.4. 2008.

[18] N. Li, B. Xu, X. Zhao, and Z. Deng, "Database conversion based on relationship schema mapping." In Proceedings of the 2011 International Conference on Internet Technology and Applications (iTAP),Wuhan, China, 16–18 August 2011; pp. 1–5.

[19] M. N. Vora. "Hadoop-HBASE for large-scale data." In Proceedings of the 2011 International Conference on Computer Science and Network Technology (ICCSNT), Harbin, China, 24–26 December 2011; pp. 601–605.

[20] J. Yoo, K. H. Lee and Y. H. Jeon. "Migration from RDBMS to NoSQL using column-level denormalization and atomic aggregates". *J. Inf. Sci. Eng.* Vol. 34, pp.1–17. 2018.