

## A Review Of Non Relational Databases, Their Types, Advantages And Disadvantages

Harpreet kaur,

Student of M.Tech(CSE)  
Sri Guru Granth Sahib  
World University,India

Jaspreet kaur,

Student of M.Tech(CSE)  
Sri Guru Granth Sahib  
World University,India

Kamaljit kaur

Assit.Prof.in CSE deptt.  
Sri Guru Granth Sahib  
World University,India

### Abstract

*Relational databases are widely used in many applications to store data. But there are many problems with relational databases like scalability, handling real time data and handling unstructured data like data on web is not properly structured, it is semi structured or unstructured. To overcome these problems non-relational databases come in to existence. This paper discuss about the non relational databases,their types,advantages, disadvantages and comparison with relational databases.*

**Keywords:** Databases, non relational, Key value stores, Document stores, Column stores, comparison, advantages .

### 1. INTRODUCTION

Databases are the data stores which are used to store the data. Every application needs a data storage unit called database to store its data. Every application has different storage requirements to store its data. Relational databases are widely used from past many years to store the data[3]. They store the data in relational or we can say tabular form. They mostly use SQL as their query language to store

and retrieve the data from the databases. Non relational databases are those databases which donot store the data in relational form and mostly they donot use SQL as their query language that is why they also called NoSQL databases. These databases handle various types of data like graph, objects and many others. These databases handle semi-structured and unstructured data very efficiently which cannot handled by relational databases. Also, these databases can handle very large amount of data in distributed environments because they are highly scalable and cost and complexity of sharding is very less as compared to relational databases. Some non relational databases provide auto sharding feature. The NoSQL or non-relational community has widely adopted CAP theorem. CAP stands for[1] :

**Consistency:** Consistency means after an update or write operation the database must be in consistent form. If any user update some information or data then all other user must see that update after update operation in distributed system or in shared resource[1].

**Availability:** Availability means that a system should continue to operate even if its some

hardware or software parts crash or not working[1].

**Partition tolerance:** A system should work in network partitions also. It should cope with dynamic addition and removal of nodes in a network[1].

Non-relational databases do not have transactions so they do not follow ACID instead they follow the BASE properties which are given below[1]:

**Basically Available:** It means an application should work basically all the time[1].

**Soft state:** An application need not to be consistent all the time[1].

**Eventual consistency:** An application should be in some known state eventually[1].

The table below shows the difference between ACID and BASE[1]:

ACID	BASE
Strong consistency	Weak consistency – stale data OK
Isolation	Availability first
Focus on “commit”	Best effort
Nested transactions	Approximate answers OK
Availability?	Aggressive (optimistic)
Conservative (pessimistic)	Simpler!
Difficult evolution (e. g. schema)	Faster
	Easier evolution

**Table 1.1:** Difference between ACID and BASE[1]

## 2. NON RELATIONAL DATABASES TYPES

There are following types of non-relational databases given below:

### 2.1 Key Value Stores:

The key value store databases store a value corresponds to a key. They have very simple structure. These are also highly scalable. They have high query processing speed than the relational databases. It provided support for query and modify operations for data through primary key. Key values represents buckets of data.[3]. There are many key value stores available in the market, some of which are described below[1]:

#### Amazon's Dynamo

Amazon's Dynamo is a very famous key value store which is developed at Amazon and also used at Amazon for various purposes. It has many advantages over other relational databases.

For partitioning it uses consistent hashing. It provides advantage of Incremental scalability. To provide high availability for writes it uses vector clocks with reconciliation reads technique[7].

#### Project Voldemort

Voldemort is a distributed data store that is designed as a key-value store used by LinkedIn for high-scalability storage.<sup>[1]</sup> Voldemort is still under development. It is neither an object database, nor a relational database. It does not try to satisfy arbitrary relations and the ACID properties, but rather is a big,

distributed, fault-tolerant, persistent hash table[8].

There are many other Key value stores which are listed below[5]:

- Riak
- Redis
- Memcached
- Berkeley DB
- Tokyo cabinet

## 2.2 Document Stores

Document stores are considered as next step of key value stores. They store structured, semi-structured and unstructured data. They store data as a document and in more meaningful manner than the key value stores. They are schemaless means that the schema is not previously defined as in case of relational databases. The main two document stores MongoDB and Apache CouchDB are explained below:

### Mongodb

Mongodb is an open source document oriented database comes under the category of non-relational databases. It is developed and supported by 10gen. MongoDB stores structured data as JSON like documents with dynamic schema. Schema is not rigid. The documents are stored in a collection. A typical MongoDB document is look like following[6]:

```
{
  "_id":
  ObjectId("4efa8d2b7d284dad101e4bc9"),
  "Last Name": "DUMONT",
  "First Name": "Jean",
  "Date of Birth": "01-22-1963"
},
{
  "_id":
  ObjectId("4efa8d2b7d284dad101e4bc7"),
  "Last Name": "PELLERIN",
  "First Name": "Franck",
  "Date of Birth": "09-19-1983",
  "Address": "1 chemin des Loges",
```

```
"City": "VERSAILLES"
}[6]
```

### Apache CouchDB

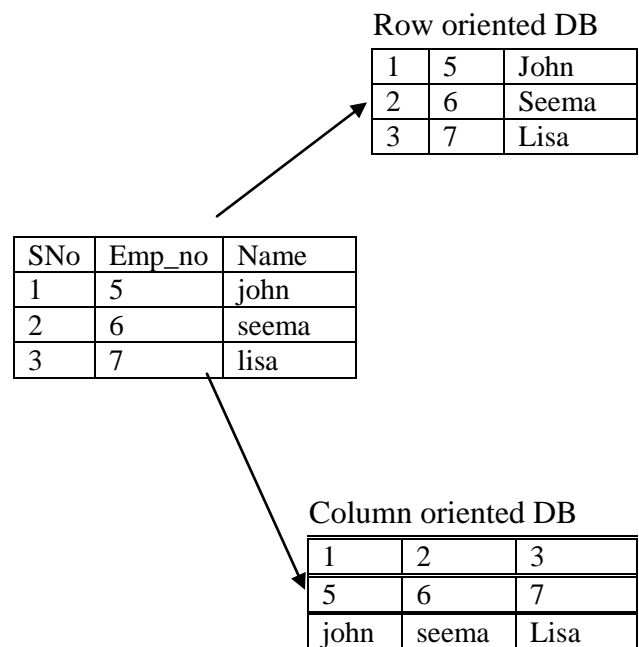
Apache CouchDB, commonly referred to as CouchDB, is an open source database that focuses on ease of use and on being "a database that completely embraces the web. It is a NoSQL database that uses JSON to store data, JavaScript as its query language using MapReduce and HTTP for an API.[9].

There are many other document stores given below[5]:

- SimpleDB
- OrientDB
- Jackrabbit
- IBM Lotus Domino
- Couchbase server

## 2.3 Column Stores

Column stores are non-relational databases which store the data in column oriented tables instead of row oriented like relational databases. Columns have family in which data is stored.



**Fig.2.1** : Column oriented vs. Row oriented

The two main column oriented databases Google's bigtable and cassandra are explained below:

### Google's Bigtable

BigTable is a compressed, high performance, and proprietary data storage system built on Google File System, Chubby Lock Service, SSTable (log-structured storage like LevelDB) and a few other Google technologies. BigTable maps two arbitrary string values (row key and column key) and timestamp (hence three dimensional mapping) into an associated arbitrary byte array. It is not a relational database and can be better defined as a sparse, distributed multi-dimensional sorted map[10].

### Apache Cassandra

Apache Cassandra is an open source distributed database management system. It is an Apache Software Foundation top-level project designed to handle very large amounts of data spread out across many commodity servers while providing a highly available service with no single point of failure. It is a NoSQL solution that was initially developed by Facebook and powered their Inbox Search feature until late 2010[11].

Other Column stores are listed below[5]:

- Apache HBase
- MonetDB
- Hypertable
- Mnesia
- Apache Accumulo

### 3. Advantage of Non relational Databases

Non relational databases have many advantages which are listed below[2]:

- Non-relational databases process data faster than the relational databases because they do not use ACID properties.
- Non-relational databases have simpler

data models than the relational databases.

- Non-relational databases are highly scalable than the relational databases.
- Non-relational databases are very flexible than the relational databases because they are schemaless.
- Non-relational databases can handle a very large amount of data.
- Non-relational databases have high performance than relational databases.

### 4. Disadvantages of Non-relational databases

Non-relational databases also have many disadvantages which are listed below[2]:

- Non-relational databases are less reliable than the relational databases because they compromise reliability for performance.
- Non-relational databases also compromise consistency for performance unless manual support is provided.
- Non-relational databases use different query languages than relational databases. So, people find it an overhead to learn new query languages.
- In many non-relational databases security is lesser than relational databases which is a major concern. Like MongoDB and Cassandra both databases have lack of encryption for data files, they have very weak authentication system, and very simple authorization without support of RBAC.[4].

## 5. Comparison between relational and non-relational databases

There is a lot of difference between relational and non relational databases which is described through a table below:

Relational databases	Non-Relational database
Low scalability	High scalability
Performance is lower than non-relational databases	Performance is high.
Reliability is high.	Reliability is low.
Use ACID properties	Use BASE properties
Data processing is slower than non relational databases	Data processing is faster than relational databases.
Complexity and cost of sharding is higher .	Complexity and cost of sharding is lower than relational databases.
Security is higher.	Security is lower in many non relational databases.

**Table 5.1:** Difference between relational and non relational databases.[1][2][4].

## 6. Conclusion

We have reviewed the non relational databases, their major types, various advantages and disadvantages as compared to relational databases and finally we compared the non relational databases with the relational databases through a table.

## 7. References

[1] C. Strauch, "NoSQL Databases," February 2011. [Online]. Available: <http://www.christof-strauch.de/nosql dbs.pdf>.

[2] Neal Leavitt " Will NoSQL Databases Live Up to Their Promise?" IEEE Computer Society 0018-9162/10/\$26.00 © 2010 IEEE.

[3] Clarence J M Tauro, Aravindh S, Shreeharsha A.B "Comparative Study of the New Generation, Agile, Scalable, High Performance NOSQL Databases" International Journal of Computer Applications (0975 – 888) Volume 48– No.20, June 2012.

[4] Okman, L.; Gal-Oz, N.; Gonen, Y.; Gudes, E.; Abramov, J.; , "Security Issues in NoSQL Databases," *Trust, Security and Privacy in Computing and Communications (TrustCom), 2011 IEEE 10th International Conference on* , vol., no., pp.541-547, 16-18 Nov. 2011 doi: 10.1109/TrustCom.2011.70.

[5] Stefan Edlich, "List of NoSQL Databases," July 2011. [Online]. Available: <http://nosql-database.org>

[6] MongoDB. Mongoddb. [Online]. Available: [http:// en.wikipedia.org/wiki/Mongoddb/](http://en.wikipedia.org/wiki/Mongoddb/)

[7] Amazon's DynamoDB [online]. Available : [http://en.wikipedia.org/wiki/Amazon\\_Dynamo DB/](http://en.wikipedia.org/wiki/Amazon_Dynamo_DB/)

[8] Project Voldemort [online]. Available: [http://en.wikipedia.org/wiki/Project\\_Voldemort](http://en.wikipedia.org/wiki/Project_Voldemort).

[9] Apache CouchDB [online]. Available : [http://en.wikipedia.org/wiki/Apache\\_CouchDB](http://en.wikipedia.org/wiki/Apache_CouchDB)

[10] Google's Bigtable [online]. Available:  
<http://en.wikipedia.org/wiki/BigTable>.

[11]Apache cassandra [online]. Available:  
[http://en.wikipedia.org/wiki/Apache\\_Cassandra](http://en.wikipedia.org/wiki/Apache_Cassandra).

IJERT