

*NoSQL overcoming BigData Failures*

**Abstract-**More products, more consumers create the need to use more data for prediction and analysis. Insights derived from data is the heart of the e-Commerce industry as it helps to gain an edge over the competitors. Huge data leads to the usage of the big data. Some e-Commerce industry fails to implement this. This article explains how they fail and what can be done to overcome from the problem.

**Introduction:** - 80 percent data in the Retail industry is considered from in-store transactions and e-Commerce whereas consumer and production information is never considered [1]. The challenge arises while managing the volume, velocity and variances in the data. Some of the examples of the big data projects which the problems faced by them:

**Data Volume-** It becomes difficult for a well established company with one billion consumers who buy its products weekly to maintain personalized targeted messages.

**Complexities-** Data velocity creates problem for the retail XYZ to manage its supply chain globally, analyze real time sensor data, data from RFID chips which is established to manage each item independently having one specific code for each item, maintaining the receipts and the shipment notifications.

**Customer loyalty database-** For a Retail company A, managing a database of 70 million consumers which includes their demographic, preferences, purchases history as well as location information and use the information for the live transactions and promotions both online and in-store.

Maintaining the variability in the data and all the above is an everyday challenge for retailers. Retail industry basically rely on mainframes or relational database management systems. But its difficult for both of them to handle this kind of complexities of the data management. Both lack the variability, velocity and volume handling issues which are inherent to big data.

Businesses wants to understand Big Data and leverage it. It does not only take skills but also finesse to understand and use Big Data to its fullest. Due to poor implementation, businesses fails and some of the reasons which fails the implementation of Big Data.

**Organization imbroglio:****1.Challenges with structure and schema**

Retail industry always utilize RDBMS using programming languages driven by SQL, Oracle. The data needs to be stored in a structure. For an example, United States address which consists of a structured format like street number, city, state and Zip code. As the aggress format is always going to be consistent so we need a structured data consisting of columns within tables.

**Managing unstructured data source:** Maintaining the schema driven RDBMS database is helpful for ERP and POS. The problem arises when we have to use relatively unstructured data including:

**1.External unstructured data-** This kind of data includes digital images, videos, XML and HTML web pages which are rich but doesn't fit well into rows and columns.

**2.Internal structural data –** Data like PDF, PowerPoint presentations, different kind of sites post which are stored at multiple legacy systems and databases which are not used properly. It provides vital information to the organization and about the users.

**3.External structured data-** This kind of data is used for analytics and product development and consists of location, loyalty, credit and demographic information related to consumers.

**2.Getting data into the database-** Fitting data into a pre defined row and column structure needs a lot of analysis and then mapping them into the rows and column is not only time-consuming but also costly. Managing data for billion consumers across multiple countries, brands and channels is impossible and hence most of the Big Data projects are not able to pass through this stage.

**3.Enabling Analytics-** In retail industry, making live promotions offers online or in-store, checking product availability, responding customers queries accurately and within shorter span of time becomes difficult as data is stored in multiple legacy systems and databases which are not designed to talk to each other. Hence, analyzing the data and generating an insight out of the data which are rarely linked to the store location or online-purchase data is a big challenge for the retail industry to use Big Data.

**1. Unabated Requests [2]-**

It becomes critical to create a group to funnel, prioritize and sequence requests. Big data becomes a risk in solving all the

organization problems which is being expected by the top notch managers of the organization. Due to lack of synergy and accountability, which occur due to communication gap between the vertical synergy between different teams, a lot of information is lost in translation making Big Data a curse for the project.

2. **Falling for Big Data, Unprepared**  
Big Data Analytics tools are meant to provide inferences such as trading, process and asset recommendations. It crosses path with the work of senior business analytics and functional experts. It is quite important to be prepared and organized for the positioning of the Big Data tools.
3. **Expanding without foresight**  
A heavy equipment manufacturer wanted to use predictive analytics to provide maintenance recommendations to customers by brute forcing the data Warehousing and Business Intelligence infrastructure. Plan was to build a store procedure engine for predictions by re-labeling their DW/BI investments to Big Data Analytics. But Big Data uses deeper mindset which needs to be introduced while expanding the offering and should include the predictive analytics at that time. Hence, expansion of predictive analytics for the recommendations resulted in engine seize.  
A proper planning, algorithm development and proper instrumentation can only help the projects to survive.

#### **Challenges faced by Retail Industry:**

In 2015, there was an increment of 14.2 percent of online sales in the retail industry whereas only 1 percent increment in the retail industry. One of the biggest giant Walmart, their global e-commerce sales increased by 22 percent to 12.2 billion whereas total retail sale increased by only 1.8 percent. Even after that, e-commerce is facing significant challenges and opportunities to convert the visitors into buyers. Even now only 2.84 percent e-commerce visitors bought anything as well fewer than 1 percent due to traffic made a purchase using smartphones. For the e-commerce giants (such as Walmart, Best Buy and Target) even increase from 3 to 5 percent sales results in \$2 billion gain.

Most retailers are unable to manage and extract value from Big Data with databases technologies which are quite common nowadays due to:

#### ***1.Product data complexities:***

A retailer's product data is quite complex as it contains structured and unstructured data which are integrated together into the database. Multi-structured sources like digital images, videos, ratings, reviews, customer loyalty information and product information or related products and services. A typical electronics retailer might have more than thousands SKUs while a part distributor will have over a million.

#### ***2.Search and Update challenges:***

It is very difficult to provide the information which the customer are looking for. Matching the specific terms as per the customer requirements produce different results for different inputs provided by the consumers. Online retail search suffers from "Goldilocks syndrome" which displays some irrelevant products. Online search delivers a very different experience than what we get into the stores.

Same with the updates of the data related to the products. When few more options are added to a new group of products. That is why most of the e-commerce search technologies are bottled into relational databases because of the pre-existing schemas which determines the storage and the restriction of the search criteria by attaching the number of attributes to be associated. It helps the customers to filter the data. Whereas developers face a lot of problem updating predefined schemas and requires additional code to make the feature work as per the requirement.

#### ***3.Relationship between product data and information:***

Customer buy solutions (such as recipes for dinner, DVD for the TV, cartridge for the printer) to meet the needs. If there is no link between the product data, retailers lose a lot of money online and in-store when lacks in providing the recommendations related to the products and the supporting products.

One of the key factors to keep in mind while using Big Data is to invest less time being defensive and more time being inventive. Due to the problems faced by the industry, we need to look for a better solution.

***NoSQL as a rescuer***

Due to the data disconnect, we need a more flexible and scalable database. Traditional databases lack both of the qualities to inherent in big data. NoSQL loads up the data first and help to find the problems instead of getting schema just before the execution phase. This approach helps to focus on querying rather than the structuring of data to fit in the traditional RDBMS.

**Attributes influencing to use NoSQL as a Platform:**

While considering e-commerce, retailers need to look for the below capabilities- *Product Data management*- The scalability and complexity of the product data that needs to be ingested is the biggest concern. The database must be optimized to store JSON, XML and geospatial data as well as the relationships, binary video files and PDF's without conversion to display a better output to the users.

***Ease of changes and development***

Conduction the searches based on the attributes is one of the biggest problems of the relational databases. While using NoSQL, changes should be made without mapping the data and saving the development time. As well data storages are indexed as documents, no normalization is required and should avoid the shape of the data. Hence, it helps in saving the time and energy which can be used in the ETL process.

***Enhanced Search Capabilities***- NoSQL database should also provide non attribute related search that is expected to be provided by an enterprise application so that alternative product and their features also becomes easy to search. The search should be focused more on relevance ranking, snippeting, proximity boosting and language support. The search functionality should be integrated into the database to help the developer's and the DBAs work easy as they do not need to worry about an extra platform.

***Semantic capabilities***- For an example, if you want to buy a piano, semantics should determine the products like stands, earphones, books which supports the specific brand. The best NoSQL database

uses semantics to store relationships between linked products. With semantics, searches become easy as it provides relevant information as per the term entered, publish information dynamically to the different kind of devices.

**Conclusion**- For the retail industry e-commerce is the most beneficial platform but due to search and update issues customers get frustrated and hence the industry end up losing 66 percent of the customers. If we are using the Big Data tools, then we should be using NoSQL platform which simplifies the product data management and intelligent search. NoSQL indexes the data automatically leading to intelligent query by the consumers. This simplification helps ecommerce increase their sales and also provide customer satisfaction.

**References:**

- 1) <http://tech.firstpost.com/news-analysis/7-301424.html>
- 2) <http://www.marklogic.com/resources/e-commerce-and-nosql-in-retail/>