BIG DIGITAL MEDIA DATABASES IMPACTING E-COMMERCE

Ştefan SFICHI
Ştefan cel Mare University of Suceava, Romania
stefans@usm.ro

Abstract:
Big data is the combined collection of digital data and old - traditional data inside or outside an entity. Media is being encapsulated in this big data offering powerful insights for e-commerce businesses. This article intends to review the good and maybe the bad part inside this evolution and to overview the crypto part inside future e-commerce businesses.

Key words: Big data, big digital media databases, media, media databases, e-commerce

JEL classification: A12, A19

I. INTRODUCTION

This paper intends to review the importance of big digital media databases in nowadays e-commerce all over the world. Databases evolved from the simple text information contained in many different forms that include much more info then the old ones.

This market has a lot of money involved. It has great advantages for both the sellers and the buyers. This is why it is very possible that it should evolve to getting a greater piece of the market.
We shall explain what big data is and we shall speak about increased shopper analysis. As we've mentioned, understanding big digital media databases is good for the sellers, but also for the buyers. This is why we shall overview improved customer service, on-line payments, the advances in mobile commerce [1]. At the end, virtual reality should be taken in the spotlight for it is already helping sellers and buyers get closer.

II. WHY BIG DATA

Big data is everywhere. Data is coming in from different sources in different forms. Facts are now integral to our future. This information can improve your workforce. It can transform both the lives of businesses and costumers. It doesn't do that all by itself. Big data needs context and connection. Innovation depends and that context, on that connection [2]. There are three important things when we we think of big data. There is volume, velocity and variety.

Big data has high volume of information that needs to be processed. Being also low-density, unstructured data, this is also quite hard to take into consideration when processing. For some organizations, it may get to almost a thousand petabytes.

Velocity speaks about the rate data is received and accessed. There are some companies that need real time evaluation and action upon. In that case, the speed is very important and most of them get to work with this data in the memory. After that it get written to disk, maybe.

Figure 1. Big data servers are different than the regular hosting centers [3]

Big data is the combined collection of traditional and digital data from inside and outside an entity. Its purpose is to be a source of analysis and continued discovery [4]. Figure 1 presents us the way big data servers look like. This is how businesses got to the point where they can access huge amounts of data, all ready to get analyzed.

Big data almost always wins [3]. Why is that? Because regular hosting can be viewed as unscalable but with simple operations, big data servers are the best choice for expanding the years to come. So if someone should want to evolve and get bigger, this is the way to go. We can view the big data servers as scalable. They can have modules added, to be short.

III. BIG DATA SERVERS: KEY COMPONENTS

Let's try and build out a sample of a big data servers. It's critical that we put the right type of components on the data analytic environment. The server has processors. That means at least 2, but usually more. The processors are multiple cores. We also have memory channels. There must be at least 3 channels dedicated to each processor socket.

For this, we need task manager, data node, MapReduce and hadoop bluster (Hbase).

For the configuration, we are taking into considerations multiple core processors, at least 3 times memory channels (per processor) [5].
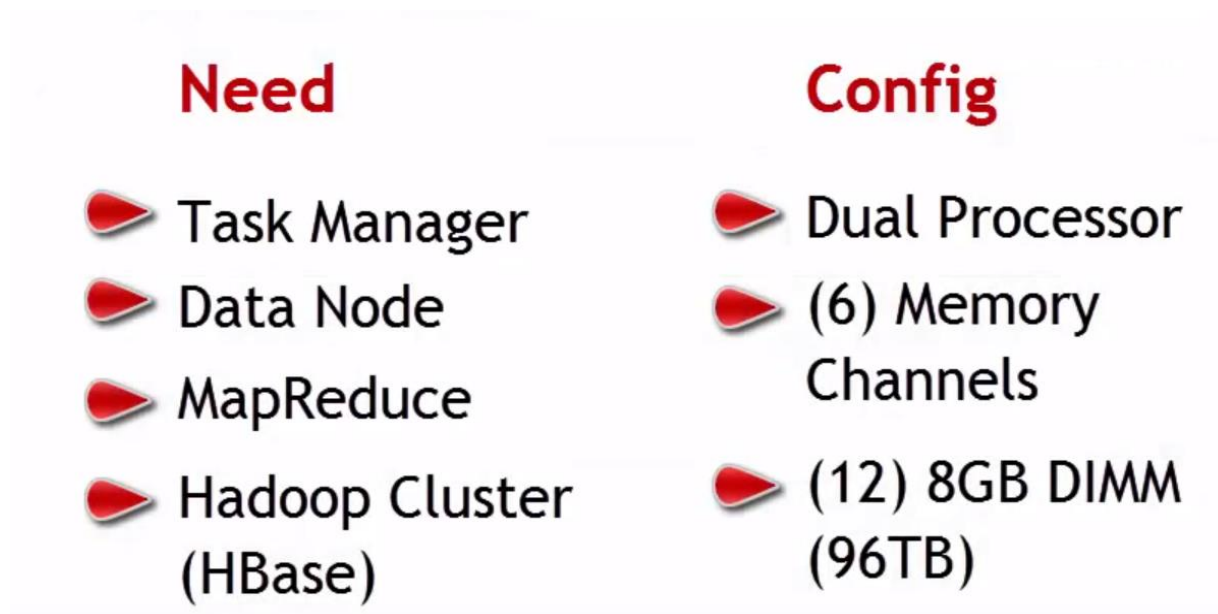
Need
- Task Manager
- Data Node
- MapReduce
- Hadoop Cluster (HBase)

Config
- Dual Processor
- (6) Memory Channels
- (12) 8GB DIMM (96TB)

Figure 2. Big Data Servers: Key Components [5]

It is important that all worker servers should scale together. What we presented until now is a worker server. The manager (node servers) are a little bit different. The key components of a big data server may be seen in Figure 2.

What is the difference between server computers and desktop computers? Understanding this could get us a step closer to the future that big data computers prepared for us. Fundamentally they are not that different. They all have the same basic components: processor, working memory, storage, power source, etc. They all have the same assembly line. If we think that we have a I7 that is working at 3.8 GHz and we also may think of a Xeon running the same 3.8GHz. At base, they are all the same, but the Xeon (witch is a server processor) costs way more. Why is that? The server components have inter compatibility with other server components. The last and more expensive ones (that run on servers) have ECC Memory Supported. It's a specification that usually is not taken into consideration. This means that servers use memory that is slightly more resistant to errors. For a typical desktop user, this is not critical, but for a server it could result in a lost of a patient medical record.

The server clients will spend money on: fancy networking capability, remote management interfaces and hardware redundancy.

The server grade hardware is optimized for different workloads. A server needs maximum performance, minimum space and minimum power use. All of these make the servers needed by the big data cost more, but still less when thinking overall.

IV. E commerce and databases

It is not very easy to find the largest e-commerce companies in the world. But because one thing a person should know best is to search, I came up with a top 8 largest e-commerce companies since February 2018 [6]. For now, it should work. That companies are:
1. Amazon, Inc., founded in 1994 in Seattle
2. Jingdong, founded in 1998 with a well over quarter of a billion registered users (2018)
3. Alibaba Group Holding Ltd., founded in 1999 and is the world's largest retailer (over 200 countries)
4. eBay Inc., the first successful dot-com bubble, a symbol of the 90's, founded in California in 1995,
5. Rakuten, Inc., Japan's largest online bank operator, founded in 1997
6. B2W Companhia Digital, founded in 2006, the bigest of its kind in Latin America – market share of over 50%

163

7. Zalando, the first European company in this list
8. Groupon, american e-commerce marketplace launched in Chicago in 2008.

Figure 3. Top ecommerce databases driven platforms in the world

These companies are evenly distributed geographically, covering each corner of the globe. Regardless where you live, if you decide to purchase something on-line, you will come into contact with one of them [6].

Because Amazon, Inc. is no 1, we shall analyze a fraction of its interest in databases. One thing is sure: Amazon may never stop investing for growth. Its razor-thin profit margin and roller-coaster earnings results are a result of the continually going through new investments strategy. The major thing is the e-commerce business, but it also has the largest public cloud computing service in the world. Someone might ask himself why!!! Because the two of them go together. Amazon needs this cloud computing service and has bought itself at the best price. And because the bigger you are, the better price you may obtain, Amazon has the largest public cloud computing service in the world. At least for now.

IV.1. The biggest one for now

At the moment of this article being written, Amazon, Inc., has the bestB infrastructure when thinking of big media databases. Because of its needs, it invested in cloud and now is a big seller on this market. They give away 1 year free membership – enough time to figure out the needs of each company.

Where your working in retail or realestate, financial data or photos, your business relies on applications that run reliable and securely at the cost and scales of your needs. AWS (Amazon Web Services) provide a complete set of cloud computing service, accessed by the internet, to help you build and run this applications [7].

Cloud computing offers all that nowadays businesses need:
- compute services
- storage services
- database services.

The great thing is that depending on the needs, anyone may access these services. This is why a big firm like this, that can be accessed by clients all over the world (hundreds of thousands of clients), can offer powerful resources at a low operational cost. One thing is for sure, these services are imperious and far more simple to use for e commerce firms. It's all about scaling. Their customers can scale up and scale down anytime. This is very important for growth and for profit margins, helping removing firms from margin calls. It is clear that cloud computing has emerged as

an extremely popular implementation option for a wide range of computing services [8]. To understand more, let us review in short figure 4 – an example of a global network architecture.
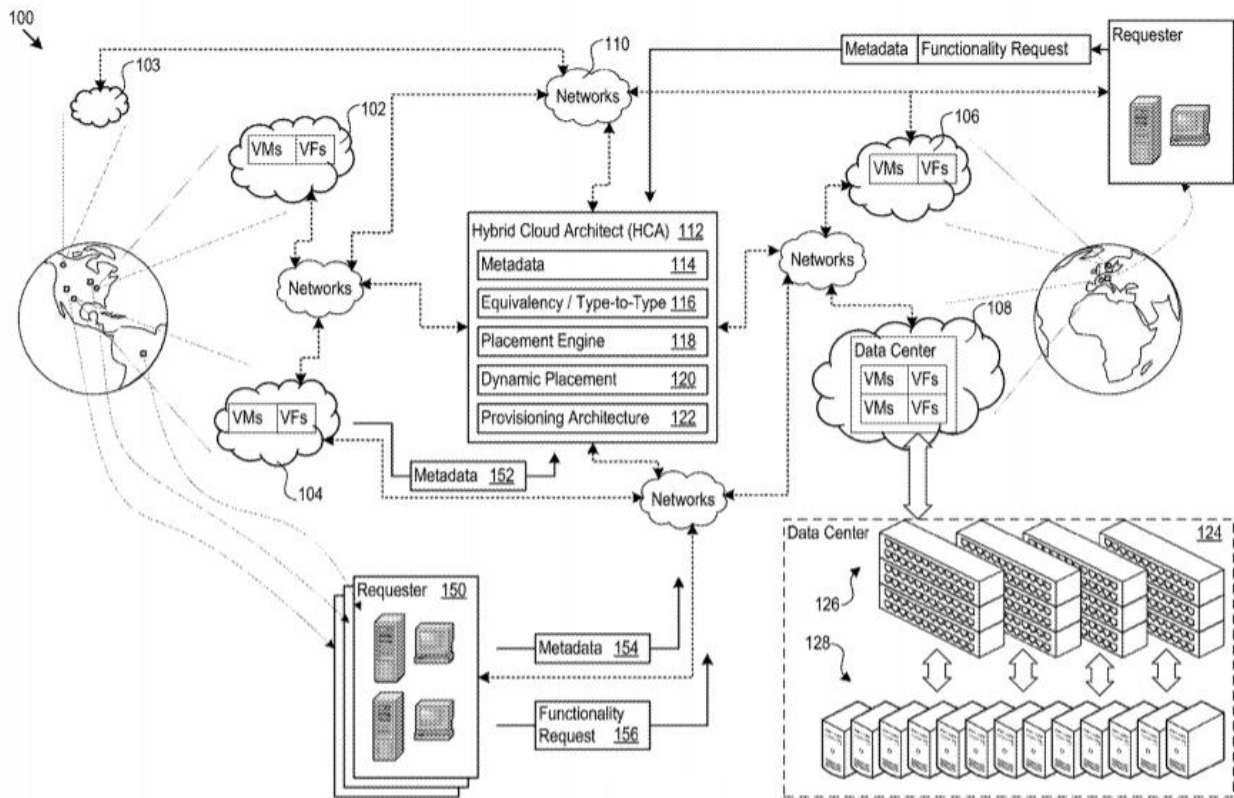

Figure 4. Example of a global network architecture [8]

In Figure 4 you can view an example of a global network architecture. This architecture implies a data center that stores data coming from the network. The data center consist in VMs and VFs. VM comes from Virtual Machine and VF from Virtual Function. The Data Center connects to the Networks. We have Requester-s from all over the world that also connect through the network so they could get to the HCA (Hybrid Cloud Architect) that has address 112 in this figure. It encapsulates Metadata, Equivalency / Type to Type, Placement Engine, Dynamic Placement and Provisioning Architecture.

V. E commerce and big data

The In-Store Experience can be safely, very well personalized with the information that this kind of databases may offer. In the past, merchandising was evaluated as a form of true art. It couldn't be really measured at the end. We are on the verge of a real change. Now, shoppers search in stores, preview in stores and tend to buy on-line. The on-line solution helps shoppers find the best prices.

A data engineering platform can help retailers make sense of their data to optimize merchandising tactics, personalize the in-store experience [9].

Retailers have a big competitive advantage because of data engineering. The data sources help sellers increase the market reach. The big data may be retained from: websites, mobile apps, cameras, selling systems, supply chain systems. It is now very simple to test and quantify the advertising techniques used (somebody can view and calibrate the return out of each tactic). They can now really easily personalize in-store services like never before.

It is now easy and efficient to use databases for the increase of conversion rates. The price for this is better and more point-driven.

Nowadays, customers expect the firms to provide information they need to be more driven to buy the right product. This makes marketers understand and connect with customers. This is why big data engineering technologies can bring in one place information and then analyze it for a greater interest.

The retailers may now find out what's happening across every step in the customer journey, who are their best customers and the info on reaching them.
This is why IBM's Institute for Business Value reported that 62% of retailers report that the use of information (including big data) and analytic is creating a competitive advantage for their organizations [10].

Smart mobile devices and smartphone have made mobile commerce spread very quickly, creating a revolution for the e-business. The search costs in e-commerce puts business under increased price competition raising the profits [11].

Smart phones are now important big data sources, helping customer relationship management improve. Retailers can use data to make better decisions improving their performance [12]. The way to do that is better data exploitation.

VI. Privacy preservation and GDPR in data analytics

There are various organizations (hospitals, banks, e-commerce, retail and supply chain) that generate huge amounts of data [13]. Social media and smart mobile devices generate that also every day, every hour, every minute, every second. This data may be analyzed to help better support decision making. This is where GDPR (General Data Protection Regulaton) comes in. It is a regulation approved by the European Parliament in April 2016. This European Regulation is already enforced and being applied in the whole European Union.
Companies process in a way or another personal data. It doesn't really matter if they do that for their own interest or for other companies. Any entity works on this kind of data. You may have info on your employees, on your clients (for marketing), on sensible matters (for your clients: health, fiscal or juridic data). Although the EU made this law, it implies non-EU firms also, firms that sell in the EU and have EU customers.

GDPR comes up and enforces:
- DPO (Data Protection Officer) – a new and important function.
- new rules for consent
- new rights for the data subject, portability of personal data
- transparency extended
- non-compliance may come with very high fines.

The privacy threats when thinking at data analytics weights on the shoulder of the data holder. Data holder is the entity that holds the data and can be: mobile apps, e commerce web pages, banks, hospitals, social networking applications, and many more. The data holder is responsible for the privacy of the data [13].

The GDPR may be easily understood if one should overview Figure 5 that presents the 6 big GDPR myths.

Figure 5. GDPR False affirmations (myths) [14]

Because of big digital media databases, the EU came up with GDPR. Many businesses that found the honey of this kind of databases grew. This kind of evolution clearly came up with many good things for the firms and also to their costumers. Still, the buyers can encounter difficulties if the seller does not have the best interest for their clients. So, although GDPR is redundant, it is very much needed.

CONCLUSIONS

Data engineering can bring together and explore massive sets of structured and unstructured data, data that may uncover hidden patterns. It can make new correlations, predict trends and help businesses grow as never before. All of this may happen if big media databases are used the right way. If the right way is convenient for only one part of the business deal, it can not be right. This is why ethics are so important and this is why only a combination of good services or products and the good will for the costumers may become a success story on the long run. In this direction GDPR helps and businesses may find the best way to accomplish the plan.

BIBLIOGRAPHY

[1]   https://dataconomy.com/2018/02/5-ways-big-data-analytics-will-impact-e-commerce-2018/
[2]   https://www.oracle.com/big-data/guide/what-is-big-data.html
[3]   https://channels.theinnovationenterprise.com/articles/why-big-data-servers-are-different-from-regular-hosting
[4]   https://dataconomy.com/2018/02/5-ways-big-data-analytics-will-impact-e-commerce-2018/
[5]   JeffreyLush.net
[6]   https://axiomq.com/blog/8-largest-e-commerce-companies-in-the-world/
[7]   *What is cloud computing? AWS* - https://www.youtube.com/watch?v=jOhbTAU4OPI
[8]   U.S. Patent, Jul. 24, 2018, US 10,033,597 B2
[9]   Five   Big   Erin   Hitchcock,   *Data   Use   Cases   for   Retail*,   2018, https://www.datameer.com/five-big-data-use-cases-retail/

[10] Analytics: The real-world use of big data in retail, How innovative retailers extract value from uncertain data – ibm.com

[11] Big Data in Mobile Commerce: Customers Relationship Management, Muhammad Anshari and Syamimi Ariff Lim, Exploring the Convergence of Big Data and the Internet of Things, DOI: 10.4018/978-1-5225-2947-7.ch006

[12] Using Data and Big Data in Retailing, Marshall Fisher, Ananth Raman, https://doi.org/10.1111/poms.12846

[13] Privacy preservation techniques in big data analytics: a survey, P. Ram Mohan RaoEmail, S. Murali KrishnaA. P. Siva Kumar, https://doi.org/10.1186/s40537-018-0141-8

[14] www.gdprcoalition.ie