

Big Data Storage, Collection, & Protection with Islamic Perspective

Ibrahim Nasreldin Ibrahim Ahmed, Mohamad Fauzan Noordin

Kulliyyah of information and Communication Technology, International Islamic University Malaysia, 53100, Selangor, Malaysia

Abstract: *The paper examines the concept of Big Data storage, collection and protection and its Islamic perspective. It also supports the idea of scholars with evidence from the Noble Quran in terms of use. The paper asserts that Big Data in ICT is view in different aspects. These aspects were extensively dwelled with supportive verses from the Holy Quran. The paper concludes by ensuring proper back up is done by all and promoting the good use of Big Data.*

Keywords: Big Data, Islamic Perspective, ICT and Quran, collection protection storage of Big Data.

1. Introduction

Big data may refer to exponential growth and availability of data, both structured and unstructured while also as the technology including tools and processes that an organization may require in handling the large amounts of data and storage facilities.

There are many types of data in Big Data such as:

- Traditional Enterprise Data – this would include customer information which is obtained from CRM systems, transactional ERP data, web store transactions, and general ledger data.
- Machine Generated/Sensor Data – this would include weblogs, smart meters, manufacturing sensors, equipment logs trading systems data.
- Social Data – this would include micro-blogging sites, customer feedback streams, and social media platforms like Google Plus

In this paper big data collection, infrastructure and protection will be reviewed with accordance to the Noble Quran as well as looking at the Islamic perspective.

2. Collection of Big Data

An enterprise developing a throughout and insightful understanding of a business needs to distill and analyze big data in combination with traditional enterprise data. This leads to a stronger competitive position with enhanced productivity and a greater innovation impacting the enterprise significantly.

(وَإِنَّ عَلَيْكُمْ لَحَافِظِينَ - كِرَامًا كَاتِبِينَ - يَعْلَمُونَ مَا تَفْعَلُونَ)

Translation:

But verily, over you (are appointed angels in charge of mankind) to watch you, Kiraman (Honorable) Katibin – writing down (your deeds), they know all that you do.) (82:10-12)

When looking at the Ayat from Surah Al-Infithar (82:10-12), it is noticed that every action is recorded and this in turn is used in the collection of big data, for example; retailers are most of the time aware of who usually buys their product thru the use of social medias and web log files from within their ecommerce site, this in turn would be used to also help them

in understanding who didn't buy what product and why they chose not to buy it thru the use of the information that was not provided to them, this is carried out by viewing the logs provided when a customer login to the ecommerce site thus allowing a much more effective micro customer segmentation and targeted marketing campaigns, which could then also allow for an improvement in the supply chain efficiencies by setting a more accurate demand planning.(1)

Another example is social media sites like Google+, Facebook and Myspace which simply exists due to the use of big data. The business model used by these social media website requires that a personalized experience on the web is met in order to capture and use all available data about a user or member signed up into their website.

In a sense everything is being recorded of what the user is or has done throughout the social media website and is then used as valuable data in order to draw a pattern for advertisement towards the user that might interest them.

3. Storage of Big Data

The range of big data storage technology consists of solid state, hard disk, and tape type tiered storage, these storages help in extreme-scale data environments by their flexibility, capacity, performance and cost effective solutions.

Hierarchical Storage Management (HSM), is used due to the inherent complexity of tiered storage, which is a software that provides transparent data access across multiple storage subsystems. Virtualization of the storage components from the application, delivering automated data management and the user viewpoint is its principle goals while also encompassing a set of tools used by the system administrators in implementation of site specific data and storage policies.(2)

As we can see from this Ayat (Surah Al-Kahf 18: 49):

وَوَضِعَ الْكُتُبَ فَتَرَى الْمُجْرِمِينَ مُمْسِكِينَ بِمَا فِيهِمْ وَيَقُولُونَ يَا وَيْلَتَنَا مَا لَ هَذَا الْكِتَابِ لَا يُغَادِرُ صَغِيرَةً وَلَا كَبِيرَةً إِلَّا أَحْصَاهَا وَوَجَدُوا مَا عَمِلُوا حَاضِرًا ۗ وَلَا يَظُنُّمْ رَبُّكَ أَحَدًا (٤٩)

Translation:

And the record [of deeds] will be placed [open], and you will see the criminals fearful of that within it, and they will say, "Oh, woe to us! What is this book that leaves nothing small or great except that it has enumerated it?" And they will find what they did present [before them]. And your Lord does injustice to no one.

The Ayat tells us that in this book or storage every detail is recorded, be it wrong or right and there is no injustice done, which is the equivalent of the system administrators making sure to check the storage for errors or whether the system is running smoothly or needs maintenance.

4. Protection of Big Data

Using traditional backup process may sometimes take weekly or even daily backup due to the large amount of data that needs to be copied and therefore it can be impractical, this is why the infrastructure of big data should always provide assurance for data protection in order to ensure that its investment is appropriately cared for such as a process that continuously copies files from a primary storage.

Flexible is a necessary attribute of a big data infrastructure it is needed in order to provide support to multiple file types as well as client-side operating systems on the front end with multiple storage platforms for the back end.

If file accesses stop, a data protection copy commences which is then moved to a secondary storage tier, a second policy will then mark that primary copy as a candidate for truncation through a certain amount of days. As soon as the capacity thresholds in the primary storage are met the actual truncation of the file in the primary disk tier will be carried out. This will then result in at least one full-time and size copy of each file being stored on long term archive media, provided the file has become inactive. It is then possible to restore a copy on the primary storage from the archive when accessed again, thru maintaining the copy for as long as possible.(3)

According to the ayat (Surah Fuṣṣilat 41:42):

لَا يَأْتِيهِ الْبَاطِلُ مِنْ بَيْنِ يَدَيْهِ وَلَا مِنْ خَلْفِهِ تَتَنَزَّلُ مِنْ حَكِيمٍ حَمِيدٍ

Translation:

Falsehood cannot approach it from before it or from behind it; [it is] a revelation from a [Lord who is] Wise and Praiseworthy.

Data protection has been created such a way that it is impossible for falsehood to breach the data storage.

5. Islamic Perspective of Big Data

In Surah Al-Fatiha, the first ayat starts off with:

([All] praise is [due] to Allah, Lord of the worlds), meaning that not only in reality but this extends to the virtual world, with that in mind as a Muslim it is essential to apply the Shariah rules in the virtual world as well. This means we must follow the tawhidic paradigm in any conditions. For

instance to not take something without permission, in the virtual world this can be applied by making sure to have a terms and condition in which it is stated what data is being used in the Big Data collection. It is a must for us to adhere to the Islamic principles anytime, everywhere. Never to be used with the intention to violate the privacy of others. And to always remember that everything you do is being recorded.(4)

6. Conclusion

Big Data just like any other data is used in order to give an advantage to an enterprise by the gathering of data carried out, The Noble Quran has taught us many things we can use especially when it comes to Information Technology, however amongst everything we know today the most important and emphasized on is the backing up of Data, without this data can easily become distorted or tampered with by natural cause or 3rd party cause and therefore making sure to keep the data safe in such a way that it cannot be corrupted has become a common practice by almost all companies. It is essential to identify and protect the primary storage thru the different types of big data storages provided, and to not repeat any past mistakes of failed storages.

References

- [1] Jean-Pierre Dijcks (June 2013), Oracle: Big Data for the Enterprise, World Headquarters 500 Oracle Parkway, Redwood Shores, CA 94065, U.S.A.
- [2] Michael Feldman, (November 2013), The Big Data Challenge: Intelligent Tiered Storage at Scale, White paper, Actionable Market Intelligence for High Performance Computing, Intersect 360 Research
- [3] Eric Slack, (May 2012), Storage Infrastructures for Big Data Workflow, Quantum
- [4] Noordin, M. F., (2013), ICT and Islam, International Islamic University Malaysia. Kuala Lumpur: IIUM Press
- [5] <http://quran.com>

Author Profile



Ibrahim Nasreldin Ibrahim Ahmed received his B.Sc. (Computer Science) from the University of Greenwich, U.K, and is currently a student in IIUM for Masters in I.T.



Mohamad Fauzan Noordin received his Ph.D. from the University of Wales, U.K.; MBA from the Central Missouri State University, U.S.A.; and B.Sc. (Computer Science) from the University of Missouri Kansas City, U.S.A. Currently.