Volume 5, Issue 8, August 2015 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

Handling Big Data Issues in Mobile Networking and Computing Data

Navya H Ajjar
Assistant Professor, MCA
AIMIT, Bangalore, India

Nagalambika Swami
Assistant Professor, MCA
EWIT, Bangalore, India

Abstract: ***Big data is the term for data sets so large and complicated that it becomes difficult to process using traditional data management tools or processing applications. Mobile networking is becoming a more and more important counterpart of traditional Internet and big data. The mobile networking is becoming larger and larger due to releasing of hundreds of thousands of cell phones and pads. Moreover, the evolution of cellular network has enables mobile devices to be connected fast and reliably, Handling streaming high-rate data in relational models remains as an open problem, statistical analysis and machine learning algorithms for big data need to be more robust and easier to use and an ecosystem-alike mechanism should be built around the devised big data algorithms such that data Management and usage can evolve sitting on top of the researching algorithms***

Keywords: BigData, Data Management, mobile Networking, swiftstack

I. INTRODUCTION

Mobile phone technology is growing at incredibly faster rate, the fastest growing industry in the history of mankind and in science has to be mobile phone industry. Frequently introduction of new computerized phone in the market with latest software and accessories has surprised the people, which they never dreamt and also Capability of a mobile phone to play games and access to Internet brought an impact on the industry - then immediately came inbuilt computerized and highly sensitive camera. Capturing a photo in the mobile phone was a surprise to its users. In the past few years smartphones remarkably started to carry sensors like GPS, accelerometer, gyroscope, microphone, camera and Bluetooth. Related application and service offering covers e.g. information search, entertainment or healthcare

This new research, whose findings are clearly important to society at large, has been often conducted within corporations that historically have had access to these data types, including telecom operators [7] or Internet companies [5], or through granted data access to academics in highly restricted forms [6]. Some initiatives, like [4], have collected publicly available but in some extent limited data sets together. Clearly, government and corporate regulations for privacy and data protection play a fundamental and necessary role in protecting all sensitive aspects of mobile data. From the research perspective, this also implies that mobile data resources are scarce and often not ecologically valid to test scientific hypotheses related to real-life behavior.

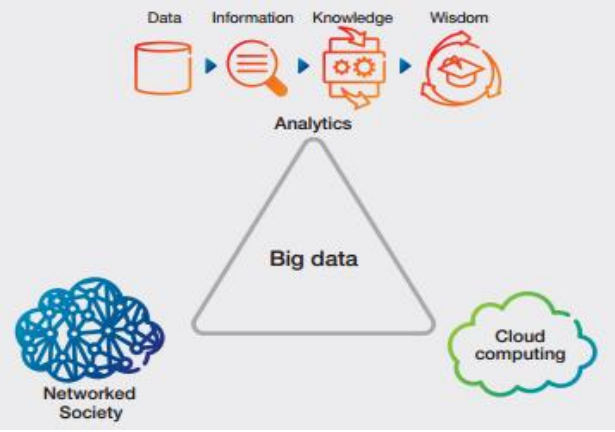

Figure1: efficient extraction of value from data

Big-data technologies are usually engineered from the bottom up with two things in mind: scale and availability. Consequently, most solutions are distributed in nature and introduce new programming models for working with large volumes of data. Because most of the legacy database models cannot be effectively used for big data and consistency it provides is acceptable for many applications. As a result, requirements for consistency need more careful analysis during the design of applications that run on top of big-data platforms.

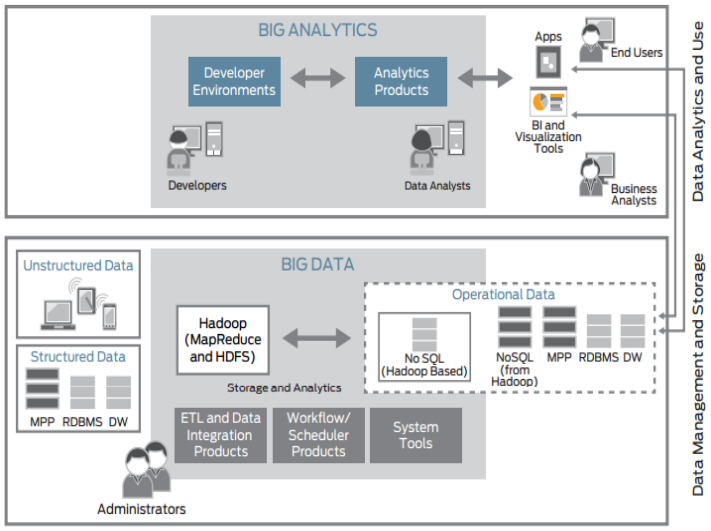

Figure 2: Structural view of handling data

Big data mirrors the growth in content and data source, as well as the pervasiveness of technology in our everyday lives. As more and more of what we do is both connected to and often empowered by a network—and the devices that we connect to are themselves powered by an array of sensors—we should expect that the ongoing stream of data will grow. Within data centers, every node (servers, storage, and applications) generates a tremendous number of log files and isolated data streams that also can be collected, collated, and analyzed. With storage costs dropping, the cost associated with saving and leveraging even the most mundane data becomes a nonissue.

As reviewed in Jupiter.net[1] predominant traffic pattern has changed dramatically. Big data represents just the latest application environment to drive this architectural shift. As data becomes more horizontally scaled and distributed throughout network nodes, traffic between server and storage nodes has become significantly greater than between servers and end users. Interestingly enough, the data itself can be generated by servers, applications, or storage environments as opposed to an external source (system log files, for example). This machine-to-machine network traffic and data sharing is often referred to as east-west traffic. Building a data center optimized to provide high-speed connections optimized for east-west traffic is critical in developing scalable, high-performance big data implementations

II. OBJECTIVES

Hashing was applied to a variety of text entries appearing in the MDC data, including Bluetooth names, WLAN network identifiers (SSID), calendar titles and event locations, first names and last names in the contact lists and media filenames (such as pictures). For anonymization of the WLAN and Bluetooth MAC addresses, we split them into two parts. First, the MAC pre- fix, also known as the "Organizationally Unique Identifier (OUI)

Phone numbers appearing in the call logs and in contact lists were also split in two parts. First, the number prefix, which contains the country and region/mobile operator codes, was left as clear text. Then, the rest of the phone number was hashed as described above. Also the cell ID and the location area code (LAC) of the cellular networks were anonymized using the hashing technique as described above

Several java language binding are available, but the jclouds[8]seems to be the favorites support open stack swift and nova and amazon S3 object storage.

Watermarking release of the MDC data set to a large community of researchers motivated an additional step in which each distributed copy of the data set was watermarked individually in order to identify it if necessary. The watermarking process introduced negligible alteration of the data that did not interfere with the results. Operators everywhere look to minimize risk, maximize profitability, and achieve competitive advantage. Tier One operators ramp up use of Big Data and Network Analytics to improve network operations, and operators of every size

III. CONCLUSION

Conclusion The data available to operators through their networks presents them with an opportunity and a business-intelligence edge over other players. As is often the case, with opportunity comes challenge, and for big data this challenge comprises the volume and diversity of the data – and the fact that both are expected to grow substantially in the next few years. The value of the information in the data is significant, but the costs involved in obtaining it using current technology are inhibitive. Consequently, big-data technology is an important part of the puzzle for operators wanting to leverage value from the large volumes of data in their possession in a cost-efficient way. Applying big-data technologies has the side effect of transferring some complexity from the database to the application. Within the big-data sphere, Ericsson holds a unique position by being among the few companies that truly understand networks. Ericsson can provide the glue needed to convert vast volumes of raw data into information and knowledge that can be monetized. At Ericsson Research considerable knowledge has been gained through the development of big-data technology proofs-of-concept, expanding our existing competence to build large distributed systems with high availability to include big-data technologies.and also understand swift.

REFERENCES

[1] www.juniper.net.
[2] Survey of Recent Research Progress and Issues in Big Data
[3] http://www.idiap.ch/~gatica/publications/LaurilaGaticaEtal-mdc12.pdf
[4] http://crawdad.cs.dartmouth.edu/.
[5] L. Backstrom, E. Sun, and C. Marlow. Find me if you can: improving geographical prediction with social and spatial proximity. In Proc. World Wide Web Conf. (WWW), Apr. 2010
[6] M. C. Gonzalez, C. A. Hidalgo, and A.-L. Barabasi. Understanding individual human mobility patterns. Nature, 453(7196):779–782, June 2008
[7] S. Isaacman, R. Becker, R. C´aceres, S. Kobourov, M. Martonosi, J. Rowland, and A. Varshavsky. Identifying important places in people's lives from cellular network data. In Proc. Int. Conf. on Pervasive Computing, San Francisco, Jun. 2011.
[8] https://github.com/jclouds

