Big Data in Capital Markets

Manpreet Singh
Associate technology
Sapient Global Markets

ABSTRACT
In this modern era of technology, we are witnessing a wave of technological advancements. These advancements are as diverse in the areas like collection, storage, aggregation, processing and analysis of financial data. Moreover, with the advent of the Internet of things," there is an exponential rise in huge volumes of unstructured data from several different sources. In this regard, Capital markets are one of many industries which have voracious appetite for big data; they collect and analyze data to mitigate risk, understand their customers, and generate more profitability. Currently, the financial markets are challenged by exponentially increasing volumes of data, inconsistent data quality, and heightened regulatory scrutiny. Remedy to this problem is Big data. Focus of this paper is to analyze the growing need of Big Data technology in financial domain, especially in Capital Markets. Why Capital market sector is much interested in this technology, what all benefits could be incurred? It also highlights the existing modules where big data is already in use, areas where implementation is ongoing and also the hindrances in the path of implementation of the same.

Key Terms- **Big Data, Hadoop, Map Reduce**

1. INTRODUCTION
Since the last decade, the concept of big data has been applied to a broad area of diverse industries but capital markets have been comparatively slow to adopt changes due to security as well finance handling matters. Within the domain of financial services, big data has gained huge traction in retail banking especially; because of the increasing lust of financial companies to profile the customers and their clients in a manner like Google, Facebook and other such companies who have voluminous data. Giants like Google, Amazon were few to adopt big data at early stages. Whereas, the institutional aspect of the capital markets has traditionally far more customer holding stickiness; hence there is less flexibility to apply big data strategies to avoid any financial loses. However, Over the last decade big data has begun to influence in various areas of the capital markets, listing a few would be risk analysis, analysis for sentiments in trading and management, and surveillance regarding market trends and issues.

These findings not only identify challenging opportunities big data initiative both for current as well as future scenarios, but also analyze its relevant benefits potentially of adopting big data strategies and technologies in the respective capital markets community. Finance domain has always evolved in parallel with the advancements in technology. Adopting a new technology is might be thought of as gaining competitive advantage in the market. It is an important requirement for many financial institutions in this industry. For instance, utilizing IT services in capital markets trading automation is one of the major factors for the market efficiency and competitiveness as time has a huge impact on the costs incurred by financial institutions. Systems used by financial institutions are proprietary thus creating hindrances in the interoperability between potential business partners decreasing. While this was not a major problem for decades, the recent widespread use and acceptance of the Internet results in communication between geographically dispersed entities should no longer be a challenge. Businesses can now share information in an established environment, take part in value chains and involve business processes from other business entities.

2. DEFINING BIG DATA
Big data has been a much misused and misunderstood term within the financial services industry for some time, applied to diversity of things from traditional relational databases (RDBMS) to web-based sentiment analysis tools. It is often not marked as a standalone concept, but when it is discussed in a theoretical manner, it is often poorly defined. What exactly constitutes as "big" is also somewhat a heated argument; are we talking in terabytes, petabytes, or exabytes? Does the strategy only has to be focused on a project related to a huge volume of data, or can other factors determine its application?

The concept of "big" in a scientific or retail environment is somewhat different from what is taken to be big in a capital markets context. The profiling of customers in the retail markets, for example, involves analytics of unstructured data collected from a vast variety of sources like a broad range of social media, on the other hand, capital markets tend to deal mostly in structured data sets from a compact set of defined sources such as market infrastructures, market data vendors and their respective counterparties. Some unstructured data sets have, however, become important to capital markets in areas like sentiment analysis and market surveillance for profiling certain trends and activities in the market and within firms, but these have not traditionally been the data sets of primary importance to the business. Thus, Big data is a concern now.

3. CAPITAL MARKETS
Having on-demand access to the right data and the ability to analyze it in real-time and at the right time are essential for intelligent decision-making investments. One of the high values big data offers is its looping ability to process real time feedback, react to it, analyze the results, and continue the loop to iterate continuous optimization. [3] This loopback feature will allow testing large amounts of data in real time, using a variety of data from several sources, to be conducted in shorter spans of time. In capital markets, the speed at which data can be tested, analyzed, and applied will likely provide a competitive edge, specifically in the areas of prediction and forecasting.

4. EVOLVING BIG DATA TECHNOLOGIES IN CAPITAL MARKETS

1. Data grids- It uses distributed caching to control data which is large volumes of data across a servers network. In-memory data grids are one way to handle the velocity, variety and volume of big data. By moving data into memory and distributing it across multiple servers, it aims for faster and smooth access to data with improved scalability and better data analysis. In respect with Capital Markets, it serves purpose for trade management and analysis.

2. Compute Grids- Compute Grids offer a way of parallelizing processes across multiple servers for handling capacity, efficiency and failure issues. Also orchestrating tasks across the grid to make it more reliable and faster.
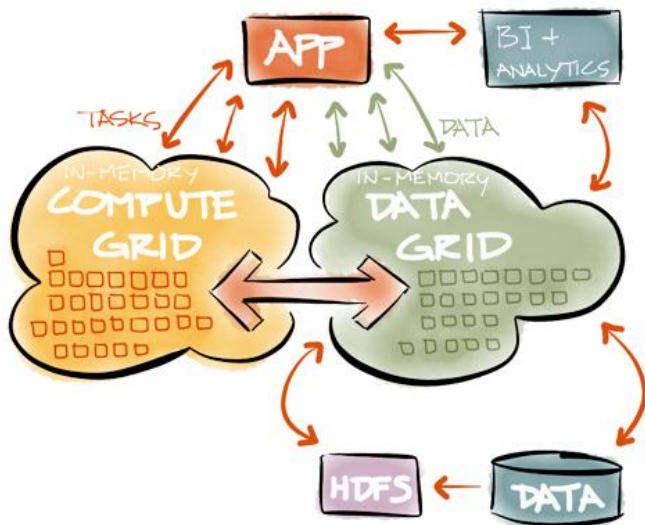

Fig. 1- Grids use in Capital Markets

3. Parallel processors-Parallel Processing involves the coordinated processing of a program which is executed in such a way that multiple independent computers are assigned to it, each with its own Operating system and memory to execute a task by dividing the load

4. In-memory databases- The advantage of in-memory processing is that it delivers the fastest data-retrieval speeds available today, which is alluring to companies facing difficulties with high scale online transactions, forecasting and planning synchronized. [7] Disk-based storage may still be enterprise standard, but since the price of data storage is declining steadily, so memory-intensive architectures will eventually replace slow, mechanical spinning disks.

5. Hadoop- This is a tool used to query the unstructured data which is a major part of Big data analytics.

5. AREA WHERE CAPITAL MARKETS IS USING BIG DATA

1. Financial Data Management and Reference Data Management- Data Storage for historical trading, internal data Management and overall control on reference data is big task for financial domain. Organizations must blend historic and new data to generate insights about delivering better, more cost-effective solution. Advancing this strategic need; creates even more demand for efficient, accurate reference data management. It can be difficult to maintain, handle and process the data from different asset classes coming from a whole host of vendors. Thus, big data provides an aid for the problem.
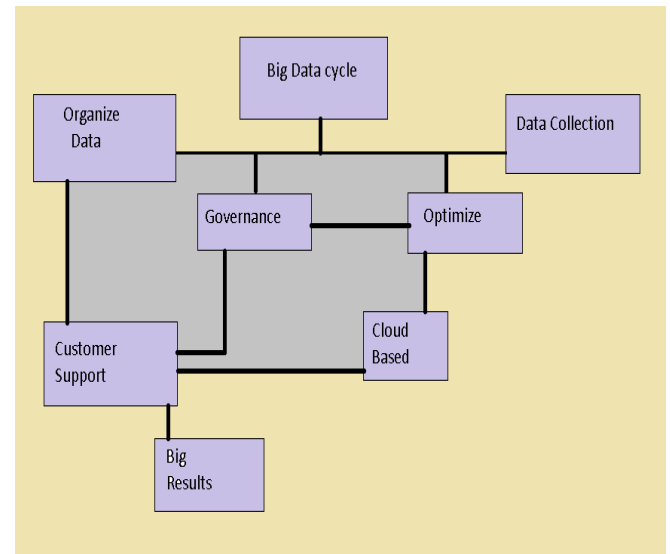

Fig. 2- Financial Data Management

2. Regulation- Financial Services experience a lot of regulatory changes over the previous decade. Capital Markets confront new regulations and are preparing for the upcoming regulations Like EMIR, Dodd Frank (CFTC).

3. Risk Analytics- Risk analysis contain fraud mitigation, anti-money laundering, rogue trading and continuously monitoring asset performance and invest capital to the areas of the business that are delivering a higher risk-adjusted return on capital investment.

4. Trading Analytics – Includes various analytics areas in trading for high frequency trading, pre-trade decision, supporting analytics for frequent trades, sentiment measurement and temporal analytics etc.
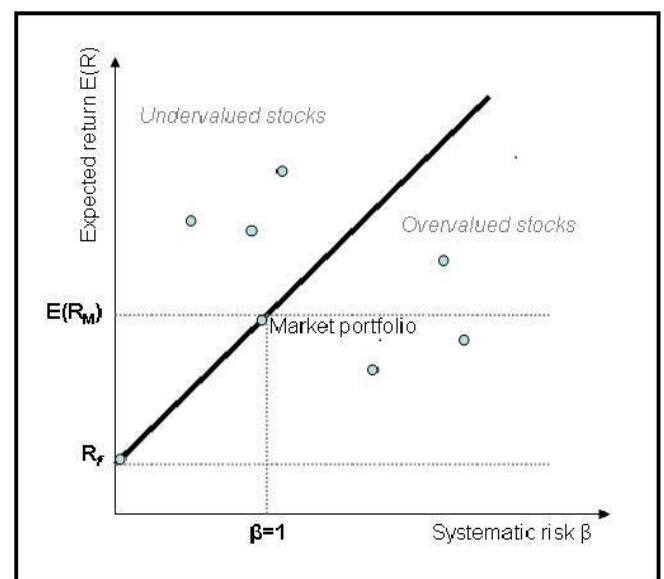

Fig. 3- Trade Analytics

5. Credit Scoring- Credit scoring or Predictive analysis is another area where big data can be useful. Comparing to the older models used to predict the credit worthiness of a company can be supplemented with input and analysis from big data. [5]Since, big data offers a wide variety of information; these variables can be integrated into more traditional credit rating models to identify hidden patterns that can lead to better predictive abilities regarding credit worthiness. This hybrid model would allow a more accurate way to apply credit ratings to companies and governments.

6. Data Tagging- In enterprise range of monitoring and reporting, it's very difficult to match and reconcile trades from various systems built on different technology standards resulting in duplicated, invalid and missed out trades. Data tagging can resolve it by easily by keeping track of all the trades and the events like corporate actions enable regulators defect stress signs early.

6. PITFALLS AND CHALLENGES

One of the biggest reasons for the failure of big data projects applicability in the past has been compatibility issues between business requirements of capital markets and the capabilities of big data. For instance, Hadoop implementations have many a times been unsuccessful; reason being the technology is reliant on offline batch processing through Map Reduce, which is incompatible with real-time analytics. [1]Furthermore, Hadoop closely couples resource management with processing data; therefore prioritizing tasks was not possible when running multiple applications simultaneously. Hadoop is however, aimed at tackling these resource allocation challenges.

Firms also need to change their data storage strategy via the addition on tiered storage architecture, placing the data sets of utmost importance on comparatively fast devices and allowing other sets to be less readily accessible though cheaply stored. The prioritization of these data items is a significant challenge when it is against the background of a functionally of capital markets. With different priorities and budget sensitivities capital markets have a different architecture. There is another component which is a more proactive approach of retaining data, where firms are able to retire and delete data after it has reached the expiration of its retention period.

The intent of the big data project overall also needs to be carefully considered strategically. [1]If an target relevant to the business is not identified, then big data strategies will not deliver returns on investments made. In this end, analysis could result in unpredicted data gaps or because of lack of contextual information. Respondents also have inadequate technical knowledge which becomes an encountered challenge during a big data project; therefore it highlights the need for appropriately skilled staff to implement such a strategy. The biggest issue in this regard is finding employees or consultants that understand the business and the technology.

Data privacy is also a key concern for big data projects especially when it comes to financial markets. This is a challenge that is common with cloud technology implementations. There must be an adequate level of protection and security for commercially sensitive data, safeguarding it from unauthorized access and information should also be kept private by organizations because of legal and regulatory requirements, which may vary from jurisdiction to jurisdiction. Some countries are stingier than others about the treatment and storage of data; hence these factors must be carefully considered before any implementation can take place.

7. THE FUTURE OF BIG DATA

Implementation of Big data in capital markets is a long way to be considered mature; it has yet to reach a stable and turning point in usage within top-tier sell-side firms, even if not across the industry as a whole. Though there are many viable reasons for investment, these tend to be focused on specific use cases instead of an over-arching big data strategy.[10] Those that are currently aware of the benefits of big data and have successfully realized some of these internally despite facing hurdles, are certainly keen to extend their usage of these strategies and technologies. These implementations will, however, likely remain in growing stage for the near future.

It shows that 69% of frontrunner respondents expect to start actively on big data technology, or for the first time in the case of those that have yet to invest. A laconic 13% feel they will not evaluate big data technology in the next two years and 18% yet are not certain if their firm has any definite plans to conduct such an exercise. Given this respondent group is an educated subset of the Respondents have insufficient technical knowledge as the most commonly encountered challenge during a market at large, most firms in the capital markets sector are likely to be in the "not likely" and "don't know" categories for some time to come, though this will change as more use cases for big data become prevalent in the market.

8. CONCLUSION

Since, big data is new in capital markets though, firms have acted to be early adopters of the new technology but they often do not have comprehensive big data strategies implemented for proper efficient working. Only a few banks and hedge funds have been the frontrunners in initiative of implementing big data strategies and technologies. Also, these deployments are restricted to very specific areas such as trade analytics, risk modeling and market surveillance.

It could be concluded that education and viable resolutions with future awareness will attract industries to adopt the technology. When more capital market related scenarios for big data become prevalent in the market, firms will become more comfortable hence flexible with these strategies. A changing point in adoption will need firms to get a better hold on what comprises a big data strategy and how efficiently such an approach can deliver return on investment timely.

Big data usage reflects the steeply rising importance of functional data management. These functions in financial institutions are in a state of continuing evolution because of factors like regulators and market infrastructure changes or an increase in requirements related to transparencies. Firms are now divergent towards data from a business value perspective rather than solely as a control function, which enhances the importance of implementing big data.

Firms are interested in insights, response speed, and also the future scalability on implementing big data. The most important aspect to respondent firms for big data deployment signal is its actionable information. Insights are equally secured with scalability for increases in data volume for future. Delivering reasonable return on investments in key areas will attract further investment in big data across other branches of the financial business domain.

Even though many recent deployments may be in the front office however, the future will witness more middle as well as back office implementations of big data. Investments in big data are largely divergent towards revenue generating opportunities, but in future there would be much more focus on retaining clients, functional support and other such opportunities. Higher volumes of historical data storage requirements will force firms to examine and compare the relative cost of implementing big data to current database technologies.

9. REFERENCES

[1] M. Aitken, A. Frino, E. Jarnecic, M. McCorry, R. Segara, and R. Winn. The microstructure of australian stock ex- change: An introduction. Securities Industry Research Cen- tre of Asia-Pacific (SIRCA), 1997.

[2] T. Andrews, F. Curbera, H. Dholakia, Y. Goland, J. Klein, F. Leymann, K. Lui, D. R. D. Smith, S. Thatte, I. Trick- ovic, and S. weerawarana. Business process execu- tion language for web services (BPEL4WS) specification. http://www.siebel.com/bpel, 2003

[3] T.-H. A. Cheung and S. Y.-H. Wu. Trade data service for capital markets. Honors, school of Computer Science and Engineering, UNSW, Nov 2003.

[4] U. Wahli, M. Drobnic, C. Gerber, G. G. Ochoa, and M. Schramm. WebSphere Version 5 Web Services Hand-book. IBM Redbooks, 1st edition, March 2003.

[5] Trading technology survey of exchange technologies. http://www.tradingTechnology.com, 2003.

[6] Capital market cooperative research centre (CMCRC). http://www.cmcrc.com.

[7] F. A. Rabhi and B. Benatallah. An integrated service ar-chitecture for managing capital market systems. IEEE Net- works, 16:15–19, 2002.

[8] Daniel C. Quan and Sheridan Titman, "Do real estate prices and stock prices move together? an international analysis," Real Estate Economics, 1999, 27, 2, pp.183-207.

[9] Jay C. Hartzell, Jarl G. Kallberg, and Crocker H. Liu, "The role of the underlying real asset market in REIT IPOs," Real Estate Economics, Spring 2005, 33, 1, pp.27-50.

[10] R. Blundell and S. Bond, "Initial conditions and moment restrictions in dynamic panel data models," Journal of Econometrics, Vol.87, 1998, pp.115–143.