

Research paper on assumption1: machines can be reliable

[Business](#), [Company](#)



Management Information System

- According to Mark Hurd, why does his typical customer's IT budget increase by two to three percent per year (even when his typical customer does not do any innovation)? (5 points)

There are tons of changes the application undergoes in its structure. The companies are trying to update the old applications to be used with today's gadgets and cloud as the users today are not limited to, system based apps, but use a variety of products like tablets, smartphones and tabs. Clouds have become modern day's database and means to communicate, circulate and publish information.

Some applications are 21, 22 or 23 years old. On an average the applications are 23 year old which CIO's are trying to update . These applications are not built for the cloud or mobile access so are incompatible with modern day gadgets . Since the applications are undergoing such updates, it requires an investment. Apart from this there has been a huge inclination towards data in present application age and customers' demands are also increasing . All this factors contribute towards price increase.

- According to the video, forty three percent of workforces in USA will retire in the next 10 years. Why does this matter, according to Mark Hurd?

After the forty three percent retire in the next 10 years, this work force will be replaced by another forty three percent workforce in the future. This is a matter of concern because the data that will be stored by the next forty three percent work force would be dramatically different from the one that retired . There are two main agenda that company's need to take care of.

Firstly, the company needs to get access to the data that needs to be stored

<https://assignbuster.com/research-paper-on-assumption1-machines-can-be-reliable/>

as well as the customers. Secondly, company needs to reduce costs as the maintenance cost would increase.

- The video illustrate two examples where big data are collected but not utilized. Describe the two examples.

Ans. The data getting generated in this world is increasing exponentially. Large amount of data is captured by organizations that work in financial, logistics and health sectors. Large social sites are also generating digital materials. Computers are now able to get meaningful data from videos or still images. More smart gadgets are getting developed that are able to connect to the internet and hence online marketing has expanded. And finally, several areas of scientific advancement have started to generate large quantities of data that has multiplied recently. Consequently, the data that resides with organizations are mostly ignored. Most of the gathered data remains unprocessed and are exhausted.

First example related to this is the data accumulated by the loyalty cards given by retailers that remains unprocessed.. Second example can be the video data captured in the hospitals for healthcare purposes. The video data recorded during surgeries in the hospitals are deleted after a few weeks. .

- What are two main components of Hadoop? (5 points)

Hadoop is one of the leading technology of big data. It is available with open source and its software library offers reliable and scalable computing platform for analytics. Many data pioneers use Hadoop. For example, around 1 billion personalized recommendation are stored by LinkedIn each week. The storage and distribution of such large amount of data is done among a cluster of servers. The traditional computers need hardware with high

tolerance. Hadoop fixes the application related problems and can also detect the hardware failures. This gives a continuous service to each computer which could have otherwise failed. There are two components of Hadoop. The first component is Hadoop Distributed File System (HDFS) which has cluster storage with high bandwidth. The second component is MapReduce which is a data processing framework. It is based on Google technology where data sets are distributed on multiple servers. Each server stores the summary of data it is allocated. All of the summary information is aggregated to a reduced state.

- Are there any ways for small companies (who cannot afford internal big data) to use big data tools? How?

For organizations that cannot afford an internal big data infrastructure, solutions can be available with the cloud. Here, the data is not needed to be downloaded for utilization. . For example, Amazon hosts public data sets that contain medical and government related information. Further, quantum computing may improve big data processing. Quantum computers process and store the data by using quantum mechanics and excel in processing unstructured information.

- What are three key differences in the big data movement when it is compared with analytics?

Big data is often characterized by three V's: volume, velocity and variety. Here volume poses as the greatest challenge and the greatest opportunity. As big data can help many organizations to understand people better and allocate resources more effectively. However, traditional computing solutions and databases are un-scalable to handle the data of this magnitude.

Big data velocity also raises a number of issues with the rate at which data is flowing into many organizations are now exceeding their capacity in their IT system .

In addition, users are increasingly demand data to be streamed to them in real time and delivering this can prove to be quite a challenge.

Finally the variety of data types to be processed is becoming increasingly diverse. Gone are the days when data centers had to deal only with documents, financial transactions, stock records and personal files. Today, photographs, audio, video, 3D models, compact simulations, and compact data are being piled onto many corporate data cellar. Many of such data sources are also unstructured and hence not easy to categorize let alone process with traditional computing techniques.

- According to the case, it is estimated that Walmart collects more than 2. 5 petabytes of data every hour from its customer transactions. How many filing cabinets' worth of text are 2. 5 petabytes equivalent to?

The report says, Walmart collects more than 2. 5 petabytes of data from its customers every hour . One petabyte is equal to one quadrillion bytes. This is equal to about 20 million filing cabinets' worth of text.

This means, 2. 5 petabytes of data is equal to 50 million filing cabinets' worth of text.

- How did MIT Media Lab estimate Macy's sales on Black Friday? (5 points)

MIT media Lab estimated Macy's sales on Black Friday by using the location data from mobile phones to know how many people were there in Macy's parking lot, which was the start of the Christmas shopping season in the United States . This helped in estimation of sales on that critical day even

before it was recorded by Macy's itself.

- Erik and Lynn Wu's prediction about housing-price changes in metropolitan areas across the United States proved more accurate than the official one from the National Association of Realtors. What data did they use?

Erik and Lynn Wu had no special knowledge of the housing market. Yet, they started researching about it through the available web search data. They reasoned that the real-time search data would be able to forecast the near time housing market prices. They were proved right and in fact their research was more accurate than the official report of the National Association of Realtors that was based on a complex model of the relatively slow changing historical data .

- How could researchers at the Johns Hopkins School of Medicine predict surges in flu-related emergency room visits a week before warnings came from the Centers of Disease Control?

The Researchers at the Johns Hopkins School of Medicine were able to predict the surges in flu related emergency rooms a week before the arrival of the warning through the data they found on Google Flu Trends, which is a free and publically available . It had the aggregation of the relevant search data related to this which helped them predict the warning early.

- Provide evidence that using big data intelligently will improve business performance.

For example, a major U. S airline company noticed there were at least 10 minutes gap in the actual and estimated arrival time of its planes and 30% had a gap of 5 minutes. So, it turned to PASSUR Aerospace for help which provided them with their service RightETA.

The service calculated arrival times based on data about weather, flight schedules, feeds from passive radar stations etc. This helped in the elimination of gaps between the planes and hence improved the airline's performance.

- PASSUR developed a service RightETA. What does ETA stand for?

ETA stands for Estimated Time of Arrival. PASSUR provided RightETA service for a U. S based Airlines Company which had gaps between estimated and actual arrival times of its planes . It used information about the weather, flight schedules, and feed from radar stations etc. to collect data about every plane in the local sky. Every 4. 6 seconds, it collected vast amount of information about the planes it witnessed. This yielded huge amount of digital data. This collected to form a multidimensional information body and helped in sophisticated pattern analysis and pattern matching. This eliminated the gaps that occurred between plane arrivals and saved several million dollars a year.

- How was Sears able to reduce the cycle time to generate personalized promotions from 8 weeks to one week?

Sears needed customer data for promotional purposes to increase the number of customers and generate greater value. The collection process took eight weeks since the data was fragmented with data warehouses of different brands . So, Seard Holding turned to technological practices of Big Data and set up a Hadoop cluster. This comprised of a group of simple servers coordinated by Hadoop framework. The data collected from different brands' warehouses were collected here and analyzed for promotions. This reduced complexities and the time consumed got reduced from eight weeks

to one.

- What does "HIPPO" stand for?

HIPPO stands for Highest Paid Person's Opinion. Since decision making is the most critical part of any business, the important decisions are left to the top most people of the organization. This concept is based on the theory that experience and intuition are more reliable for business decision as compared to collected data. For very important decisions, the high up people of the organizations or expensive outsiders are brought in as they are experts in track records. They make these important decisions based on their experiences and so many big data community companies rely on HIPPO for decisions.

- What are two techniques that executives can employ if they are interested in leading a big data transition?

The two techniques that executives can employ while leading a big data transition are, by asking what does data say and allowing data to overrule. Asking what data says is important as it helps in following up more questions like where the data came from and what kinds of analysis were done and how confident are we with the results.

Secondly, by allowing data to overrule. Senior executives have conceded when data has disproved a hunch.

- Illustrate at least two barriers to the success of big data implementation.

One of the barriers in implementation of Big Data can be lack of effective leadership. The company cannot survive without an effective leadership that devises strategies for the organization and so it needs a leader who has clear goals, creative thinking, understands market well and can spot great

opportunities.

Another challenge can be company's culture. This includes following ethical practices in data management. If the people included in the implementation work don't understand the importance of ethics, it can become a barrier.

- The speaker challenges three conventional assumptions. What are they?

One can't rely on machines. If a machine fails once on every four-five years, it can be pretty good and the machine can last more than the expected lifetime. But if one has 100 such machines, the average failure would increase to 4-5 failures a day. If one cuts down that mean times to half, then it would mean at least 10 failures a day. so one will have to be dependent on softwares. The hardware alone cannot be reliable. So by depending on the software reliability for hardware can save a lot of money.

Assumption2: Machines have identities .

Suppose one has different servers meant for different purposes like data storing and data processing. One need to acknowledge that hardware is fundamentally unreliable. It will fail at times so one should not think about them as individuals but as commodities.

Assumption 3: A data set can fit on a single machine .

Data sets are now ranging to be 100s of teratbytes such as in the field of science, biotechnology, etc. large amount of data sets is produced. These data cannot fit on a single machine and this won't change. So one needs to understand that a data is going to expand on many different machines.

- List at least three names of regular enterprises that use Hadoop.

The three regular enterprises that use Hadoop are Amazon and Yahoo and LinkedIn.

LinkedIn uses it for storing 1 billion personalized recommendations every week. Hadoop distributes the storage and processing of large data sets across groups or clusters of server computers, while the traditional large scale computing solutions rely on expenses of the hardware with high tolerance ..

Amazon hosts many public data sets containing government and medical information. Hadoop helps in storing and processing in massively parallel processing of unstructured data .

Yahoo also stores many information using Hadoop that contributed to its online information services .

- The speaker talks about where data come from. List at least two sources of data

Ans. The first source of data among the two sources the speaker has talked about is the users who are extensively using innovative technology such as cloud and internet. As more users have access to these technologies, more data is being produced .

The second source is the storage devices. Since the cost of storage devices like hard disks are falling dramatically, it has become easier to store more and more data everyday rather than losing them .

- Can Hadoop serve data in real time? Is Hadoop a competing product with DB software, such as Oracle?

Hadoop does not process data in real time. Hadoop is not a competing product with database software like Oracle. It is a batch data processing

system. Users cannot interact directly with the Hadoop clusters . It absorbs data from different sources and process them. This data can then be loaded in interactive databases. For example it can generate indexes for interactive search boxes which complete the words or sentences whenever a user types in it.

Part II. Bullwhip Effect

- Explain what the Bullwhip Effect is.

Control of inventory is a key issue in the supply chain management.

Members of a supply chain optimize the use of their inventories by adopting policies and operation procedures. This minimizes investment while keeping customer services high. Uncertainty is inherent in all supply chains due to variability in demand, lead times, breakdown of machines and local politics. So companies often keep inventories in buffer called safety stock. It has been observed in supply chains that small variations in customers' demand result into large variations in demand as we move up the chain. This phenomenon is called the Bullwhip Effect .

- Explain why Volvo manufacturing department believed that consumers had started to like green cars in the mid-1990s.

In the mid 1990s, the Swedish car manufacturer, Volvo found itself with excessive stocks of green cars. In order to move them, the sales and marketing departments began offering attractions and special deals, so the green cars started to sell. However, nobody told the manufacturing department about these promotions. By noting the increase in sales, it read it as a sign that consumers had started to like green cars and ramped up the

production .

- Discuss how the Bullwhip Effect can be reduced in general and how IT (Information Technology) can be used to reduce the Bullwhip Effect.

For minimizing bullwhip effect it is required to cut several decisions across the supply chain and have greater collaboration among the supply chain partners. Bullwhip effect can be minimized by reducing the number of layers in the supply chain . Recently, several global firms have resorted to this approach and have tried to directly reach the customers through local distribution centres with the help of third party logistics such as Federal Express. Another strategy to reduce the Bullwhip Effect is by reducing the delay of information flow in the supply chain. Use of point of sales data capturing systems Electronic Data Interchange (EDI) and Information Technology can help organizations to cut the delay in the information flow . Supply chain partners should be encouraged to share sales, capacity and inventory data amongst themselves. Suitably modifying the incentive structure can be one method to achieve this objective. For example, Bangalore based CTV manufacturer BPL had linked top distributors and retail outlets through IT and the finished goods inventory level fell dramatically almost immediately.

The Bullwhip effect can be further monitored by reducing the lead time of the business process. This can be done using better business process, information technology and closer working arrangements with logistics providers and distributors .

References

- Daoliang Li, Y. L. (n. d.). Computer and Computing Technologies in Agriculture .
- Economist. (2002, January 31). Managing a supply chain is becoming a bit like rocket science. Retrieved from Economist. com: <http://www.economist.com/node/949105>
- ExplainingComputers. (2012, June 16). Explaining Big Data. Retrieved from Youtube. com: http://www.youtube.com/watch?v=7D1CQ_LOizA
- Hurd, M. (2012, September 30). Oracle Big Data and Innovation: President Mark Hurd. Retrieved from Youtube. com: https://www.youtube.com/watch?v=F6YGZZeG_2M&feature=related
- IV, O. E. (n. d.). smallbusiness. chron. com. Retrieved from smallbusiness. chron. com: <http://smallbusiness.chron.com/reduce-bullwhip-effect-3908.html>
- Jorgwel. (2011, August 28). Hadoop and Big Data 2/6 Processing Petabytes. Retrieved from Youtube. com: <http://www.youtube.com/watch?v=xQOKOl6IKJM&feature=relmfu>
- Jorgwel. (2011, August 28). Hadoop and Big Data 5/6 Ferrari vs Freight Train. Retrieved from Youtube: <http://www.youtube.com/watch?v=-QdCABPyu1k&feature=relmfu>
- Jorgwell. (2011, August 28). Hadoop and Big Data 1/6 Challenging Old Assumptions. Retrieved from Youtube. com: <http://www.youtube.com/watch?v=y8DRKd4SKWo>
- Mahadevan, B. (2010). Operation Management: Theory and Practice. New Delhi: Pearson Education India.

- McAfee, A., & Brynjolfsson, E. (2012, October). Big Data: The Management Revolution. Retrieved from Harvard Business Review: <https://hbr.org/2012/10/big-data-the-management-revolution/ar>
- Moll, J. (2013). The Bullwhip Effect: Analysis of the Causes and Remedies. Amsterdam: University Amsterdam. Retrieved from http://www.few.vu.nl/en/Images/werkstuk-moll_tcm39-354834.pdf