

Teradata

Dhiran R. Karale

Department of Computer Science and Engineering
Jhulelal Institute of Technology, Lonara
Nagpur, India

Rana Syeda

Department of Computer Science and Engineering
Jhulelal Institute of Technology, Lonara
Nagpur, India

Abstract – A database is an organized collection of data. The data is typically organized to model aspects of reality in a way that supports processes requiring information. A database is a very large collection of a big database has no minimum absolute size. Although a very large data base is a database like smaller databases, there are specific challenges in managing a very large database. Formally a "database" refers to a set of related data and the way it is organized. The management system provides various functions that allow entry, storage and retrieval of large quantities of information as well as provide ways to manage how that information is organized. Databases are used to support internal operations of organizations. The relational part comes from entities referencing other entities in what is known as one-to-many relationship, like a traditional hierarchical model, and many-to-many relationship. A database management system (DBMS) is an aggregate of data.

Keywords: - *Big Data, Managing Big Data, Teradata database difference, Shared everything architecture, Shared nothing architecture.*

I. INTRODUCTION

The Teradata is a much more scalable effectiveness relational database management system (RDBMS). It is mostly used to manage large data warehousing operations and used to solve large data warehousing queries. The Teradata is capable of handling large amount of data and complex queries. Or big data, big data that must be used to show high capacity to predict conclusion, with low cost consumption, increase efficiency and enhance decision-making in various fields like traffic control, weather prediction, disaster prevention, finance management, fraud control, improve business transaction, control on national security, education improvement, and health care. They all are shows or stores huge amount of data. Teradata is RDBMS and Oracle is Object RDBMS. Teradata has been mainly designed to OLAP whereas Oracle is for OLTP. But teradata is used for DWH because of its parallelism. The DWH process of constructing and using a data -warehouse. Data warehouse is constructed by integrating data from multiple heterogeneous sources that support analytical reporting structure. Teradata is "Shared Nothing Architecture" where as oracle is "Shared Every Thing Architecture". In teradata millions of instructions parsec, fast

retrieval and better storage. In teradata. Thousands of instructions per sec fast retrieval in oracle. Teradata many bulk load facilities and terabytes of storage. Oracle limited bulk load facilities and GIGA bytes of storage.

Teradata is designed for DWH, (Data warehouse) data mart, internet, e-commerce appliances. There are Teradata utilities to import and export data to and from host-based and client-resident data sources, or data base sources which are generally divided into a several small parts which we may call that data marts or departmental data marts. Teradata Columnar partitions tables by column, row or both. Teradata is a Relational Database Management System. (RDBMS) for the world's largest commercial database. The Teradata can be store data up to Teradata bytes in size. The teradata is a fully scalable relational database management system. It is widely used to manage large data warehousing operations. Teradata is traded on the New York Stock Exchange (NYSE) under the stock of the symbol. The Teradata product may be called as the "data warehouse system" and stores and manages all the data. And takes addition of some more servers and nodes takes the increasing quantity the database software keeps on top of the amount of data that can be stored. The success of your data warehouse has always rested on the performance of your data sources or data structured engine. But as your data requirements grow increasingly complex, performance comes more vital. So does finding a much faster, simpler way to manage your data warehouse. You can find that combination of undefined performance and efficient management in the Teradata Database. The transmission of the successful Data Warehouse (DW) project in a heterogeneous landscape of various data base sources, limited resources and lack of customer requirements and unfocused and budgeting constraints is always challenging and risky – many long term DW project failures and related to the Requirements defined needs and information requirements for decision making. Extract transform load process that process is a database usage process which is generally or widely used in the data warehouse field.

The Teradata Multimedia Object Manager is a general purpose content analysis multimedia server designed for symmetric multiprocessing and massively parallel processing in environmental process. The Multimedia Object Manager defines and manipulates user-

defined functions with invoking in parallel process to analyze or manipulate the contents of multimedia objects.

This paper analyses the concept of big data in terms of structured and unstructured data, big data tools and effectiveness of big data in small & medium businesses. The main aim of the paper is to understand the theme of big data and the operations on big data, addresses that challenges the data management through big amount data tools, comparing big amount of data or database operations with the Enterprise Content Management instructions or operations, effects of big amount of data in the business environment or business in industrial nature, and the advantages or benefits of using big amount data in organizations or industrial area. The huge amount of piling of data or database is also a concern amongst all industry sectors and in particular the Small and Medium Business (SMB) who need to compete with large competitors in computational area. Therefore, there is most important tools to find the effectiveness of big data among small and medium business. The generated result or output is basically based on various or several different types of research articles and case studies in order to yield a supportive discussion with recommendation and conclusion. Teradata shows two basic architecture which are as follows –

1. SHARED NOTHING ARCHITECTURE

A shared nothing architecture is a distributed computing architecture in which each node is independent and self-sufficient. The data warehouse use a “shared nothing” in

1. Shared Nothing Architecture: - A shared nothing architecture is a distributed computing architecture in which each node is independent and self-sufficient. The data warehouse use a “shared nothing” in which that process or procedure, which means that we can say each server takes its own personnel memory and the processing power.

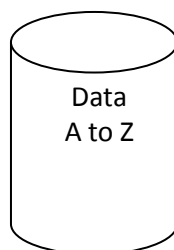


Fig. 2 Shared Everything Architecture

which that process or procedure, which means that we can say each server takes its own personnel memory and the processing power.

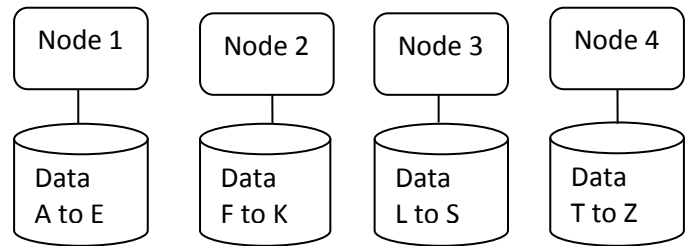
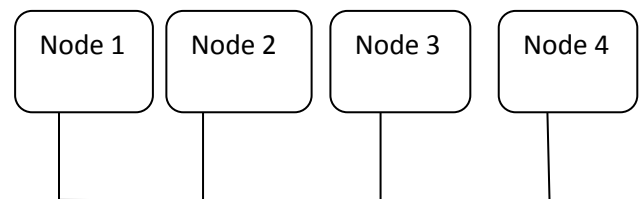


Fig. 1. Shared Nothing Architecture

2. SHARED EVERYTHING ARCHITECTURE

A shared everything is also known as Shared memory or shared disk. Shared-Everything architecture, also referred to as SE architecture or SE model, is a multiprocessor computer system architecture. in which all the processors share same memory address space for read/write access.



II. LOGICAL VIEW OF TERADATA ARCHITECTURE

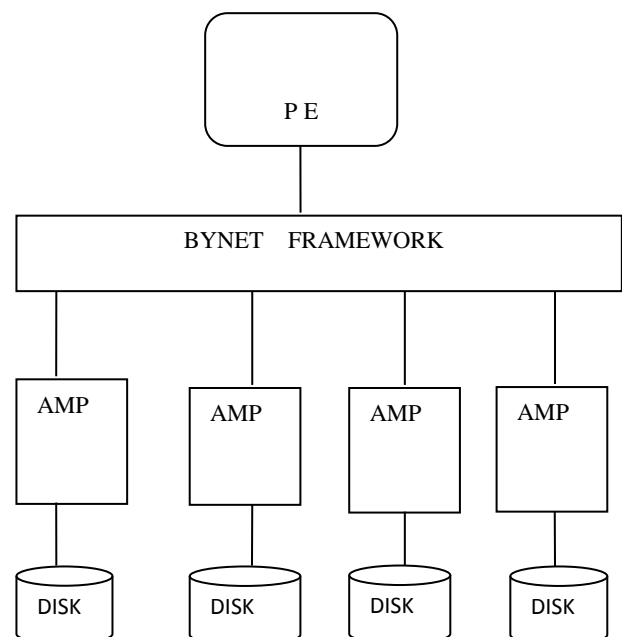


Fig.. 3 Logical View of Teradata Architecture

The Biggest strength of the teradata is the parallelism. so the architecture of the teradata is designed in such way to keep this strength in mind. Teradata is unique from any other database the main component of the teradata architecture is as follows -

- 1) PE (Parsing Engine)
- 2) AMP (Access Module Processor) and
- 3) BYNET.

We also look into these components in details after looking the logical view of the architecture.

1. Parsing Engine: - Whenever a user login to teradata it actually connect to parsing engine (PE). When a user submit a query, then the PE takes action, it creates a plan and instruct AMPs what to do in order to get the result from the query PE checks the syntax of the query, check the user security rights. Then PE Come up with best optimized plan for the execution of the query.

2. Access Module Processor (AMP):- Each AMP attached to the teradata system listens to the PE via the BYNET for instruction. Each AMP is connected to its own disk and has the privilege to read or write the data to its own disk and has the privilege to read or write the data to its disk. The AMP can be best considered as the computer processor with its own disk attached to it. This is known as the "SHARED NOTHING" architecture.

3. BYNET: - The BYNET is the communication channel between PE and AMP is correct and on right track.. In teradata system there are always two BYNET systems. They are called as BYNET 0 and BYNET 1.

What is Big Data?

According to Sagioglu and Sinanc (2013), Big Data can be defined as enormous amount of data with complex structures that excerpt data values by capture & analysis technique. The term big data refers only to data size and the properties of big data such as variety, complexity and volume that takes care of the challenges (Katal et al., 2013). Big data is huge information, it not just handles the information being put away in customary warehouse but also the unused information stored in the distribution centers. Big data helps in accessing load or lots of information which helps in forming better business strategies and techniques and makes investigating information more effective.

III. MANAGING STRUCTURED & UNSTRUCTURED DATA WITH BIG DATA

In data management, data can be categorized in to two types - structured and unstructured data. Near about the data 80 to 85percent of total volume of data content is

unstructured (Daemon, 2013). Generally, structured data applies to relational database that deals with straightforward, clear internet searches and static operations. Unstructured information is basically the inverse which is difficult to manage due to the increase in volume of data. Big data deals with both structured and unstructured content and is more focused on the unstructured approach which targets on transforming unstructured content to structured content. Many organizations have started investigating unstructured data, and numerous unstructured data is scaling out their answers in handling expansive volumes of unstructured information. Likewise new innovations are developing to help backing unstructured data and the investigation of unstructured data. According to Doan et al. (2009,) numerous organizations accept that their unstructured data stores and incorporates data that could help them make better business choices. Some of these acceptances help both structured and unstructured data by incorporating frameworks like Hadoop and Map Reduce.

IV. BIG DATA TOOLS

Big data can be termed as the next generation of Data Base Management System (DBMS). This does not imply that the traditional DBMS are out of the market as Teradata, metadata are being used across various organizations. It is the volume and speed of data transmission which needs to be addressed. The growth in Healthcare & Manufacturing sectors is rapid and the volume of data has increased drastically and this is where big data frameworks can be of great help (Russom, 2011). Hadoop is one of the extensively used open source frameworks that are categorized in to Distributed Computing. Hadoop instruments a computational archetype called Map Reduce in order to process a large amount of data.

Advantages of Big data:

Big data can help corporate identify large amount of hidden data which implies organizations now have entry to more precise data which can impact their business. Better business marketing decisions can be implemented with the help of big data targeting the right user. This helps in reduced cost and better customer satisfaction. Big data analytics can help in multiple sectors by means of managing different portfolios of respective industries resulting in saving billions of dollars and yielding higher profit. Hadoop has the capacity to process data both internally and in cloud without any authentication.

Technology and Products

Teradata is a massively parallel processing system running a shares nothing architecture. Its technology consists of hardware, software, database, and consulting. The system moves data to a data warehouse where it can be recalled and analyzed.

The system can be used as back-up for one another during downtime, and in normal operation balance the work load across themselves.

In 2009 ,forester research issued a report ,” the forester awake : Marketing research company in the marketing area or industrial area Gartner Group placed Teradata in the “leaders quadrant ” in its 2009,2010,and 2012 reports “Magic Quadrant ” For data warehouse , Data warehouse Database Management System.

The Teradata Database Difference

Your data warehouse solution depends on its database engine. So make sure that the database you select offers:

- Effortless or hardwork or man power scalability through fully parallel operation.
- Mission-critical availability.
- Complex and ad-hoc query performance.
- Mixed workload management (ability to keep data current and handle strategic operational queries).
- Columnar.
- Cost-based optimizer.
- Intelligent Memory.
- Virtualized interconnect.
- Virtualized storage.
- Virtualized CPU.
- File system.
- Ease of manageability.
- Simplified mainframe integration.
- Complete support infrastructure.
- Reference accounts.
- Lowest total cost of ownership.

An Array of Powerful Analytics

When it comes to powerful analytics, Teradata Database is unmatched. That’s because it’s optimized or observe to offer arrange of analytical data or database or analytical capabilities, everything from geospatial and temporal to slice-and-dice and data mining. You can extract more value from your data, from more parts of your business, than ever. Plus, you can perform both general-purpose and specialized analyses in a database where all of your supporting data are available to enrich your analytical processes. And Teradata Database includes a variety of built-in features that makes it faster and more efficient to run a wider scope of analytical processes or analytical processing activities, features that let you: Store and process geospatial data. With Teradata Database, your business intelligence (BI) and other applications can use location and proximity along with all the data dimensions you’re used to. show that the Teradata Database can analyze geospatial data inside the data warehouse. That means all your enterprise data are available for showing the result. In addition to common geospatial capabilities, Teradata Database allows or permits you to store the series

of points with timestamp information to understand and analyze movement tracked by GPS or other means. Easily maintain and base analyses on the changes in your business data. The temporal option available with Teradata Database helps you cut through the costly, complicated or difficult to handling, and time-consuming effort that needed to keep pace with changes in your business.

Built-In Business Benefits

Teradata Database delivers benefits that provide a single view of your business or your business process, a view that lets you make intelligent, Faster or confident on your work decisions and realize top or up - and bottom or down -line growth. It will help you: Extend Decision Support Capabilities Teradata Database can enable much more sophisticated or more difficult Business intelligence analysis by optimizing increasingly complex analytics and simplifying their development process. That increased performance or score gets answers or report to users faster and processes more work in the same time frame. We also offer enhanced security and privacy capabilities that meet all consumer. And our enhanced accessibility for external sources, external resources, targets, and processes enables multiple or many of the language data warehouses for global corporations or big organizations.. Extend the Capabilities of Teradata Database Each industry, business , and application or functions are uniquely describe. Many incorporate unique types of data and analytic functions. Within Teradata Database’s high-performance and scalable parallel architecture, you can extend the native capabilities with user-defined types, functions which are designed by users also user-defined SQL operators, stored procedures, and table operators. Embed script program logic with special script table operators for Ruby, Perl, Python, R, or shell scripts to put application logic in the database for architectural flexibility. Store any type of data your business uses or have a query automatically reach out for data stored in Oracle transaction systems and analyze it with all of the scalability and ease of management you expect from Teradata Database. Increase Performance for Operational Workloads Teradata Database lets you acquire data faster than ever for more effective, timely decisions. We’ll show that you have to extend the operational use of your data warehouse to more easily support front-line applications and decision makers.

Need More Reasons To Choose Teradata

The Teradata is the over all world’s leading analytical data or instruction solutions focused on integrated data warehousing, big data analytics, and business applications. Teradata’s innovative products and service deliver integration or collection and insight to empower organizations to achieve the various competitive advantage.

CONCLUSION

The teradata is one of the technique or an application or software which are used to handle or store large or huge amount of data from database. The Teradata from above all sentences prove the sophistication of database capturing method. It is clearly evident that organizations have been influenced greatly by big data. The issue of unstructured content is addressed by big data with the help of open source frameworks and tools. There are more advantages in big data when compared with other Business Intelligence or Data Warehouse tools . The review of literature clearly shows that big data is a developing trend in the current market and more tools and frameworks are being upgraded to yield better results. The Multimedia Object Manager is designed to meet the processing challenges of the emerging generation of multimedia content-based applications. Teradata used for the handling or storing big amount of data.

REFERENCE

1. www.sas.com/en_us/insights/analytics/big-dataanalytics.
2. [www.ibm.com/big-data/us/en/big-data-andanalytics/\(2015,january\)](http://www.ibm.com/big-data/us/en/big-data-andanalytics/(2015,january))
3. "TERADATA CORP /DE/ 2013 Annual Report Form (10-K)" (XBRL). United States Securities and Exchange Commission. February27,2014.
4. www.techopedia.com/definition/...../teradata
5. en.wikipedia.org/wiki/Teradata
6. Sagioglu, S., & Sinanc, D. (2013, May). Big data: A review. In Collaboration Technologies and Systems (CTS), 2013 International Conference on (pp. 42-47). IEEE.
7. Daemon, W. (2013, November 18). Combining Unstructured and Structured Data to Deliver Big Data Business Value. Retrieved from wiki on:
http://wikibon.org/wiki/v/Combining_Unstructured_and_Structured_Data_to_Deliver_Big_Data_Business_Value A.B.2014
8. M. Stonebreaker, "The Case for Shared Nothing," IEEE Data Engineering Bulletin, Vol. 9, No. 2, pp. 4-9.