Data Warehousing:
SAP HANA, Comparison, Tools and Issues

Raymond Delgado, Savita Yadav, Davide Tita, Ninelia Talverdi

CIS 5430 – Databases and Data Warehousing

Dr. Ming Wang

College of Business and Economics

California State University, Los Angeles

ABSTRACT

Data warehousing has become the most critical tool for the decision makers of any industry and it has changed the way global organizations conduct business. Many people believe that developing a business strategy without a data warehouse is impractical. The purpose of this paper is to research and explain the importance of data warehousing and SAP HANA database.
We began by defining data warehousing and describing the business uses for it. Our discussion then focused on the key implementation issues of data warehousing. We examined SAP HANA database and compared it with two other data warehouse products. In addition, we discussed a case study explaining how SAP BW/4HANA was implemented successfully for a leading utilities company and the impacts of that implementation. We also investigated the top five tools available that are used to design, implement and maintain the data warehouse.

Keywords: Data Warehouse (DWH), Data Warehousing, Online Analytical Processing (OLAP), SAP HANA database, SAP BW/4HANA

1.0 Introduction

Historically, organizations were limited on the amount of data that it can retain. As a result, management had to rely on poor information and "trust their gut" when they needed to make a decision. Moreover, performing data analysis on their limited data was a very slow and difficult process. The individual departments also had to retain their own databases to store the data they need to operate. Those individual databases may have met the need for the department but they were incapable with other department's databases. This means that organizations lacked a complete view of their data. Many organizations now implement data warehouses (DW) to solve this issue and find a competitive edge in the market.

Data warehouses are customized data storage that aggregate data from different sources and store it in a central location so that reports and queries can be run on it. This technology appears from the need to integrate enterprise data through various application repositories that an organization may have, so that it would make the data available to anyone who needs to use the information and make decisions. Many companies use data warehouses to compile financial reports and business metric analyses. Due to advancement in cloud computing and storage, data warehousing is becoming more affordable and manageable for many organizations. One other advantage of cloud-based data warehouses is that the organizations do not have to invest into the hardware and technical staff that a traditional on-premise data warehouse would require.

There are many data warehouse solutions and providers that an organization can choose from. One of the most viable options for a data warehouse solution is SAP S/4 HANA. SAP S/4HANA is an enterprise resource planning (ERP) software that includes both data warehousing (SAP BW/4HANA) and advanced data analysis (SAP Predictive Analytics). The fact that SAP S/4 HANA is an ERP software ensures that the organization's data is stored in a uniformed and effective manner. Organizations' members can use SAP Predictive Analytics to query data from their data warehouse and perform advanced analysis and machine learning to gain actionable insights from their data.

2.0 Review of Literature

In his article in journal article in Issues in Information Systems, James F. Kimpel states that management's inability to make effective decisions and the silo effect are driving the need to implement a data warehouse (Kimpel, 2013). It must be noted that lack of quality data and the fact that it is slow and difficult to perform data analysis is hindering the decision making of upper management. This means that management is forced to make decisions on incomplete information or postpone their decision until they have the information needed. However, there is still a real possibility that the opportunity will be unavailable once they have the information to act. The Silo effect also hinders the effectiveness of the whole organization. For instance, there might be major inconsistency among the data stored in a department's data repository. Moreover, the organization data is decentralized—thus obtaining a holistic view of the organization is very difficult. Lastly, the silo effect can greatly reduce the flow of data among the departments. Thus, the departments work in isolation and run the risk of conflicting with each other.

Kimpel states that data warehouse resolves these issues by integrating the whole organization's data into a central data repository. Kimpel also states that the data within the data must be stored in a format that is consistent, reliable, timely, and readily available (Kimpel, 2013). We can define a data warehouse as a massive data repository that stores the whole organization's data in a uniformed

format. It must be noted that utilizing a data warehouse enables online analytical processing (OLAP). This means that management can use OLAP techniques (e.g. slice and dice) to view their data by many different dimensions. Moreover, many data warehouse providers provide simple-to-use middleware for people to perform advanced data analysis in near-real time. Thus, the manager will always have the data needed to make a quick and decisive decision. Data warehousing can also eliminate the silo effect by providing a single and consistent data repository to all department members.

Kimpel also cites some benefits that an organization can achieve when they implement a data warehouse. Here are the benefits that we will discuss further (Kimpel, 2013):

- Improved Productivity
- Better Quality information
- Speedy information retrieval
- Better decision making
- Increased competitive position
- Easy to use

An organization that employs a data warehouse tends to be more productive in many ways. For instance, a firm can use data warehousing and OLAP techniques to automate some routine decisions (e.g. loan amount). The organization's information will be more complete because the data warehouse will contain historic data that is consistent, accurate, and available. The time to access data will be greatly reduced because all of the data is stored in a single location. The management can also be more decisive in their decision making because they will have the information needed. Firms that utilize a data warehouse will be more competitive because they will be able to quickly identify opportunities to exploit. The last benefit that we would like to discuss is the fact that it would be easier to gather the information needed. This is because all the data is in a single location and many data warehouse vendors provide middleware that is easy to use.

3.0 Research methods and questions

The methodology of the project consists of a linear process that is divided into six phases. These phases are:

- Group forming
- Research question selection
- Conduct research
- Record findings
- Process findings into a report and presentation
- Finalize report and presentation

The group forming phases consist of the professor randomly dividing the class into multiple groups. The professor wanted the groups randomly selected to reflect the fact that in a corporate setting people will not have the privilege to select their co-workers or project members. Moreover, this challenge will force the students to organize and divide the work in a fair manner.

The research question selection phase began with us discussing the questions as a group. The purpose of this initial discussion was to gage our previous experience with the subject matter and how hard the questions would be to answer. Then, we selected the question that we would prefer to research. Fortunately, we all wanted to research different questions. Thus, we did not have to resolve any major

disagreements. This phase concluded with us confirming on what questions we are going and who is responsible for answering these questions. The research questions that we have selected are:

- ❑ Compare SAP HANA with two other database and data warehouse products
- ❑ Key data warehouse implementation issues
- ❑ A case study of SAP BW HANA data warehouse implementation and benefits obtained, lessons learned, and pitfalls to avoid.
- ❑ Types of software to support data warehouse design, implementation, and maintenance

The research phase of the project first consisted of initial research. This means that each member would first do their own research about their questions. We all tried to find at least three sources that are related to the research questions. The main sources that we utilized were Google Scholar and the CSULA library database. The main issue that we found was the lack of sources that dealt with both SAP S/4Hana and data warehousing concepts. We were able to resolve this issue by sharing our sources with each other and finding supplementary sources.

After we got a sufficient number of sources, we proceeded to record the findings of our source. We began this process by simply listing all the main points of our sources. We then filter out the main points that are not relevant to our questions. We organized the remaining factoids by its importance to the questions that we are trying to answer. It must be noted that we retained all of our findings in case we might need them for a different section of our paper.

Once we got a working list of potential answers, we transformed our findings into a report and presentation. In terms of our report, we organized our findings into a one page answer to our research questions. Then we review our own answer to make sure that our response is the most informative and concise as possible. We also made sure that all visuals used directly answer our question and are legible. For our presentation, we made sure that we highlighted the findings that are most relevant to the class. At the end of this phase, our report and presentation should only contain all the information that we want to share with the class.
The last phase of the project is to finalize our report and presentation. Please note that during this phase we are focused on identifying errors like spelling and grammar. The way we conducted this phase is to have all members read the report and presentation. The purpose of this meeting was to ensure that all members agreed to how the report and presentation are developed (e.g. font size and colors). The final product of this phase is the complete report and presentation for the professor and class.

4.0 Discussion/Analysis

4.1 Compare SAP HANA with two other database and data warehouse products

Data Warehouse Providers and Solutions

While a database is a key component of a data warehouse, it is not the only component. It is more helpful to think of a data warehouse as a software solution that comprises many components. The following table lists some of the top data warehouse solutions and their providers:

Provider	Solution
Amazon	Redshift
Google	BigQuery
Snowflake	Snowflake
IBM	Db2 Warehouse
SAP	Business Warehouse
Microsoft	Azure SQL Data Warehouse
Oracle	Autonomous Data Warehouse

Table 1. *Data warehouse providers and solutions*

Following is the overview of two Data warehouse solutions:

- **Oracle Autonomous Data Warehouse** runs on Oracle's Exadata cloud platform. It uses the same Oracle Database software as on-premises installations.
- **IBM Db2 Warehouse** is a cloud data storage platform that integrates the Apache Spark analytics engine and IBM's Db2 in-memory database.

Comparison of SAP HANA with ORACLE and IBM DB2 databases:

Features	SAP HANA	Oracle	IBM DB2
Initial Release and Latest Version	2010, 2.0 SPS05	1979, 19c	1987, Db2 11.5.5
Description	SAP HANA (high-performance analytic appliance) is an in-memory, column-oriented, relational database management system developed and marketed by SAP SE.	Oracle Database is a multi-model database management system produced and marketed by Oracle Corporation.	DB2 is a database product from IBM. It is a Relational Database Management System (RDBMS).
Written In	C, C++	Assembly language, C, C++	C, C++, assembly, Java

Scalability	Sap Hana is highly scalable and achieved by using a columnar data organization and massively parallelizing the workload.	Oracle Database In-Memory's unique dual-format enables transparent scale-up and scale-out for analytics and OLTP workloads running together.	IBM DB2 is very scalable, offering a multi-choice of possibilities massively parallel processing and also clustering.
Disaster Recovery	HANA System Replication (HSR) is a high-availability and disaster recovery solution provided by SAP HANA.	Oracle ensures the integrity of the data in case of system failure. It includes a powerful tool called Recovery Manager (RMAN) – allows DBA to perform cold, hot, and incremental database backups and point-in-time recoveries.	No matter wherever the data resides, DB2 ensures high availability for complete or partial site failures and has the characteristic to support up to 3 remote standby servers.
Transaction Concept	SAP HANA is an ACID-compliant database with advanced data processing, application services, and flexible data integration services.	All Oracle databases, Oracle RDB, and InnoDB transactions comply with ACID properties.	IBM DB2 is an Acid compliant database.

Table 2. *Comparison of SAP HANA with ORACLE and IBM Db2*

4.2 Key data warehouse implementation issues

For most SME and large organizations, the data warehouse is the backbone of business intelligence. Businesses face warehouse problems all over the world in their quest to make warehouses more efficient and productive. There are quite a few issues involving Data warehouse implementation. The following discussion includes the key issues.

Why do companies need implementation and data quality?

Data quality in a data warehouse is fundamental to achieve efficient processes. The increasing regulatory environment and the increasing complexity of data warehouse solutions is becoming a major topic. Data quality doesn't have to be "error-free" but it depends on the user requirements. Nowadays companies have different policies on data quality and privacy called "data government policies". These are policies stipulated to maintain the right data quality and to make sure data privacy rules are not violated. Furthermore, regulatory environments such as TCPA (US Telephone Consumer Protection Act) and the GDPR in Europe, put pressure on data quality.

Data Implementation

The warehouse implementation is usually a huge effort which must be designed and carried out according to well-developed methods. Few specific issues which arise with data warehousing are Construction, Administration and Quality control.

The project must be in line with the company's overall strategy and goals. It's essential to fulfil the user's requirements and expectations about the completed project. Data Warehouse building has to be incremental. Building adaptively is essential to adapt to evolving business requirements. Data warehouse implementation projects should be managed by both IT professionals and business professionals. This way the implementation will have a better outcome and all the requirements will be fulfilled. Manual data entry or processing can put the accuracy of the data at risk. Data users are required to possess a high level of Data quality and knowledge. Failing to comply might result in a misinterpretation of data. Prior to the development of the data warehouse, usage forecasts should be conservatively calculated and updated on a regular basis to match the current requirements. A challenge during data implementation is integrating the available source data into the warehouse's data model. Because of the constant rapid shift in technology, the warehouse's requirements and demands can change over time. Some other major tasks in Data Warehouse Implementation are, managing the warehouse in a large company, designing a management function and selecting a data warehouse management team.

Data Quality

As warehouse data processes change frequently there are few important points to maintain data quality and improve processes. First the Data Owner should keep the database up to date to satisfy a good quality. Moreover, enhancements and implementations to new rules for errors that were not discovered yet is vital. Lastly, disabling data quality rules that are no longer needed is important to avoid overlapping rules. If a company does not yet have a strategy for data quality, it is imperative to build a prototype into a business unit to show the benefits. Furthermore, maintaining good data quality constant over time after implementation, is important to ensure a long-run advantage.

With the ever-changing market climate and data warehousing to meet your data needs, you need the ability to embrace and incorporate analytics to streamline the business intelligence process. Business managers may make well-informed business decisions to help their organization's strategic priorities if they have access to advanced analytics insights. This allows them to deceive and exaggerate.

4.3 SAP BW/4HANA in Action: Leading Utilities Company Case Study

The following customer case study, published by Eckerson Group LLC, illustrates how SAP BW/4HANA was implemented for a leading utilities company and the impacts of that implementation. (Eckerson Group)

Based in Melbourne, this leading utilities company is a $10 billion distributor of gas and electricity. It delivers gas to more than 1.3 million homes and businesses, and electricity to more than 350,000 homes and businesses.

Why the Company Chose SAP

The problems the company faced with its existing data management system were a study in irony, as users pulled data out of a “data management system” that would not support their needs and created their own new data sets in Excel. The company’s legacy Oracle data warehouse provided no single source of truth. There were issues with data transformation and the system was not mature enough in terms of BI and analytics. As a result, teams struggled to get the information they wanted, and what information they did get they didn't trust. This lack of trust created yet another issue: business users requested raw data extracts in order to manipulate the data themselves in Excel, resulting in many different versions of the same data. There was no self-service reporting. Compliance reporting was developed and delivered using Excel; completing the monthly compliance KPI dashboard took a team of four up to a month to source and compile data from multiple platforms.

The Solution

The company chose SAP to implement a database consolidation strategy that would create a single source of truth and meet demands from business users for self-service reporting, thereby optimizing customer experience, cost, and operations. The SAP BW/4HANA system ingests 45 to 50 billion data records every day from various SAP and non-SAP systems and devices, including meters and other customer facing touchpoints. The company did not consider any other vendor for its new solution because so many of its existing systems were from SAP. Replacing Oracle with SAP BW/4HANA provided the missing piece. The SAP BW/4HANA system supports the company’s employee base of approximately 1,700 people, 1,200 of whom use the system for basic housekeeping such as timesheets and 500 of whom are heavy system users, performing a variety of tasks such as work orders and billing. Different teams such as customer service, finance and asset management created and used the dashboards to measure KPIs and various reporting purposes.

Impact and Benefits

On a scale of 1 (least satisfied) to 10 (most satisfied), the company rates its satisfaction with SAP BW/4HANA as a solid 7–8 out of 10, with beneficial results as outlines below.

From a technical point of view, the SAP BW/4HANA implementation has resulted in significant efficiency and flexibility improvements for the company. The energy company ingests two to three years’ worth of data from multiple SAP and non-SAP systems into its dashboard, yet the new system manages all of that data, so people can obtain the analytic insights they need to do their jobs in a minute or so compared to the hours, days, or weeks it took using the previous system. Real-time data duplication in HANA eliminates the need for complicated and inefficient extractors, and the organization now has a variety of data mash-up and reporting solutions. Users can process year-to-date and month-to-date volumes from multiple SAP and non-SAP sources in seconds. These data volumes are significant and previously could not be supported in real time. Users can also analyze annual and monthly usage and performance trends.

From a business perspective, the company has not quantified ROI, TCO, or other metrics, and it believed the assessment of ROI is at least a year away. However, one thing was certain: deploying SAP BW/4HANA was the cornerstone of its path to become an insights-driven company, as it assesses consumer demands and its own capabilities to market and sell new products and services through its energy distribution network. One easily quantifiable result was the ability to track and ensure compliance. Once the system detects issues that may lead to breaches in compliance, the company can now proactively respond to specific steps within the process. The old system served data that was up to a month old, which made it impossible to detect current issues and respond accordingly. That resulted in regulatory violations and fees, which the firm believes it can no longer incur with the new system in place.

Lessons Learned, Drawbacks and Improvements

The company started its SAP BW/4HANA deployment in late 2016 and went live December 2017, and as with virtually every technology implementation, the process was not without its challenges. In the early stages, product stability was an issue, including when deploying patch upgrades and bug fixes.

One of the major areas of concern was moving data between hot and cold storage. The company would like to see a smoother transition between the two. They have mentioned this to SAP, and the situation has been improving. Lower-region refreshes, such as copying data from production back to quality assurance, are not as smooth as the legacy BW process and involve more moving parts (e.g., IQ, SLT, HANA schema remapping, and SAPI/ODP source systems). One of the data flows relies on the BW app server for exception aggregation. The company also experienced issues with differences between SAP BW/4HANA and SAP BW. While data tiering optimization (DTO) and data temperature control have not been as smooth as the business would want, the latest SP8 feature list appears to be encouraging.

On the people side of the deployment, the company found that experienced BW/4HANA resources were not always readily available, but because they are part of SAP's early adopter program, the project team was able to escalate issues.

Conclusion and Future

The company chose SAP to implement its database consolidation strategy, to provide a single source of truth that would bring a user-friendly, feature-rich environment for its people. To a great extent it appears that SAP has delivered. Users can now obtain the analytic insights they need to do their jobs in minutes where getting the same results used to take hours, days, or weeks.

Things are less clear on the implementation and usage sides. The company continues to experience a few technical issues, including its backward compatibility between BW/4HANA and plain-vanilla BW. SAP messaging around HANA is that all data is in memory, which avoids data movement and is a major ingredient in its high performance. The company, however, uses hot and cold storage, which is an optional feature that requires moving data in and out of SAP HANA. Nevertheless, the companies are working closely to address issues, and that process will continue.

4.4 Types of software to support data warehouse design, implementation, and maintenance

Data warehouse solutions enable users to gain critical insights into their data through improved seamless self-service business intelligence (BI) capabilities. Though the purpose of the software remains the same, it differs in the mode of design and implementations. The data warehouse solution can be deployed on-premises or on the cloud.

With cloud data warehouses, businesses can scale horizontally to hold increased storage and compute requirements. A data warehouse deployed on the cloud provides an improved infrastructure that lets companies focus more on delivering better and faster insights rather than managing a full house of servers on premises. These solutions provide cost control as organizations pay for what they use. On the other hand, an on-premises data warehouse software lets organizations buy one time, deploy in-house and enable control over their hardware and software infrastructure. This implementation solution requires a consultant to help with installation and ongoing support. One benefit of on-premises data warehouse solutions is that they provide full control and access to data within an organization, helping minimize security risks. (G2, 2021)

Nowadays, the data warehouses are moving from the data center to cloud-based data warehouses. Data warehousing tools are important for managing and maintaining these cloud-based Data warehouses. There are a few factors that should be considered when choosing a data warehouse solution. Figure 1 shows the criteria that can be considered to make the decision easier.

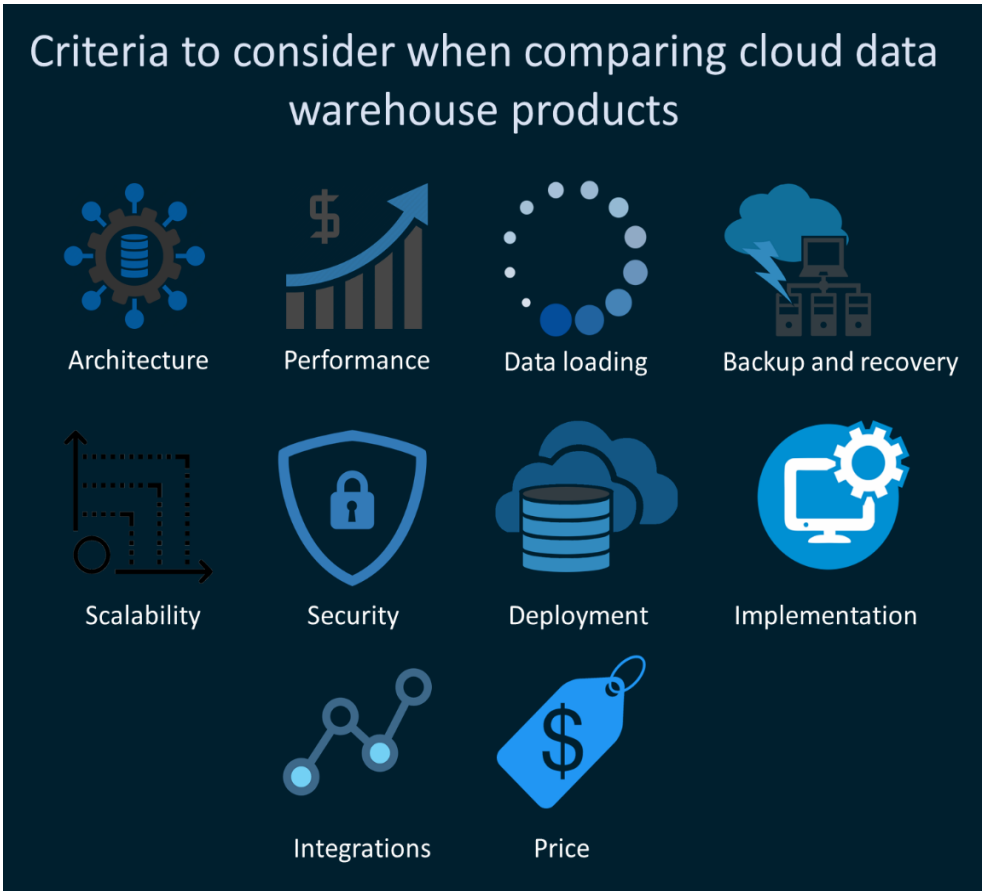

Figure 1. *Criteria to consider when choosing cloud data warehouse products (AltexSoft, 2021)*

Cloud data warehousing market solutions

In this section, the top 5 cloud data warehousing solutions and their characteristics of them are described.

Amazon Redshift

Gartner reports Amazon Web Services (AWS) is often considered the leading cloud data warehouse platform-as-a-service provider (Gartner, 2016). As defined by Amazon, Amazon Redshift is a fast, fully managed, petabyte-scale data warehouse that makes it simple and cost-effective to analyze all our data using existing business intelligence tools. (Mansell, 1988)

Amazon Redshift engine is a SQL-compliant, MPP, query processing and database management system designed to support analytics workload. The storage and compute are distributed across one or more compute nodes (Gupta et al., 2015). The core infrastructure of Amazon Redshift data warehouse is a cluster and it is composed by a leader node and one or more compute nodes

Amazon Redshift supports integration with other AWS services and built-in commands that load data and information in parallel to each node from AWS DynamoDB, S3 or EC2. In these services, we can add AWS Kinesis, Elastic MapReduce, Data Pipeline, and Lambda.

Microsoft Azure Synapse Analytics (formerly Azure SQL Data Warehouse)

“Microsoft Azure Synapse is a cloud based, scale-out database capable of processing massive volumes of data, both relational and nonrelational. At the heart of Azure Synapse is a cloud-native, distributed SQL processing engine. It’s built on the foundation of SQL Server to drive your most demanding enterprise data warehousing workloads. Similar to other cloud MPP solutions, Azure SQL Data Warehouse (SQL DW) separates storage and compute, billing for each separately. Azure Synapse saves relational tables data with columnar storage and abstracts physical machines by representing compute power in the form of data warehouse units (DWUs). This allows users to easily and seamlessly scale compute resources at will.” (Qlik, 2021)

Google BigQuery

“BigQuery is a fully managed, serverless data warehouse that automatically scales to match storage and computing power needs. With BigQuery, you get a columnar and ANSI SQL database that can analyze terabytes to petabytes of data at incredible speeds. BigQuery also lets you do geospatial data analysis using familiar SQL with BigQuery GIS. In addition, you can quickly build and operationalize ML models on large-scale structured or semi-structured data using simple SQL with BigQuery ML. And you can support real-time interactive dashboarding with BigQuery BI Engine. The BigQuery architecture is composed of several components. Borg is the computer. Colossus is the distributed storage. Jupiter is the network. And Dremel is the execution engine.” (Qlik, 2021)

Snowflake Cloud Data Platform

Snowflake is a fully managed MPP cloud data warehouse that runs on AWS, GCP, and Azure. It is an all-inclusive cloud data warehouse for structured and semi-structured data provided as Software-as-a-Service (SaaS). When you’re a Snowflake user, you can spin up as many virtual warehouses as you need to parallelize and isolate the performance of individual queries. Snowflake enables very high

concurrency by separating storage and computer to ensure that many warehouses can simultaneously access the same data source. You interact with Snowflake's data warehouse through a web browser, the command line, an analytics platform, or via Snowflake's ODBC, JDBC, or other supported drivers. The platform supports ACID-compliant relational processing and has native support for document store formats such as JSON, Avro, ORC (Optimized Row Columnar), Parquet, and XML. (Qlik, 2021)

CONCLUSION

During this paper we have identified the need for data warehouse--which are the inability for management to make decisions and the silo effect. Data Warehouse is a large centralized data repository that stores the whole organization's data in a unified manner. It must be noted that many data warehouses also provide OLAP functionality. SAP S/4 HANA provides data interrogation and advanced data analysis. Some of the major benefits of data warehousing are: improved productivity, better quality information, speed information retrieval, better decision making, increased competitive position, and the ease of use. Our research paper was divided into six major phases: group forming, research question selection, conducting research, record findings, process findings into a report and presentation, finalize report and presentation. Some of the major data warehouse solutions are Amazon's Redshift and IBM's Db2 warehouse. Amazon's Redshift is integrated within the AWS environment. IBM's Db2 warehouse integrates Apache Spark to perform analytics. The key data warehouse implementations issues are the cost and effort to implement a data warehouse and how to retain high data quality. Many companies who moved to SAP BW/4HANA are satisfied with the product since it brings a new opportunity for agile enterprise data warehousing (EDW) modeling and fast delivery of business analytics' requirements for organizations seeking solutions in reporting, planning, dashboarding, and predictive and prescriptive analytics. Most of the companies who used SAP BW/4 HANA are satisfied with the product. These companies cited improved customer relations as the main benefit of using SAP BW/4 HANA. However, they would prefer that SAP provides a simpler way to transition from a legacy system to SAP BW/4 HANA. Many data warehouse solutions are now cloud--making it more obtainable for many organizations.

REFERENCES

A. Gupta, D. Agarwal, D. Tan, J. Kulesza, R. Pathak, S. Stefani, and V. Srinivasan, “Amazon redshift and the case for simpler data warehouses,” in Proceedings of the 2015 ACM SIGMOD International Conference on Management of Data. ACM, 2015, pp. 1917–1923.

AltexSoft. “Snowflake, Redshift, BigQuery, and Others: Cloud Data Warehouse Tools Compared.” AltexSoft, AltexSoft, 14 Jan. 2021, www.altexsoft.com/blog/snowflake-redshift-bigquery-data-warehouse-tools/.

Data Flair. (2019, May 7). SAP HANA Features – 7 Unique Reasons to Learn SAP HANA. Retrieved from Data Flair: https://data-flair.training/blogs/sap-hana-features/

Ferreira, Pedro Joel, et al. “Data Warehousing in the Cloud: Amazon Redshift vs Microsoft Azure SQL.” SCITEPRESS, 2017.

Gartner, 2016. Magic Quadrant for Data Warehouse and Database Management Solutions for Analytics [WWW Document]. Magic Quadr. Data Wareh.

G2. “Top Free Data Warehouse Software.” Best 18 Free Data Warehouse Software Picks in 2021, G2, 2021, www.g2.com/categories/data-warehouse/free.

Hauskrecht, A. (2021, May 20). How to Implement a Data Quality Process. Retrieved from Totpal: https://www.toptal.com/database/data-warehouse-data-quality-process

IT Central Station. (2021, May 20). Compare IBM Db2 Database vs. SAP HANA. Retrieved from IT Central Station: https://www.itcentralstation.com/products/comparisons/ibm-db2-database_vs_sap-hana

Kimpel, J. (2013). Critical Success Factors For Data Warehousing: A Classic Answer To A Modern Question. *Issues in Information Systems*, 376-284.

Mallach, E. (2000). Decision Support And Data Warehouse Systems. New York City: Mc Graw Hill.

Mansell, Chris. “Redshift.” *Amazon*, Five Islands Press, 1988, aws.amazon.com/redshift/faqs/.

Qlik. “Top Cloud Data Warehouses for the Enterprise: Amazon vs. Azure vs. Google vs. Snowflake.” QlikTech International AB, 6 Jan. 2021.

Rudra, A., & Yeo, E. (1999). Key issues in achieving data quality and consistency in data warehousing among large organisations in Australia. Annual Hawaii International Conference on System Sciences (HICSS) (pp. 5-8). Maui: IEEE.

Santos, J. (2020, December 11). Best Data Warehouse Software & Solutions. Retrieved from Project Management.com: https://project-management.com/best-data-warehouse-software/

Wood , Brian, and Timm Grosser. “Customer Insights into Modern Data Warehousing with SAP BW/4HANA.” Eckerson Group, 2019.