International Journal of Leading Research Publication (IJLRP)

E-ISSN: 2582-8010 ● Website: www.ijlrp.com ● Email: editor@ijlrp.com

Data Integration and Orchestration across SAP and External Data Sources

Naresh Kumar Rapolu

Nareshkumar.rapolu@gmail.com

Abstract

The research paper has provided a holistic overview of the important concepts of data integration and orchestration across various SAP and external data sources. It has provided how organisations can benefit from the implementation of these two data processing techniques. The research paper has also mentioned a few disadvantages that can happen during their execution. Finally, a few recommendations have been discussed in the study that can help to effectively solve the drawbacks of each data processing technique across different SAP and external data sources.

***Keywords*- Data integration, Data orchestration, SAP, Data silos**

I. INTRODUCTION

In the contemporary business scenario, data integration and organisation are significant techniques that help companies streamline their stored information. The research paper will be instrumental in exploring the various ways in which these two data processing techniques are used across different SAP and external data sources. The analysis of their merits and demerits will help to gain a comprehensive understanding of these concepts in multiple business scenarios.

II. DISCUSSING THE IDEA OF DATA INTEGRATION ACROSS SAP AND EXTERNAL DATA SOURCES

Data integration can be described as the process of collecting and combining data from different sources. This is mainly done in order to create a concise, single and accurate data set. In the context of SAP, appropriate data integration techniques can help to considerably minimise the number of errors and optimise the different business operations. This idea is mostly pertinent for various business entities. They effectively implement the idea of data integration across the different systems, applications and products that they use for gathering and processing large volumes of data every day[1]. Proper data integration can also be beneficial for ensuring the reliability of the data that is obtained from different external sources. In this way, both individuals and large organisations can maintain a seamless flow of data. Therefore, data integration across SAP and other exterminated resources can help to make better decisions. It can be stated that the data transformation efficiency is measured by dividing the total amount of data that is transformed by the total input data. The resultant is then multiplied by 100 to obtain the desired result.

III. EXPLORING THE CONCEPT OF DATA ORCHESTRATION ACROSS SAP AND EXTERNAL DATA SOURCES

The idea of data orchestration is inherently interrelated to the concept of data integration. In fact, it goes much further and incorporates ideas like data preparation, data processing and data discovery as well. Within the domain of data processing, data orchestration plays a very important role since it allows an organisation to move siloed data from different storage locations into one centralised repository[2]. In this central location, the collected data is properly cleaned and made ready for different business purposes. It needs to be mentioned that the process of data orchestration is automated in nature. Therefore, organisations are able to carry out their business operations with accurate and up-to-date information. There are different ways in which the data construction process is implemented. The three overarching steps can be further elucidated. The first step revolves around the organisation of data from different sources. If the incoming data is obtained from sources like CRM, social media platforms and other channels, it is told within different tools and systems like data lakes, warehouses and other cloud-based tools[3]. Therefore, in this initial step, the data from these multiple sources are collected and properly formatted before transmitting it to the target destination. The next step involves transforming the different types of data into a standard format. This can be helpful in opening the data in a systematic way. In the final step, the collected and transformed data is made available for activation. Data orchestration is therefore incredibly important for compiling the various kinds of information in different SAP and external storage sources. If this technique is appropriately implemented, the business can remove operational bottlenecks and comply with the established data privacy laws. In this regard, it can be mentioned that data latency is calculated by adding the time taken for data extraction, transformation and loading. In addition, the amount of data throughput is measured by dividing the total volume of data that is processed by the amount of time taken for processing.

IV. BENEFITS OF DATA ORCHESTRATION AND INTEGRATION

Data orchestration

Reduced cost: The automated process can help to free a lot of organisational resources. Additionally, since the potential for error is considerably reduced, the company can save a lot of finances that have to be used to solve errors.

Eliminate data bottlenecks: Since the entire process is automated, an organisation can improve their operational efficacy a lot. Therefore, seamless data can flow across an organisation and can be used when required[4].

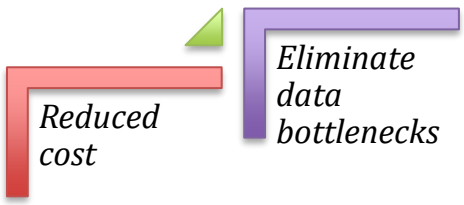

Figure 1: Benefits of data orchestration

Data integration

Improved data quality: Any kind of inconsistency and redundancy can be solved through data integration. Therefore, it can help to produce better and more accurate results. They can be useful for making better decisions.

Better collaboration: An appropriate integration across multiple SAP and external data sources can be instrumental in fostering a collaborative workspace. It can be really important for a business' growth.

V. DRAWBACKS OF DATA ORCHESTRATION AND INTEGRATION

Data orchestration

Lack of compatibility: There can be compatibility problems that can emerge if the data orchestration tools do not align with the differentdata repositories in an organisation.

Overcoming data silos: A business can face difficulties in managing the different silos in which data is stored. Effectively managing these storage units can require extensive time and effort for the business[5].

Data integration

Data security: since data is collected from internal as well as external sources, it can put a business in a vulnerable position. Therefore, strong security measures need to be taken to safeguard important information[6].

Complicated process: If the IT team within an organisation does not properly plan the integration process, it can be quite difficult to maintain. This can potentially hamper the performance of the organisation.

VI. RECOMMENDATIONS

Data orchestration

Use of interoperability tools: Businesses can use different kinds of middlewares and integration platforms like Apache Airflow and Talend to solve the compatibility issue.

Implementing a centralised repository: A central repository such as Snowflake can be implemented to store important data. It can solve the problem of various data silos.

Data integration

Data encryption: All the data that is transmitted from SAP systems and other external data sources needs to have proper encryption for better security measures[7].

Use of low-code or no-code platforms: Businesses can minimise manual coding in data integration by shifting to platforms like Microsoft Power Automate, Zapier etc.

Figure 2: Potential solutions for solving the problems of data integration

VII. CONCLUSION

From the above discussion, it can be stated that the processes of data integration and data orchestration are very important for a business. They help to streamline the data that is stored within various internal and external storage units. In this manner, the businesses are able to operate properly and fulfill their goals.

Abbreviations and acronyms

- SAP - Systems, Applications & Products in Data Processing
- CRM - Customer Relationship Management

Units

- Data transfer and network bandwidth - Kilobits per second (Kbps), Megabits per second (Mbps) and Gigabits per second (Gbps)
- Data quality - Percentage (%)
- Data volume - Bytes, kilobytes, megabytes, terabytes and petabytes.

Equations

- Tlatency = Textraction +Ttransformation + Tloading
- Rthroughput = Vdata/Tprocessing
- Etransform = (Dtransformed/Dinput)×100

REFERENCES

[1] A. Oussous, F.-Z. Benjelloun, A. Ait Lahcen, and S. Belfkih, "Big Data technologies: A survey," *Journal of King Saud University - Computer and Information Sciences*, vol. 30, no. 4, pp. 431–448, Oct. 2018, doi: https://doi.org/10.1016/j.jksuci.2017.06.001. Available: https://www.sciencedirect.com/science/article/pii/S1319157817300034

[2] C. L. Philip Chen and C.-Y. Zhang, "Data-intensive applications, challenges, techniques and technologies: A survey on big data," *Information Sciences*, vol. 275, pp. 314–347, Aug. 2014, doi: https://doi.org/10.1016/j.ins.2014.01.015

[3] C. Vitolo, Y. Elkhatib, D. Reusser, C. J. A. Macleod, and W. Buytaert, "Web technologies for environmental Big Data," *Environmental Modelling & Software*, vol. 63, pp. 185–198, Jan. 2015, doi: https://doi.org/10.1016/j.envsoft.2014.10.007. Available: https://www.sciencedirect.com/science/article/pii/S1364815214002965

[4] H. HU, Y. WEN, T.-S. CHUA, and X. LI, "Toward Scalable Systems for Big Data Analytics: A Technology Tutorial," *Ieee.org*, vol. 2, Jun. 2014, doi: https://doi.org/10.1109/ACCESS.2014.2332453. Available: https://ieeexplore.ieee.org/iel7/6287639/6514899/06842585.pdf

[5] H. V. Jagadish *et al.*, "Big data and its technical challenges," *Communications of the ACM*, vol. 57, no. 7, pp. 86–94, Jul. 2014, doi: https://doi.org/10.1145/2611567

[6] S. Shekhar, "Integrating Data from Geographically Diverse Non-SAP Systems into SAP HANA: Implementation of Master Data Management, Reporting, and Forecasting Model," *Emerging Trends in Machine Intelligence and Big Data*, vol. 10, no. 3, Apr. 2018, Available: https://www.researchgate.net/profile/Suman-Shekhar-9/publication/387794044_Emerging_Trends_in_Machine_Intelligence_and_Big_Data_Integrating_Data_from_Geographically_Di

verse_Non-SAP_Systems_into_SAP_HANA_Implementation_of_Master_Data_Management_Reporting_and_Forecasting_Model/links/677d7cf9fb021f2a47e02ad7/Emerging-Trends-in-Machine-Intelligence-and-Big-Data-Integrating-Data-from-Geographically-Diverse-Non-SAP-Systems-into-SAP-HANA-Implementation-of-Master-Data-Management-Reporting-and-Forecasting-Mod.pdf

[7] X.-W. CHEN and X. LIN, "Big Data Deep Learning: Challenges and Perspectives," *IEEE access*, vol. 2, May 2014, doi: https://doi.org/10.1109/ACCESS.2014.2325029. Available: https://ieeexplore.ieee.org/iel7/6287639/6514899/06817512.pdf