
**MANAGING DATA INTEGRITY AND CONSISTENCY IN E-COMMERCE
BACKENDS****Surender Singh Rawat**

Research Scholar, Department of Computer Science & Applications, Baba Mastnath University

Email: surilinux@gmail.com**Prof. Reena**

Professor, Department of Computer Science & Applications, Baba Mastnath University

Email: reenarajrkt@gmail.com**Abstract:**

In the world of e-commerce, maintaining data integrity and consistency is essential for providing customers with accurate and reliable information. However, managing data integrity and character in e-commerce backends can be challenging due to data duplication, inconsistency, loss, and unauthorized access. This research paper aims to examine the current challenges in managing data integrity and consistency in e-commerce backends and to identify the most effective techniques for overcoming these challenges. Through this analysis, the paper aims to provide e-commerce companies with a better understanding of how to manage data integrity and consistency in their backends, and it offers practical insights into the best practices and techniques that can be used to achieve this goal.

Keywords: *E-commerce, Data integrity, Consistency, Backend, Data validation, Access control.*

1. Introduction:

The internet is a highly dynamic environment that is continuously changing (Zhu et al., 2020). Supply chains, e-commerce, and virtual enterprises have replaced traditional manufacturing and service environments to create more physically distributed enterprise environments (Anthony & Abbas, 2021). As a result, the internet has emerged as a critical resource for modern business, and many companies are developing online identities. A "highly pervasive innovation that is driving significant changes in the conventional ways of doing business" is e-commerce (Amin et al., 2016). Customers are responding to e-commerce with remarkable consistency. Examples include Flipkart, Amazon, OLX, Myntra, and other significant e-commerce websites. Various activities, including online transactions, inventory management, databases containing customer/user information and product information, etc., are carried out through e-commerce websites. However, the issue of e-commerce database management at the backend or the websites being hacked by hackers is also significantly rising. For instance, more than two lakh systems were compromised in nearly 150 countries using ransomware attacks (Asan, 2022). This led to the leakage of crucial data and information, which also caused a significant loss. In this context, it would be interesting

to learn about the research and development work that has been done on this database management (Shukla et al., 2022)

1.1 E-Commerce- Background

E-commerce is a business model that enables small and large business organizations to complete transactions more quickly over the internet. Based on the parties involved in the transaction, there are four main types of electronic commerce: business-to-business, business-to-consumer, consumer-to-consumer, and consumer-to-business. The enormous benefits that e-commerce provides to businesses on both sides of the transaction can be used to explain its rise. The backend of an e-commerce system is the server-side technology that supports the online store's front-end operations, as shown in Fig. 1. It includes critical components such as inventory management, payment processing, customer management, order management, reporting and analytics, and integration with other systems (Yu & Duan, 2021). The backend is crucial for the success of an e-commerce business, as it provides the infrastructure and processes needed to manage and grow the company and ensure data security and reliability.

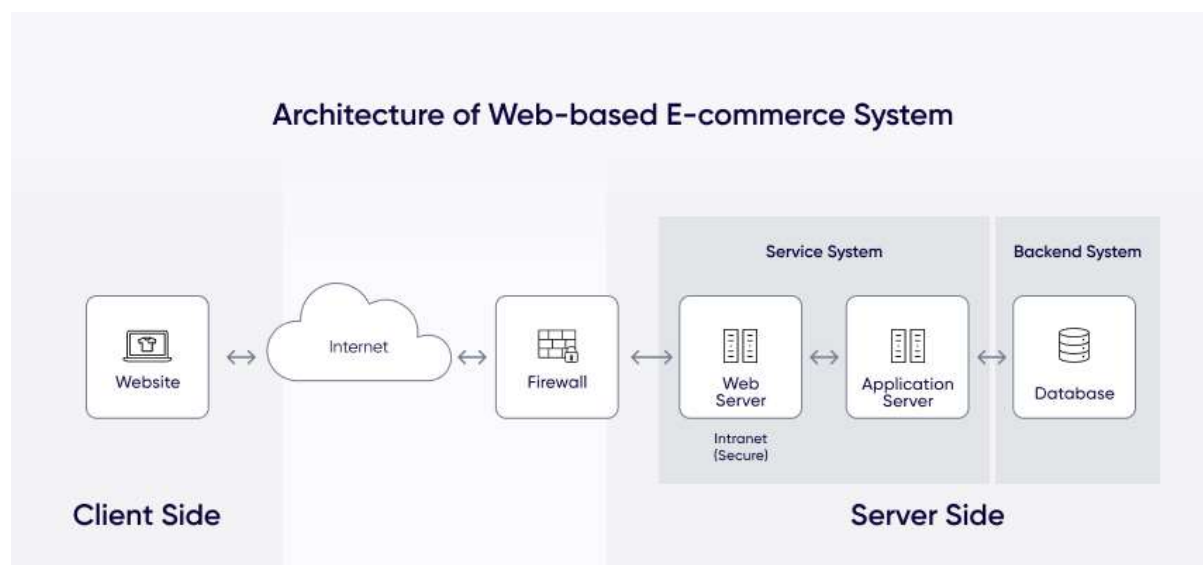


Fig. 1 E-commerce Architecture

1.2 Literature survey

The backend of an e-commerce system can face various problems, including scalability, security, integration, performance, reliability, and data management (Sultana, 2019). Addressing these issues is critical for the success of an e-commerce business, as it ensures that the backend systems are reliable, secure, and capable of supporting the company's growing demands. Therefore, as per Singh (2014), the successful operation of e-commerce depends on the security of transactions and information. However, Malaysian users of e-commerce reportedly thought that security was the

main obstacle to the practice. The author claimed it is simple to understand why people are reluctant to engage in online commerce. He thought that for E-Commerce users to feel comfortable using the internet regularly, they needed to know that their communications and data were secure from unauthorized access or modification. According to Vashi et al. (2017), security problems like hacker intrusions could make using the internet for business purposes challenging. In the same study, Choi & Mai (2018) concluded that security would continue to be the main issue for e-commerce in the future. Therefore, suppliers must invest more in creating secure sites to alleviate these worries. Furthermore, the users must take precautions by using strong firewalls to protect themselves from the concern that hackers will invade their privacy. In addition, users can use some encryption measures to allay their fears about sending sensitive information over the internet.

2. Material & Methods:

The material used to conclude this study is surveyed from the IEEE papers and similar journals.

2.1 Challenges faced by E-commerce Industry for preserving the data integrity

Here are a few more challenges the eCommerce industry faces while preserving data integrity.

1. Multiple Data Sources

The company intranet, suppliers, and vendors feed an eCommerce product data repository. By 2019, three-fourths of analytics solutions used ten or more external data sources, according to Gartner (Lo et al., 2019). Thus, businesses must verify data consistency and integrate the pieces into a central hub.

2. Interacting Applications

E-commerce stores enrich product data with APIs from many sources. They have several tools to analyze product data. Unfortunately, business and IT teams often clash. They don't communicate or understand each other, so they buy overpowered software. Multitasking wastes time and money.

3. Manual data pulls

Data retrieval takes time and knowledge. It threatens business performance. In the 2012 London Whale incident, JP Morgan Chase Bank lost \$2 billion due to Excel data copy-and-paste errors (David, 2017). Manual data functions may cause more problems than they solve unless your data flow is fully automated.

4. Uneven Reports

Inconsistent data cause erratic reporting. Authors, audiences, release schedules, and disjointed information can conflict in manually built reports. In addition, multi-departmental reporting standards can hinder cross-functional teamwork.

5. *Spreadsheet Dependencies*

Past productivity software was Microsoft Excel. Unfortunately, it still works for complex process management tasks. Excel cannot handle billion-dollar data sets or complex data analyses. It hinders cross-team collaboration.

6. *Poor Procedures*

Product data can easily be corrupted as it passes from manufacturers to suppliers to brands and retailers to third-party sellers. Information businesses using inaccurate product data compound that damage. As data changes hands, inaccuracies increase, making product content managers and consumers explode.

2.2 E-commerce Data Management Model

The division of e-commerce data according to the user's perspective is the logical model of e-commerce data. The data types of the entire e-commerce can be effectively divided based on functions and expressions using this model. It is divided into three levels from the standpoint of data warehousing: extract, transforms, and load (ETL). The information needed for analysis can be found in several places. ETL's first step is to extract data from these various sources. As the data extracted will affect the analysis outcome, it is crucial to remove the data to prevent data loss or corruption correctly. The next step is to transform the dataset once the necessary ones have been acquired.

Additionally, data transformation involves cleaning and transforming the data. Handling missing data, eliminating duplicates, and standardizing values are all aspects of cleaning. Next, the acquired data is converted into the desired format using transform functions, rules, or lookup tables. As part of the conversion process, combining data from various sources, creating aggregates, creating surrogate keys, categorizing the data, deriving creative intended values, and using sophisticated validation rules are necessary. The final step is to load the data after the required data has been extracted and transformed.

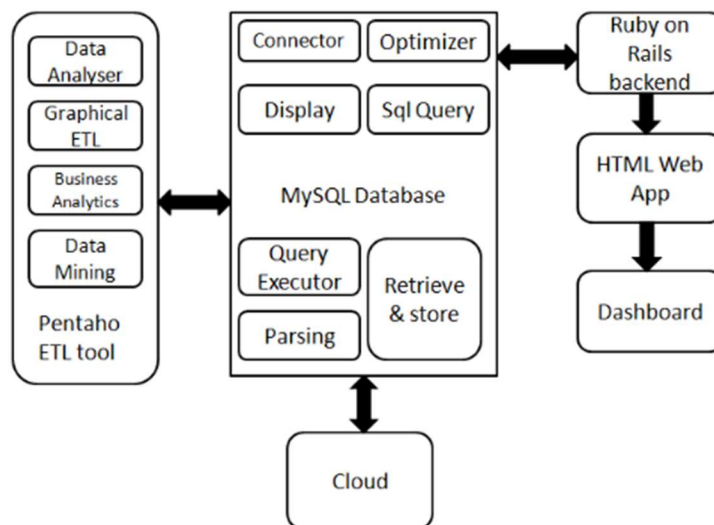


Fig. 2 Architecture of the Model

The market is flooded with ETL tools that are ready to use. By offering connectors to familiar information sources like databases, flat files, mainframe systems, xml, etc., these ETL tools are optimized for the ETL process, which is their main advantage. They offer a way to execute data transformation across various data sources without a doubt or a mishap. Pentaho is one such instrument. Pentaho Business Intelligence Server Edition and necessary tools like JDK and JRE are installed on the Linux server. Once the installation is complete, a local web browser can launch and access the Pentaho server. Port 3306 can connect to the MySQL database using the Pentaho user interface. Once this connection has been made, necessary data can be extracted using joins and SQL queries, then stored in the appropriate format.

3. Result & Discussion:

Data management is a critical aspect of e-commerce platform design and management. However, due to several factors, managing data integrity and consistency in e-commerce backends is a significant challenge. Data duplication is one of the biggest challenges in managing data integrity and consistency in e-commerce backends. With multiple sources of data input and numerous systems used in an e-commerce platform, it is possible for data to be duplicated, leading to inconsistent or inaccurate information. For example, two employees may update the same product information in two different systems, resulting in conflicting information. Inconsistent data complicates data integrity and consistency. User or system changes can make data unpredictable over time. Unreliable data can confuse and mislead e-commerce companies, making data-driven decisions difficult. E-commerce backends also struggle with data loss. It can cause data loss. Furthermore, the promotion of e-commerce data management is greatly aided by big data technology (Lo et al., 2019), cloud computing (Vashi et al., 2017), and other advanced and cutting-

edge technologies (Yu & Duan, 2021). Big data technology can provide a real-time and dynamic service experience to the operation of e-commerce with information collection and analysis capabilities; in other words, it can have favorable control over the information to conduct e-commerce-related businesses better.

4. Conclusion:

E-commerce backend data integrity and consistency require technical and administrative solutions. Database transactions, unique constraints, referential integrity, data validation, version control, backup and recovery, and access control can solve data duplication, inconsistency, loss, and unauthorized access. As a result, data integrity and consistency help e-commerce companies make informed decisions, avoid legal issues, boost productivity, and gain a competitive edge.

References

1. Zhu, T., Li, J., Hu, X., Xiong, P., & Zhou, W. (2020). The dynamic privacy-preserving mechanisms for online dynamic social networks. *IEEE Transactions on Knowledge and Data Engineering*, 34(6), 2962-2974.
2. Anthony Jnr, B., & Abbas Petersen, S. (2021). Examining the digitalization of virtual enterprises amidst the COVID-19 pandemic: a systematic and meta-analysis. *Enterprise Information Systems*, 15(5), 617-650.
3. Amin, S., Kansana, K., & Majid, J. (2016). A review paper on E-commerce. In *TIMS 2016-International Conference, At Gwalior*.
4. Asan, H. (2022). Data security. *Artificial Intelligence Perspective for Smart Cities*, 253-276.
5. Shukla, S., George, J. P., Tiwari, K., & Kureethara, J. V. (2022). *Data Ethics and Challenges*. Springer Singapore Pte. Limited.
6. Yu, L., & Duan, B. (2021). Optimization of IoT-Based Sporting Goods Consumer Service Management System. *Complexity*, 2021, 1-13.
7. Sultana, M. (2019). Online Shopping System.
8. Singh, J. (2014). Review of e-commerce security challenges. *International Journal of Innovative Research in Computer and Communication Engineering*, 2(2), 2850-2858.
9. Vashi, S., Ram, J., Modi, J., Verma, S., & Prakash, C. (2017, February). Internet of Things (IoT): A vision, architectural elements, and security issues. In *2017 international conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC)* (pp. 492-496). IEEE.
10. Choi, Y., & Mai, D. Q. (2018). The sustainable role of the e-trust in the B2C e-commerce of Vietnam. *Sustainability*, 10(1), 291.
11. Lo, S. K., Liew, C. S., Tey, K. S., & Mekhilef, S. (2019). An interoperable component-based architecture for data-driven IoT system. *Sensors*, 19(20), 4354.
12. David, M. (2017). *Statistics for managers, using Microsoft excel*. Pearson Education India.