ADVANCED DATABASES

by

Name

Institution Affiliation

Professor’s Name

Course

Date

**Abstract**

The potential of advanced database and cloud technology on a company’s backup, performance, and disaster recovery is immense. The aim of this business report is to assess the Ideal Car Company’s database and recommend the best in-house solution for its business needs. With the right combination of technology, companies can ensure that their data is secure, their performance is optimized, and their disaster recovery plans are robust. Advanced databases are essential for companies to store and manage their data. By leveraging the power of the cloud, companies can ensure that their data is stored securely and is accessible from anywhere. Cloud technology can also use to improve a company’s backup and disaster recovery plans. By leveraging the power of the cloud, companies can ensure that their data is backed up regularly and securely. This allows companies to quickly recover from any data loss or system failure. Additionally, cloud technology can be used to create robust disaster recovery plans that can help companies quickly recover from any disaster.

**Table of Contents**

Cover Page--------------------------------------------------------------------------------1

Abstract------------------------------------------------------------------------------------2

Description of the tables in the database--------------------------------------------4-5

Views----------------------------------------------------------------------------------------5-8

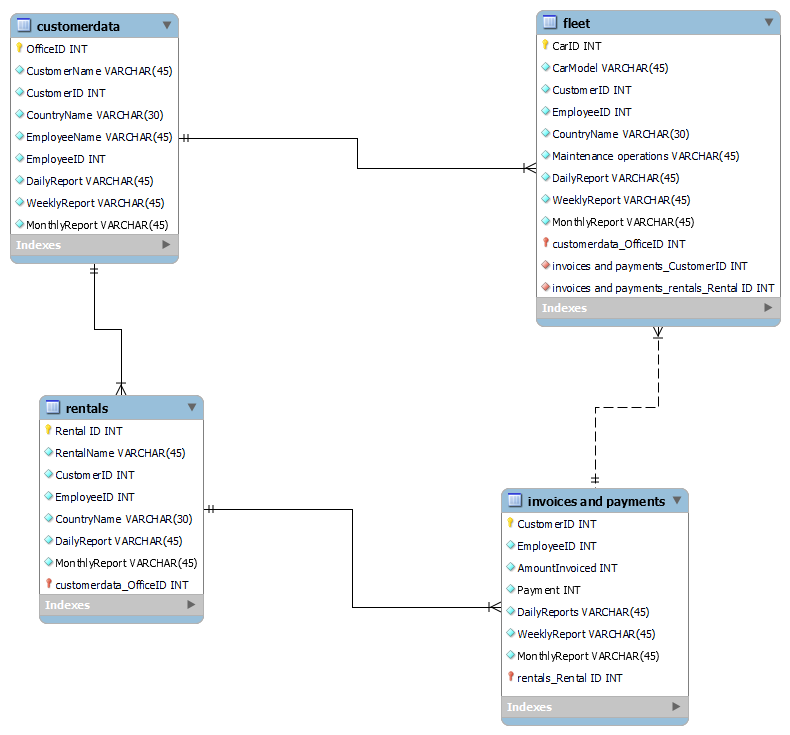
List of additional indexes---------------------------------------------------------------8-10

Security/Access rights-------------------------------------------------------------------10-13

Replication vs. cloud--------------------------------------------------------------------13-17

Bibliographies----------------------------------------------------------------------------17-19

**Description of the Tables in the Database**

****

The database above is a general structure of the company’s management system that will aid in addressing its business needs. The customer data table has columns such as the identity numbers of the customer, office and employee who served the clients. This tables relate with each other in terms of the similar columns present in each table. For instance, the customer identity number in the customer data and fleet tables indicate a specific client who has purchased or ordered a particular car model hence making it easy to trace the transactions.

Despite the tables having the daily, weekly and monthly reports, the invoice and payment relates with the other tables in the central database because the employee identity number is a column in each table hence making it easy to know the staff serving each client from different countries. The central database indicates the identity number and names of specific rentals, and the customers who booked or paid for the accommodation services. The fleet and rental tables relate with each other in terms of the country name that specific customer represented by an identity number hence enabling the company’s staff at different geographical locations to better serve clients and prevent errors when entering clients’ data on the system.

In terms of their functions, the customer data table provides details regarding the office number in a particular country that a specific customer received the company’s services. The table also has daily, weekly and monthly reports regarding the customers paying for cars and rental services. The fleet table shows the identity numbers of each car purchased, the employee who served the client and the customer buying the car or paying for the rental service. The table also provides the specific name of the country where the transactions occur. The rental table is important for showing the rental name, identity number of the apartment, specific customer and employee reporting the business transaction. The invoice and payment table indicate the customer identity the employee who records the transaction from a specific customer, the amount paid or invoiced to a client.

**Views**

The current solution has replaced the old reporting solution which uses a decentralized form of database at each branch and can only be accessed through a Microsoft Access application. Sending the daily, weekly and monthly reports by each employee can be tiresome and contribute to entering of inaccurate data in the database. According to research, the Joins and semi joins case study proved that the use of centralized database is effective when accessing data from several tables as compared top retrieving data from distributed or decentralized repository systems (Sharma, 2012). For example, the centralized database is effective in tracking every single car by using the identity number of the vehicles to check their details such as the mileage, fuel consumption, model and the year of manufacture. These details ensure that the Car ideal company delivers the right car specifications to the customers.

Another perspective is that the increasing number of customers at the different branches across the globe contributes to extra time consumption in the filling of customer forms since a new account has to be opened when a client visits another branch. The new centralized solution ensures that the initial account can be accessed at any of the company’s branches hence saving time retrieving the customer’s personal details. For example, the staff at every branch uses the customer’s unique identity number to record the transactions or the services offered to a customer’s regardless of their location hence preventing double entry mistakes in the system.

According to research, the fragmentation of decentralized databases can contribute to time wastage when retrieving data due to several segments that are related to the original data as compared to the centralized database which facilitates easy changing and tracking of data (UBAIDILLAH, 2020). For example, it will be very easy tracking customers because every customer has their own identity number which makes them distinct from other customers when purchasing vehicles and making rental payments. The company’s management team at the headquarters directly contact the employee in any given office of a specific nation in case there is need for clarification about a specific client or vehicle by simply tracking the identity number of the worker.

The firm aims at adhering with the car manufacturers’ terms of maintaining the cars to reduce the costs of repairing the vehicles by checking the levels of brake and engine oils, the wheel and steering balancing, and whether the windshield, brake pads and side mirrors operating. In terms of managing the car maintenance operations, the computerized system of every car in each branch is synchronized to the new database system hence alerting the staff in case of any sign of unexpected breakdown. Currently, each branch has saved about $60,000 on every car maintenance costs and uses about $30,000 on scheduled car breakdown operations. The staff at every branch document the daily reports of the cars and save them in the cloud for comparing with the weekly and monthly performance of the cars. The system is useful in determining the fleet of cars that the company can use every day or hire out to customers. Currently each branch in all the countries have approximately 750 fleet of cars and from this end month report, all the cars are fully serviced and ready for use. The company has hired out approximately 500 cars in each branch after several customers from all the countries especially companies dealing in corporate events, signing short-term and long-term contracts.

The unique rental IDs have been useful to the company because the staff can easily confirm from the system which apartments are occupied and which ones are not. For instance, all the rentals in New York are fully booked by international high-end customers who will be attending the Global Business Summit next week. The firm’s returning customers from this region referred their business associates from other regions to the company for accommodation services hence making the staff to assist the increasing number of clients to book for rentals in the neighboring towns. The company is experiencing a high increase in unique customers especially in regions that attract tourists such as the Bahamas and the Maldives hence all the rentals are currently booked now.

In terms of costs and profitability, the company has improved immensely after implementing the new database system. For instance, the company incurs scheduled car maintenance operations only since the car systems send alerts prior to experiencing breakdowns and the staff can monitor from the system which car models are performing greatly as compared to others. The Porsche, Land Rover and all Mercedes Benz series cars are performing well by covering the estimated mileage after hiring out to customers in all the thirty branches hence increasing the profits for the company since unique high-end customers prefer these car models to others. The company has so far made a net profit of approximately $200 million from the fleet of cars and rentals in each branch which is an equivalent of over $5 billion in total profits. In comparison to the last results prior to the implementation of the new solution, the firm made a total profit of approximately $1.5 billion due to the high costs of car maintenance.

**List of additional Indexes**

Implementation of the new solution is important because the outdated database system that the headquarters used in business reporting, experienced problems when two or several branches failed to send their reports on time. According to research, the centralized database management system is effective in terms of transmitting data on time since the users can access the same web browser regardless of their geographical location (Sareen, 2015). For example, the staff at all the branches have already sent invoices to customers and similar copies are transmitted to the headquarters for assessment. This process is facilitated through the ID numbers of the customers who receive emails requesting them to make payments and second email of their successful payments. Over 3000 customers in all the branches have made payments to the company hence the staff can track the transactions and know which customers are yet to honor their invoices.

The company’s decentralized system is effective because it saves data in terms of fragments but can cause confusion which may contribute to the entering of inaccurate data in the customer data, rentals and fleet maintenance operations tables. According to research, the distributed or decentralized database is effective in terms of storing data and reducing the costs of network communication and transmission (Cinar, 2016). Conversely, a distributed database can cost a company in terms of configuration difficulties hence increasing the total expenses of the company. For instance, the new system allows all the branches to transmit big data which may have an impact on the queries made. The branches in the United Kingdom transmitted big data last week during the Britain Expo due to the high numbers of customers who booked and paid for all the rentals in the region.

The Ideal Car Company has implemented the new centralized database because the outdated database system was not operated by a professional who understands how to manage and structure big data using tools such as SQL. According to research, one of the limitations that a decentralized system has is the lack of professional support hence increasing the risk of users entering inaccurate data (UBAIDILLAH, 2020). For instance, the SQL developer based at the headquarters provides the management with a report of the big data after assessing the business operations from the data transmitted by the employees in all the branches. Big data from branches in France indicate the changing mileage covered by various car models such as the BMW series hence the daily reports keep changing the queries in the database.

The outdated database did not show any relationships between the customers, fleet of cars, rentals and the payments made to the company. According to research, the NoSQL type of database are flexible since they are distributed but they cannot perform well by showing relational data through the queries made in the case of SQL database management systems (Venkatraman, 2016). For instance, the new database solution supports easy assessment of data by allowing the SQL developer to create tables and queries from the big data transmitted from the branches. However, research also proves that the flexibility nature of the NoSQL database systems allows easy retrieval and storage of data (Venkatraman, 2016). The company’s employees at various branches can transmit data easily to the new centralized system and also access data through the company’s web browser. For instance, the system shows that approximately 18,000 out of 30,000 customers in Australia have paid their rental charges whereas only 34 out of about 700 cars in the region have transmitted alerts to the system regarding signs of unexpected breakdown.

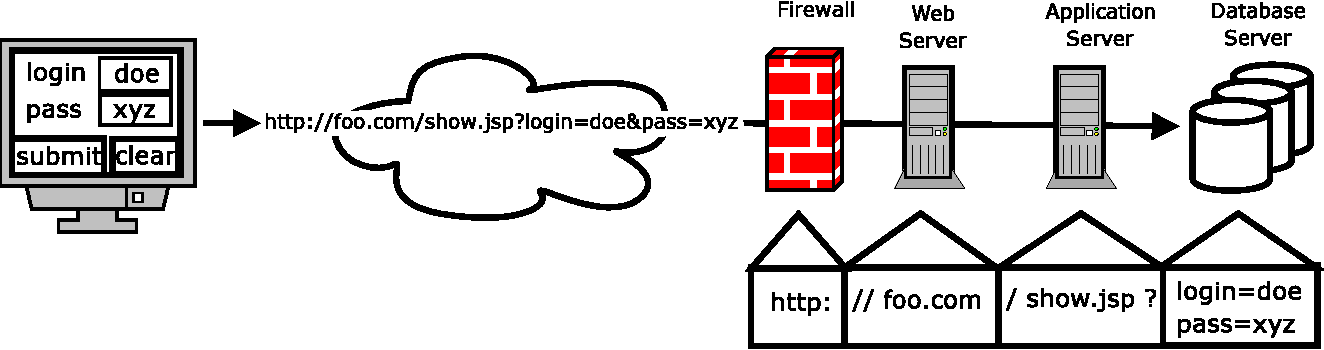
The high volume of data from different office branches in all the thirty nations can be confusing increasing the SQL developer’s chances of making errors when reporting the queries to the management. There is consideration for using a separate database that will monitor a specific type of data for accuracy purposes. According to research, the NoSQL MongoDB is an effective system because it prevents data loss especially in cases where there is high transmission of data hence ensuring partitioning of data across different web servers (Patil, 2017). Such performance by the centralized database systems is a disadvantage which can negatively impact the entire business operations of the Ideal Car Company. For instance, the fleet maintenance data of branches in Morocco can be transmitted across different web browsers to prevent damage to the system which can contribute to big data loss.

**Security/Access Rights**

The company’s employees in each of the branches have full access to the database hence the risk of data landing to unauthorized hands it too high. The new centralized database solution is effective in terms of data storage and security because there is only one person (database consultant) who will have the full access rights and approved by the firm to operate the database. However, due to the different geographical locations of different employees in all the branches, it will be easy to access data stored in the cloud through the company’s web browser. According to research, multiple users can access data virtually from a centralized database integrated with the cloud by using a web interface (Parast, 2022).

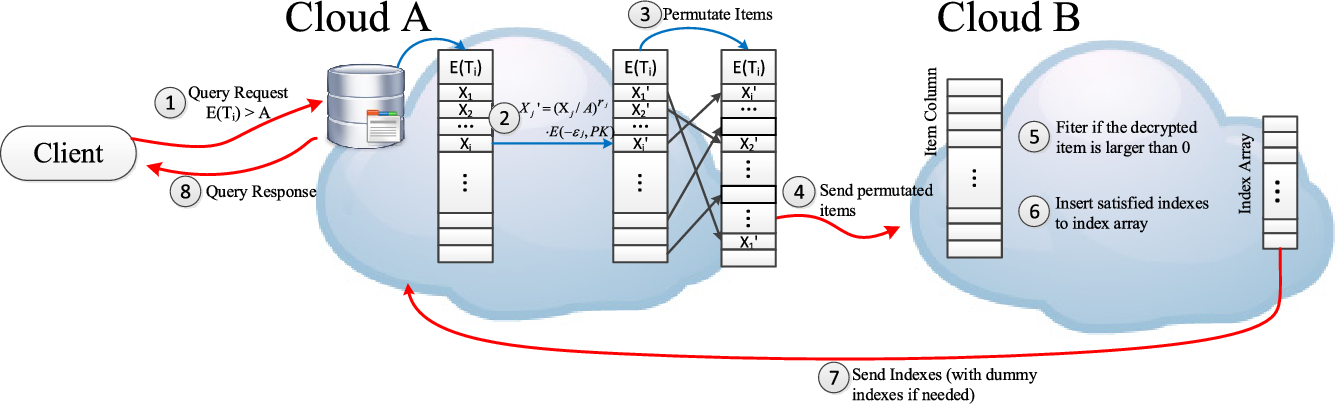
The new database solution is vulnerable to multiple threats such as cyber-attack which exposes confidential data to unauthorized personnel who may tarnish the company’s image. Research proves that the SQL injection is among one of the cyber-attacks to a database and occurs when the attacker manipulates the query statements when prompting to access data in the database (Alghawazi, 2022). For instance, an attacker may mimic the customer registration details when trying to book for rentals with an aim of accessing the data stored in the database. One of the methods that Ideal Car Company can implement in preventing SQL injections is the Analysis and Monitoring for Neutralizing SQL injection (AMNESIA). According to research, AMNESIA detect any form of cyber-attack by generating and interpreting queries before the data is transmitted to the database system (Alwan, 2017). For instance, the company can use the below AMNESIA model for detecting any form of SQL injection.

**Fig 1**. (Alwan, 2017)



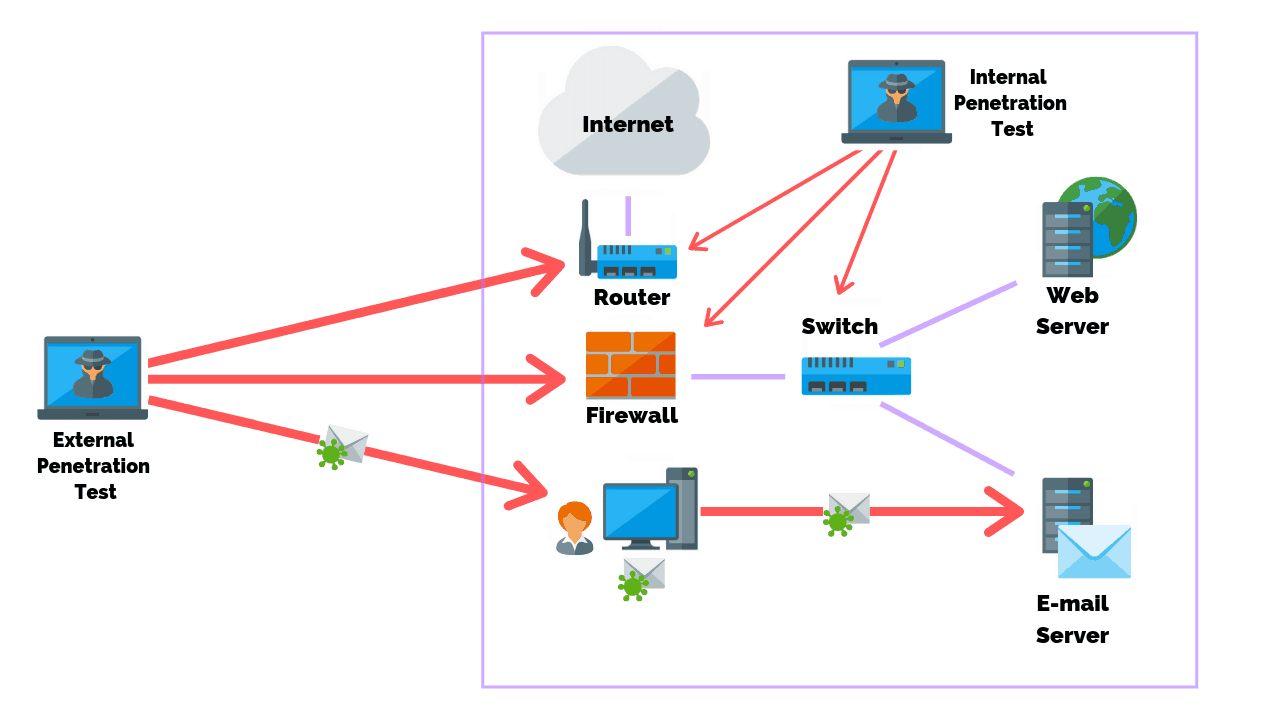
According to research, the centralized database is more secure due to its compatibility form with the cloud systems such as the Amazon Web Service (AWS) which can store and protect data (Parast, 2022). The integration of the cloud and the company’s new database solution will be effective in terms of saving essential files that cannot be easily accessed by unauthorized personnel. According to research, a centralized database is more secure because all the data is stored in a single repository hence preventing the duplication of information when data is accessible by multiple users in a distribution database (Mohammed, 2017). The cloud will be important for the Ideal Car Company because it will act as a backup in case the data in the centralized database solution is lost. However, the data owners such as the company, are not control of the safety of their stored data in a cloud system such as AWS since it is a shared cloud platform. According to research, an organization can protect its big data by implementing a two-cloud scheme whereby one cloud system stores the query statements whereas the other cloud makes computations to establish whether the command queries are encrypted and have their own security keys in relation to the data owner’s specifications (KUMAR, 2018).

**Fig. 2** (KUMAR, 2018).



The Ideal Car Company can also implement the penetration testing security measure to determine the types of threats or the new solution’s vulnerabilities. The pen testing techniques are important in determining the weaknesses in the database which can be hot spots for SQL injection hence determining the best method for securing the system (Singh, 2020). For instance, the staff at each branch sent emails in advance to all the customers via their emails to alert them of a system maintenance operation that will be taking place on March 18, 2023 starting from around 8 a.m. to noon. The company chose this date because it was on a weekend and only few clients require the firm’s services.

**Fig. 3**  (Singh, 2020)



**Replication vs Cloud**

Replication or in-house refers to the process of copying or transferring data from the outdated database to the new database solution (www.manageengine.com, (n.d.)). Data replication can be implemented through several methods. First, complete table replication copies all forms of information, from the source to the target. This technique is helpful if records are frequently hard removed from a source or if the source lacks unique keys or timestamps. One of the downsides of this approach is that complete table replication uses more CPU resources and puts more strain on the network than duplicating only modified data (Alison, 2022).

Another method of implementation is the change data capture which entails the new database receiving real-time updates anytime data is modified after the data replication program creates full first copies of the data from the origin to the destination. Fewer rows are replicated every time data is modified, making this a more effective replication method (Software, 2022). Data is replicated via snapshot replication exactly as it is at any particular time. Snapshot replication, in contrast to other techniques, ignores any data changes that have occurred in the interim. Because data changes are more likely to occur infrequently, this replication method is utilized.

Incremental replication based on timestamps only updates data that has changed since the last update. Timestamp-based replication, in contrast to full table replication, copies fewer rows of data on each update, making it more effective (Naeem, 2020). This method's limitations include its inability to reproduce or identify hard-deleted data and its incapacity to update records with non-unique keys.

Data replication ensures data is readily available to users via a web browser application hence making data accessible. This is especially helpful for multinational companies with offices spread out around the globe. As a result, data is still accessible to other sites in the event of hardware failure or any other problem in one place. The primary advantage can be in terms of data protection and disaster recovery. In the case of a disaster, hardware failure, or system intrusion that could endanger data, replication ensures that a reliable backup is stored (www.manageengine.com, (n.d.)). Replication of data can improve and accelerate server performance. Users can get data much more quickly when businesses operate various data copies across multiple servers. Moreover, administrators can use fewer processor cycles on the main server for resource-intensive writing activities when all data access procedures are routed to a replication. Storing duplicates of the same information in different locations can lower data access latency by retrieving the necessary data from the site where the process is being completed.

In terms of comparison with cloud, replication preserves more recent data than backup. Since backups are not always made, a most current backup can take a long time before change is made. Contrary, replication often takes just a few minutes to complete, minimizing the quantity of data lost. In comparison to backup, replication requires less recovery time. It may take a while and result in delays to fully restore all data from a backup (Alison, 2022). After a disaster, mission-critical and customer-facing apps can continue to run without interruption thanks to replication, whose recovery duration can be measured in minutes. Malware can spread to other websites via replication, which happens as data is copied around the system. Moreover, restoring to a malware-free condition might not be possible without sufficient backups. Nevertheless, with backup, you are safe because you are producing totally separate copies of the data at a certain moment in time.

One of the disadvantages of data replication is that an organization should be willing to incur the costs of purchasing hardware such as hard disks for storing big data (Alison, 2022). For instance, the Ideal Car Company can invest more on the data storage hardware if the management decides to split data into several servers with an aim of preventing the centralized database management system from damage which can cause big data loss. Therefore, data replication is costly since an organization will require additional software for data security purposes.

Cloud solutions are essential in storing and securing the confidential data which can be retrieved by users through a web application. Implementation of cloud solutions requires several steps such as understanding the security policies recommended by the service providers to ensure that the data owners know who should access the data (Jain, 2017). For instance, the Ideal Car Company can decide on using the Amazon Web Server (AWS) cloud system but make sure the responsibility of who can access the data strictly remains on the organization. The cloud solutions have several advantages and disadvantages.

Cloud solutions are cost-effective since they don’t need the users to have several hardware or softwares to access data. According to research, cloud computing solutions enable internet users to access data using their smart devices through a web server hence reducing the costs for softwares and hardware (Jain, 2017). The Ideal Car Company stores big data in a cloud system hence only staff can access the data for effective decision-making processes. This solution is user-friendly and does not expose data to any form of SQL injection as compared to the data replication solution which increases the vulnerability of softwares and hardware to cyber-attack.

The cloud database systems such as the Google Cloud SQL do not require any storage facility or space as compared to the replication solutions. According to research, the Google Cloud SQL is versatile and does not require the operations of the users hence this cloud solution can be accessed via the MySQL database (Bhatti, 2017). In the case of Ideal Car Company, all the employees from different branches in the thirty nations can access the cloud server via the web application in their smart devices but they cannot make changes to the web. This means that cloud server services is free to the users since the programs are remotely owned by the server provider.

One disadvantage of cloud solutions is the slow connection when accessing data especially if there are several users trying to connect to the server (Haris, 2018). The centralized database of the company can be effective but may be slow when all the employees try accessing data at the same time hence the organization may reconsider investing in an additional cloud server. Users cannot access the data from the cloud server without an internet connection. According to research, accessing the cloud database management system will require the users to have a 3G/4G/5G internet connection to access the programs and data stored in a cloud server

The best recommendation for Car Ideal Company is to implement the cloud solutions in its business operations by integrating with new centralized database. Research proves that the cloud solutions are more secure as compared to replication solutions since a two-cloud systems is effective in storing and conducting computations to establish whether the command queries are encrypted and have their own security keys in relation to the data owner’s specifications (KUMAR, 2018). Replication solutions are costly because they require hardware for storage as compared to the cloud servers which are free to use and have no hidden storage costs. Therefore, the company will operate effectively with low probability of losing data through cyber-attacks if the cloud server is part of the new solution.

# Bibliography

Alghawazi, M. A. D. a. A. S., 2022. Detection of sql injection attack using machine learning techniques: a systematic literature review.. *Journal of Cybersecurity and Privacy,* 4(2), pp. 764-777.

Alison, 2022. Database Backup vs Replication: What’s the Difference?.

Alwan, Z. a. Y. M., 2017. Detection and prevention of SQL injection attack: a survey.. *International Journal of Computer Science and Mobile Computing,* 6(8), pp. 5-17.

Bhatti, H. a. R. B., 2017. Databases in Cloud Computing: A literature review.. *International Journal of Information Technology and Computer Science,* 9(4), pp. 9-17.

Cinar, O. G. R. a. Y. A., 2016. Database security in private database clouds. International Conference on Information Science and Security (ICISS) .. *IEEE. ,* pp. 1-5.

Haris, M. a. K. R., 2018. A systematic review on cloud computing. , 6(11), pp.632-639.. *International Journal of Computer Sciences and Engineering,* 6(11), pp. 632-639.

Jain, S. a. A. M., 2017. Comparative Study of Traditional Database and Cloud Computing Database.. *International Journal of Advanced Research in Computer Science.,* 8(2).

KUMAR, M. B. K. S. M. a. S. A., 2018. Providing Privacy for Numeric Range SQL Queries Using Two-Cloud Architecture.. *International Journal of Scientific Research and Review.,* p. 39.

Mohammed, A. a. S. B., 2017. Centralized Database: A Prerequisite for Security and Sustainable Development in Nigeria..

Naeem, T., 2020. *Data Replication - An Overview and Benefits. .*

Parast, F. S. C. N. S. Y. H. K. K. a. H. S., 2022. Cloud computing security: A survey of service-based models. Computers & Security. Volume 114, p. 102580.

Patil, M. H. A. T. C. a. P. P., 2017. A qualitative analysis of the performance of MongoDB vs MySQL database based on insertion and retriewal operations using a web/android application to explore load balancing—Sharding in.

Sareen, P. a. K. P., 2015. NoSQL Database and its comparison with SQL Database.. *International Journal of Computer Science & Communication Networks.,* 5(5), pp. 293-298..

Sharma, M. a. S. G., 2012. Analysis of Joins and Semi-joins in Centralized and Distributed Database Queries. *International Conference on Computing Sciences IEEE. ,* Volume 5, pp. 15-20.

Singh, S. a. K. A., 2020. Detection and prevention of sql injection.. *International Journal of Scientific Research & Engineering Trends.,* 6(3), pp. 1642-1645.

Software, S., 2022. What is Data Replication, and Why Is It Needed?.

UBAIDILLAH, S. a. A. N., 2020. Fragmentation techniques for ideal performance in distributed database–a survey.. *International Journal of Software Engineering and Computer Systems,* 6(1), pp. 18-24.

Venkatraman, S. K. K. S. a. V. R., 2016. SQL versus NoSQL movement with big data analytics.. *International Journal of Information Technology and Computer Science,* 12(8), pp. 59-66.

www.manageengine.com, (n.d.). *What is Data Replication? | Types of Data Replication - ManageEngine Device Control Plus..*