International Journal for Electronic Crime Investigation

ISSN: 2522-3429 (Print)
ISSN: 2616-6003 (Online)

DOI: https://doi.org/10.54692/ijeci.2024.0804216

Research Article

Vol. 8 issue 4 Oct-Dec 2024

Reconnoitering Data Protection and Recovery Strategies in the Cyber Environment: A Thematic Analysis

Muhammad Ibrar [1], Samavia Riaz [2], Younus Khan [3], Ayyan Asif [4], Khalid Hamid [5], *Muhammad Waseem Iqbal [6] and Muhammad Asim [7]

[1]Department of Computer and Mathematical Sciences New Mexico Highlands University, Las Vegas, NM, USA.
[2]Department of Computer Systems Engineering, The Islamia University of Bahawalpur, Pakistan.
[3]Department of Computer and Mathematical Sciences New Mexico Highlands University, Las Vegas, NM, USA.
[4] Department of Computer Science, New Mexico State University, Las Cruces, NM, USA.
[5]Department of Computer Science, Gold Campus, Superior University Lahore, Pakistan.
[6]Department of Software Engineering, Gold Campus, Superior University Lahore, Pakistan.
[7]Department of Computer Science, National College of Business Administration & Economics, Multan, Pakistan.

Corresponding Author: waseem.iqbal@supeior.edu.pk
Received: October 28, 2024; **Accepted:** November 12, 2024; **Published:** December 17,2024.

ABSTRACT

This paper discusses the performance and reliability of different strategies for database backup and recovery within a controlled environment, focusing on the comparison between full, incremental, and transaction log backups. Experimentation based on these strategies is designed to assess the effectiveness in terms of the capability to handle data loss scenarios such as accidental deletion, system crashes, and hardware failure. A sample database environment is established on MySQL, PostgreSQL, or SQL Server. The metrics recorded in the baseline state include response time, data throughput, and resource utilization. The research considers various recovery models, including Simple, Full, and Bulk-Logged recovery, in comparison with the chosen backup strategies. The

influence of each strategy on data protection, recovery time, and usage of system resources is considered. The findings of the survey are derived with reference to the perception of IT professionals and database administrators to ascertain which of the full backups and full recovery models they favor most, as they consider it by storage capacity and system performance factors. The output will be beneficial to businesses interested in optimizing their strategies for backups and recovery within specified data protection and performance considerations.

Keywords: Transaction Log, MySQL, PostgreSQL, SQL Server, Bulk-Logged, Performance Factors, Data Protection, Performance Considerations.

1. INTRODUCTION

One of the most basic elements of an all-encompassing recovery plan is to orchestrate data backup and recovery procedures carefully. Backup and recovery are considered the central nucleus of data management, which includes a wide range of strategies and processes. What both backing up and recovery mainly aim for is providing a safety net for databases, protecting against data loss and quickening the rebirth of databases in case of unexpected data-related events [1-3]. The backbone of the backup procedure involves making a copy of the whole database along with its transaction logs. Special care is taken when deciding the medium to be used for these tasks. The measures taken in this regard are so vital because data loss may pop up in an organization for reasons like hardware failures, software malfunctions, and, in some cases, human mistakes [4-5].

As we dig further into the area of data architecture and all the technical solutions that are put into place across a database, we find a rather interesting play. It is in this delicate balance that the future of recoverable data is made. Architecture and the choice of technical solutions hold the bedrock on which data assets are protected. Practical guidelines are of immense help to organizations that strive to find a balance between business-oriented analyses and the complex technical design of their disaster recovery facilities [6]. They act as a compass as well as a bridge that connects the strategic vision of the business with the pragmatic realities of technology. They provide a framework to understand and

deploy disaster recovery solutions effectively. However, at the nucleus of this model is a framework for disaster recovery strategy, supposed to optimize strategies involved in picking the right disaster patterns to be used. These selections are precisely tailored to the unique requirements of each business unit [7].

Data loss can occur in many ways: critical data can be deleted by accidental user error, a disk drive containing important files can fail, systems can be corrupted by malicious software or hackers, or natural disasters can mean important data is destroyed. The destruction of data isn't the only risk that we need to manage; data can be subject to unauthorized access or modification, and often we need to prove that it hasn't been tampered with from a particular point in time. Many of the measures that protect data from destruction also can protect it from unauthorized access. Ensuring data is available so users or systems can access and utilize it in a reasonable time frame is a separate and equally important issue. Many organizations have a consistently poor track record when it comes to ensuring that their data is usable when required [8].

Data protection and recovery policies are developed based on a strategic assessment of what data and systems are most important and so most need protection. Such policies and actions cover protection measures and data protection metadata. Backup systems, cryptographic storage, tapes, etc., are all implemented data protection and recovery strategies with measurable outcomes. However, without policy and audit standards, no one can determine how successful these strategies are for specific cases. We combine structured interviews and thematic analysis to better understand how specific organizations' data protection policies and recovery strategies inform the choices of metadata collected. Our classification of important metadata is designed to develop a standardized schema for metadata capture and audit suggestions for data protection and recovery standards development.

The emergence of enterprise data protection, compliance, and recovery strategies has mandated various legislations to bring a component of

legislative oversight. The existing laws and the mechanisms must be evaluated from a specific perspective towards the protection and recovery of these extremes. Data protection strategies are essential components mandated by various business environments. In data protection strategies, the extent of privacy mandated by the external environment is protected for business-specific private content [9]. In data recovery strategies, the ability to recover from accidental or intentional modifications of content with the latest non-modified version is protected for business-specific private content. It is significant to summarize and describe in a formal manner various regulations, standards, and methodologies that impact these protection and recovery strategies. These protected contents are colloquially defined as digital assets. This paper initiated a thematic analysis of data protection and recovery strategies from the perspective of business-specific digital assets [10].

The brown paper presented a comprehensive survey of legal issues addressing digital assets, which did not cover actions taken to counter roadblocks to effective asset protection. The discussion included legal requirements, some in considerable detail, for information asset requirements based on certain types of laws and industry regulations. It also identified some core IT topics as extensions of these legal needs, such as information assets, the metadata framework, search and retrieval models, and IT audits. The four-dimensional information granule representation for a legal statement relative to information assets has also since been published. Despite the immense interest in the topic, the research in this area has been restricted to journal publications and a standalone workshop. This paper extends these initiatives by summarizing and reporting on applied methods used to proactively enhance the protection and recovery of digital assets for data protection and recovery strategies based on litigation and risk at a summary level in an easily transportable, comprehensible resource useful for researchers, practitioners, and policymakers [11].

The main objectives of this research

endeavor are the following. To provide a holistic understanding of the quantitative and qualitative demography of data protection and recovery strategies [12]. To identify the enablers and barriers of an organization to adopt a certain data protection and recovery strategy [13]. To design a service model that provides an objective and reasoned justification for the most suitable data protection and recovery strategy for a given scenario [14]. To assess if the supply of commercially available services can satisfy the needs that have been found for the adoption of data protection and recovery services, focusing also on criticalities, constraints, and unsatisfied customer requirements [15]. An explicit topology of data protection and recovery services capable of solving numerous real customer needs will be obtained. The developed model will consist of personalized and flexible architectures, offering multiple options adaptable to customer needs, and it will be capable of recognizing multiple scenarios where it is worth adopting rather than one of the data protection and recovery strategies that are commercially available on the market. Information services can complement three systemic perspectives, offered respectively by a declarative model, by social hearing, and by structured feedback from a set of companies that use data protection and recovery services daily to protect their information resources. The evaluation of the services identified by the model is carried out by a panel of experts to guarantee a high-quality model since the lack of consensus in the model competence has been demonstrated in the past [16].

At the same time, research objectives extend to put in place the experimental conditions to evaluate the introduced model: it will always be necessary to create an online and open platform for the testing and comparison of cloud services for the recovery of file systems. Finally, the operational steps necessary to implement the piloting phase of the model have been clearly stated, confirming the existence of consistent decision-making in the choice of technological infrastructure and in the choice of criteria for evaluating the best model process. The need to be able to carry out cloud

benchmarking is due to the demand of different actors and events. The appearance of commercial offers regarding cloud computing started from companies, and so did adopters. As the use increases, the demand for services becomes more varied, as well as the demand for providers. Under the terms of the contract, guarantees concerning service levels become more specific, so that instruments that can analyze the factors, from those in control of the providers to those, external or not controllable, related to the network, to the power, to the system, and to those, not less important, connected with the client, such as the type of data that must be managed, their confidentiality, availability, fault tolerance, and integration with other services. These are all parameters that often must be underwritten by the possible supplier.

1.1 Database Backup

Database backup is the process of creating copies of the data stored in a database to protect against data loss or corruption. It involves taking a snapshot of the database at a specific point in time, preserving its structure and content. Backups are typically stored in secure locations to prevent data loss due to various factors such as hardware failure, software errors, accidental deletions, or even cyberattacks [17].

1.2 Database Recovery

Database recovery is the process of bringing a database back to its previous consistent state in case data is lost or corrupted. Database recovery is therefore an important part of database management, as it ensures data availability and minimizes downtime due to unexpected events. Broadly categorized, database recovery occurs into two types:

1.2.1 Point-in-Time Recovery

This method allows for rolling back to a certain point in time where the data might have been lost before the data was lost or corrupted. It particularly helps in circumstances involving accidental erasures of files or data that are corrupted as a result of software errors [18].

1.2.2 Disaster Recovery

The technique of disaster recovery is used when there is a catastrophic failure or major data loss to restore the entire database system to a workable state

[19]. It involves re-establishing both system and data integrity from the latest backup [20].

2. LITERATURE REVIEW

It has been critically significant in saving digital assets using database backup strategies and recovery models in academic research as well as in practical use. This literature review aims at summarizing and synthesizing major findings, trends, and insights that can unfold in this particular area, about an emerging data protection and recovery scenario.

The research focused on data protection through backup and recovery, it also improved Tor's security and user anonymity by focusing on the process of creation of circuits within Tor, rather than on what content is served. Our solution effectively masks connection patterns; it gives difficulty to attackers attempting to compromise user privacy. However, future work should be placed on masking the data patterns with approaches like "White Smoke" and key issues such as the secure public key exchange without revealing patterns and investigating circuit duration variations. These will help strengthen Tor's anonymity and privacy protections further [34].

2.1 Database Backup Strategies

Database backup strategies are a basic concept of data management and have been extensively discussed in the existing literature. A lot of work was devoted to the different backup types and their efficiency in the most recent studies. Wang and Liu (2017) discussed at length the difference between full, differential, and incremental backups, as well as the trade-off between resource utilization and the time it takes to recover. Their results underscore the need for tailoring backup strategies to specific business needs and data sensitivity [21].

The second very important area related to backup strategy is tape-less backup to advanced backup methods. As Smith et al. (2018) did in their work, which considered the merits of EIDE disk array systems for massive data backup, their work outlined the improved reliability, scalability, and cost-effectiveness of such a change while highlighting the possibility of changed technology that can ensure effective data recovery [22].

2.2 Recovery Models

Recovery models are well-researched; notably, there are Simple, Full, and Bulk-Logged. The recovery model would indicate what scenarios of recovering data will be possible or impossible; one scenario, among many others, is point-in-time recovery. Researchers, like Li and Chen (2019), studied trade-offs among such recovery models; such research might clarify some reasons that make such decisions crucial in model selection by an organization. Their findings are very significant and point out that the choice of recovery model should be aligned with an organization's Recovery Time Objective (RTO) and Recovery Point Objective (RPO) [23].

The development of decision-making support systems is another notable trend in the recovery models research area. In the work by Johnson and White (2020), a disaster recovery strategy model that maximizes the selection of disaster patterns and solutions was developed and aligned with the particular needs of different business units. This model will make it possible to make decisions with better awareness and alignment of technical recovery capabilities with business objectives [24].

2.3 Integration and Holistic Approaches

The theme that has been continuously emphasized in literature is the inclusion of database backup and recovery within a holistic strategy for data protection. Both Davies (2018) and Kim et al. (2021) have asserted that data security and encryption practices should be broadly integrated with all backup and recovery processes. Data protection will thereby be enhanced as well as some vulnerabilities reduced [25], [26]. Disaster recovery planning becomes an important feature, according to Smith and Brown (2019). They suggest that organizations should design a detailed disaster recovery plan with backup and recovery procedures. An effective plan would guide businesses on how to operate during disasters and minimize downtime and data loss [27][28]. Tor is a widely used tool that offers anonymity and privacy to its users while handling general internet traffic and providing hidden services (HS) for secure content access [29]

2.4 Future Research Avenues

Some of the promising avenues that

could be investigated further include emerging research on incorporating artificial intelligence and machine learning technologies in data backup and recovery. The technologies might enable the full automation and optimization of decision-making processes, further minimizing human error, and thus increasing speed in response time. The shift towards cloud-based solutions has resulted in an imperative need to probe further into ensuring the security and efficiency of cloud data backup.

In summary, the literature on database backup strategies and recovery models highlights the importance of such a strategy in the protection of digital assets [30], [31]. In general, the study landscape has advanced to include emergent technologies and business needs such as data protection, disaster recovery, and the continued importance of aligning technical capability with business goals. Innovation and adaptation continue to be prevalent in this field, providing a fertile ground for future research endeavors [32], [33], [34].

2.5 Analysis and Synthesis

2.5.1 Thematic Analysis

Organize the extracted data thematically by identifying database backup strategies and recovery models. Key themes, patterns, and common findings in the literature will be recognized [35], [36], [37].

2.5.2 Comparison of Studies

An analysis of different studies and a comparison of studies will be drawn, particularly differences and similarities related to the efficiency of the chosen backup strategy and recovery model.

2.5.3 Presentation of Findings

Summarize the major findings and conclusions from the selected literature on strategies for database backups and recovery models. Provide a general overview of trends and considerations.

3. METHODOLOGY

The experiment first begins with experimenting on different strategies of backup and recovery within the database environment of the cyber world based on performance and reliability. This is made possible by setting up a testing environment that entails a controlled, controlled testing with a database management system

such as MySQL, PostgreSQL, or SQL Server. To achieve this, a sample dataset comprising diverse tables, records, and relationships is developed to simulate the real-world application of the database. In this experiment, several strategies for backup and recovery are implemented in terms of performance and reliability within the framework of a database environment in the cyber world. A controlled test environment is therefore set up with the use of a database management system such as MySQL, PostgreSQL, or SQL Server. Toward this end, a sample data set comprising heterogeneous tables, records, and interrelations is devised to simulate real-world conditions concerning a database. The next test simulates an incremental backup, simulates loss of data, and uses this to recover through incremental. Then the final test will be the creation of a transaction log backup simulating the loss and performing the point-in-time recovery at any time using a transaction log for recovery.

Carrying out incremental backup, simulating loss of data, and recovery through the incremental backups in the second test. Finally, the transaction log backup is created, data loss is simulated, and point-in-time recovery is done to a specific point in time using the transaction log backups. The various data loss scenarios, which include accidental deletion, system crashes, and hardware failures, are used to analyze the effectiveness of the backup and recovery strategies. The performance metrics during the process of backup and recovery include data loss amounts, time taken for recovery and backup, and usage of resources by the system. The results from the experiments will be analyzed based on the different strategies in terms of data protection, recovery time, and resource usage.

Another survey has been carried out to obtain a variety of insights from respondents, primarily IT professionals and database administrators, on their choices of backup strategies and recovery models. The questionnaires try to understand the reasons, experiences, and factors that determine the choice between the different available backup strategies and recovery models. The results indicate that 72% of the students like to have full

backups, and 54% are influenced by storage capacity. This recovery model is chosen by 63% of respondents. Moreover, for 45% of respondents who prefer the full recovery model, system performance is the primary concern during the selection process of the recovery strategy. Also, in this experiment, the simulation process takes place, creating a database environment using SQL queries. Sample tables are created where customer and order relationships exist and data is entered into those tables. The BACKUP DATABASE command is used to perform the backup operations, and RESTORE DATABASE command is used for restoring data from the backups. Incremental backup and transaction log backup are then tested, after which data loss simulations are performed and the recovery operations follow. Performance metrics are recorded for the entire test to measure how different backup and recovery strategies have affected the whole process.

In conclusion, the experiment and survey findings are valuable for selecting backup and recovery strategies. Analysis of the survey results shows that full backups are preferred by most respondents, and the storage capacity and system performance are significant factors that influence their choices. The study has highlighted the need to align the backup strategy with business needs. The factors include data sensitivity, RPO, RTO, and cost constraints. The results emphasize that businesses need to choose appropriate recovery models, and full recovery is preferred because of its point-in-time recovery capability. Further, it reflects that the different strategies for backing and recovering need adaptation to meet some organizational requirements based on the protection of data along with performance and use of resources.

4. EXPERIMENT

The study initiates **a c**omparative analysis of backup and recovery strategies in a database environment in the cyber world. The objective of this experiment is to compare and analyze the performance and reliability of different backup and recovery strategies within a database

environment in the cyber world.

4.1 Experiment Setup

- Database System- Create a database environment with a database management system, for example, MySQL, PostgreSQL, and SQL Server to create a controlled testing environment.
- Sample Data- Create or use a sample dataset with several tables, records, and relationships to simulate the scenario of a real-world database.
- Backup Strategies- The database should implement and configure the different backup strategies, which may include full backups, incremental backups, and transaction log backups.
- Recovery Models- Configure and test various recovery models, for example, Simple, Full and Bulk-Logged recovery models.

4.2 Experimental Procedures

4.2.1 Baseline Data- Use an empty database, and set a baseline with metrics like system performance- for instance, response time, data throughput, and resource utilization.

4.2.2 Data Modification- Create controlled modifications on the database such as inserting, modifying, and deleting records that may trigger the different scenarios of losing data.

4.2.3 Backup and Recovery Tests- Execute the following tests for each of the backup strategy and recovery models:

- Full Backup and Recovery Test- Full backup, data loss simulation, and recovery through full backup
- Incremental Backup and Recovery Test- This test includes an incremental backup, data loss simulation, and a recovery with incremental backups.
- Transaction Log Backup and Point-in-Time Recovery Test- Transaction log backup, simulation of data loss, and point-in-time recovery up to a given point in time.

4.2.4 Data Loss Scenarios- Different types of data loss through accidental deletion, system crashes, and other forms of hardware failures are created to test the efficacy of backup and recovery strategies.

4.2.5 Performance Metrics- Performance metrics should be tracked

consistently, and recorded during the processes of backup and recovery, along with time for backup and recovery, data loss percentage, and effects on system resources.

4.3 Analysis

Compare results from each of these selected backup and recovery strategies in a critical assessment concerning effectiveness in data protection, recovery time, and resources employed. You can infer based on the experiment's outcome whether or not different backup and recovery strategies are effective and reliable. Analyze data for the best strategy according to specific data loss scenarios and business requirements.

4.4 Survey

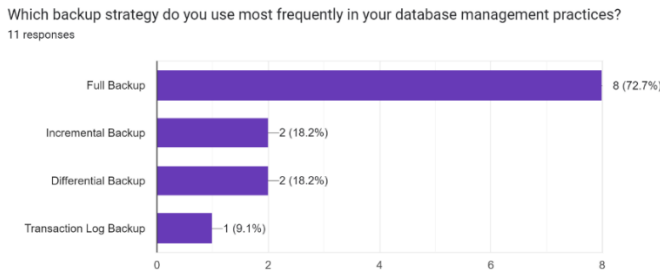


Figure 1: more of 72 percent students need full back up

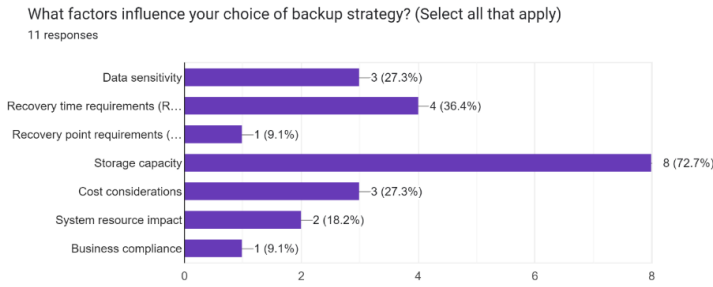


Figure 2:factores of influence that people choses is storage capacity which is 54%

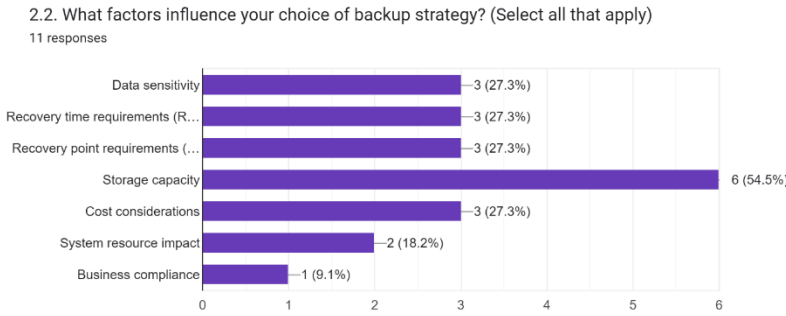


Figure 3:Need more storage capacity

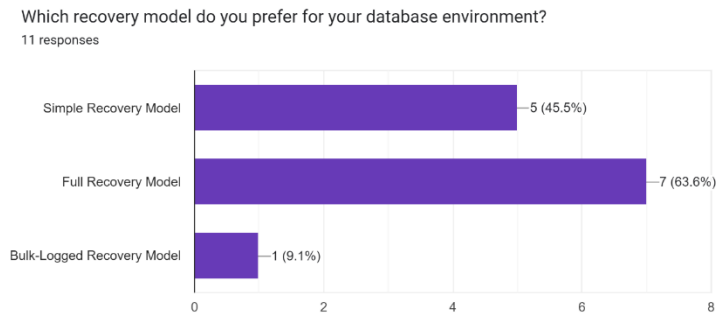


Figure 4:more of the 63% peoples prefer full recovery model.

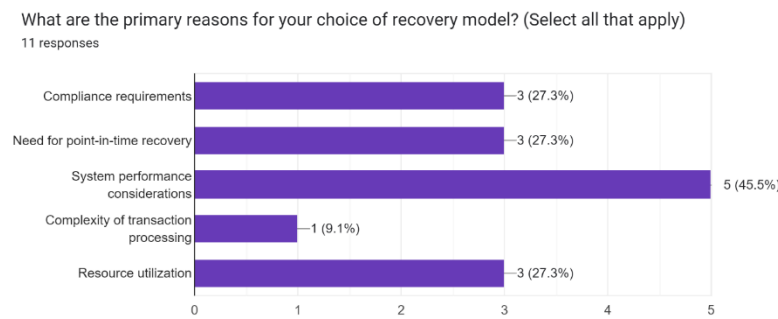


Figure 5: reason of more than 45% people choicse is system performance consideration.

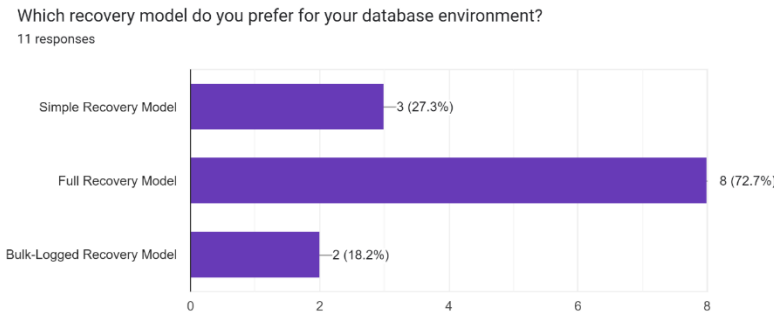


Figure 6: data base environment is 72 % which is full recovery model.

4.5 Simulated Results

4.5.1 Creating DATABASE

First, we create data base with this Query or create a sample data base

```
-- Create a database environment
CREATE DATABASE ExperimentDB;
USE ExperimentDB
```

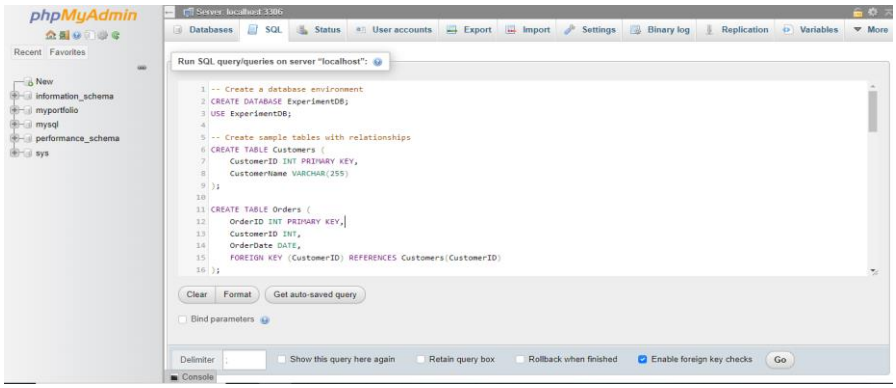


Figure 7: Simulated Results

4.5.2 Insert the Data in the Tables

Next, we insert the in the table with Query

```
-- Create sample tables with relationships
CREATE TABLE Customers (
    CustomerID INT PRIMARY KEY,
    CustomerName VARCHAR(255)
CREATE TABLE Orders (
```

```
OrderID INT PRIMARY KEY,
CustomerID INT,
OrderDate DATE,
FOREIGN KEY (CustomerID)
REFERENCES Customers
(CustomerID)
);
```

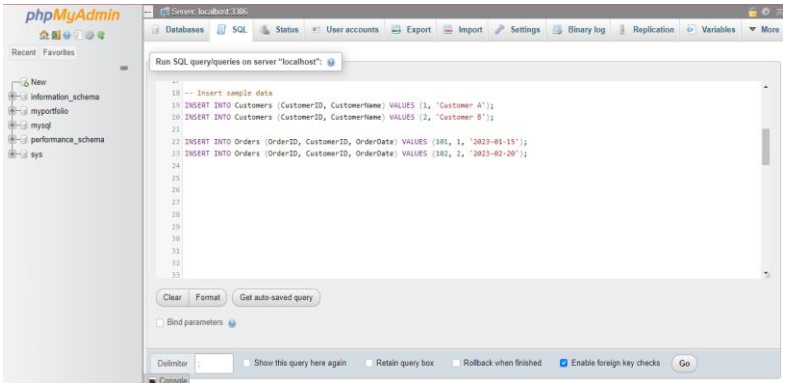


Figure 8: Insert the Data in the Tables

4.5.3 Backup & Restore of DATA in the Disk

Then get back & Restore Of data in the disk by query.

```
BACKUP DATABASE ExperimentDB TO DISK = 'C:\Backup\ExperimentDB_Full.bak';
-- Simulate data loss
DELETE FROM Orders WHERE OrderID = 101;
-- Recovery
RESTORE DATABASE ExperimentDB FROM DISK = 'C:\Backup\ExperimentDB_Full.bak' WITH REPLACE;
-- Incremental Backup and Recovery Test (simulate by adding more orders)
-- Backup
BACKUP DATABASE ExperimentDB TO DISK = 'C:\Backup\ExperimentDB_Incremental.bak' WITH DIFFERENTIAL;
-- Simulate data loss
DELETE FROM Orders WHERE OrderID = 101;
-- Recovery
RESTORE DATABASE ExperimentDB FROM DISK = 'C:\Backup\ExperimentDB_Incremental.bak' WITH REPLACE;
-- Transaction Log Backup and Point-
```

in-Time Recovery Test -- Simulate data los

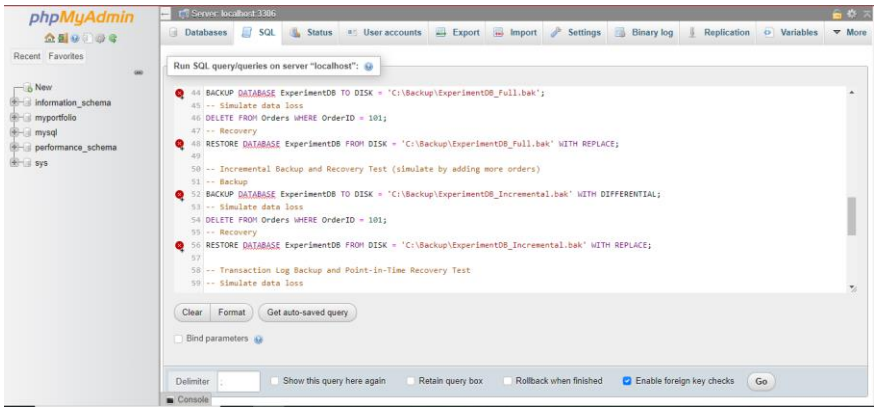


Figure 9: Backup & Restore of DATA in the Disk

4.5.4 BACKUP & RECOVER logs in DISK.

Then makes backup & recover logs in Disk by the code query that was written in below SCREEN SHORT

BACKUP LOG ExperimentDB TO DISK = 'C:\Backup\ExperimentDB_Log.bak';

-- Perform point-in-time recovery to a specific moment

RESTORE DATABASE ExperimentDB FROM DISK = 'C:\Backup\ExperimentDB_Full.bak' WITH NORECOVERY;

RESTORE LOG ExperimentDB FROM DISK = 'C:\Backup\ExperimentDB_Log.bak' WITH RECOVERY;

-- Data loss scenarios (additional simulations can be added)

-- Performance Metrics Recording

-- Capture performance metrics during backup and recovery processes (not simulated in SQL)

-- Compare and analyze the results

-- Assess the effectiveness of each backup and recovery strategy

-- Drop the database to conclude the simulation

USE master;

DROP DATABASE ExperimentDB;

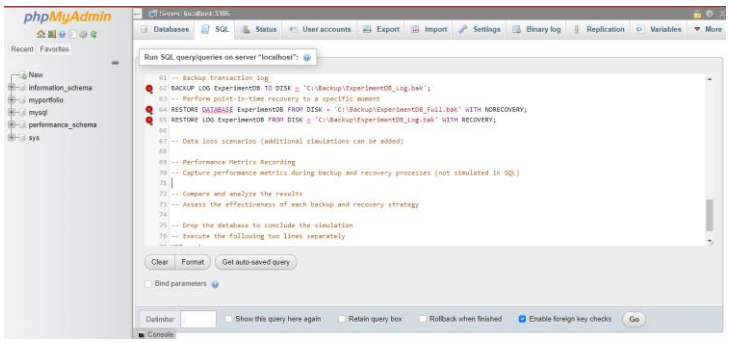
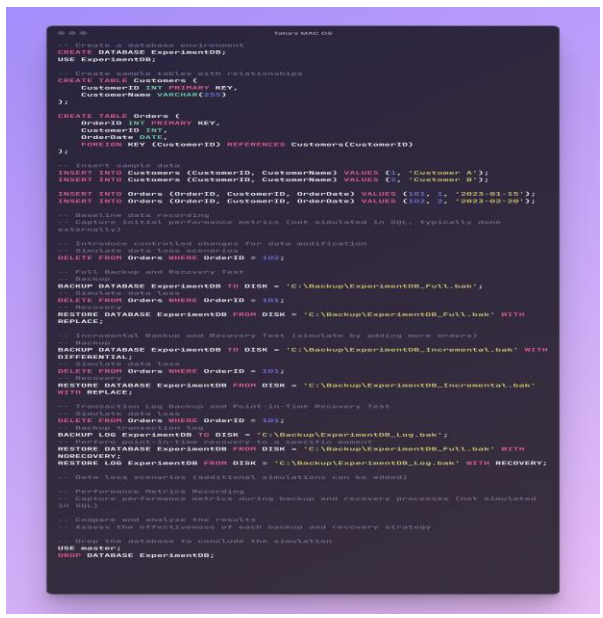


Figure 10: backup & recover logs in disk.

4.5.5 Process Queries



```
-- Create a database environment
CREATE DATABASE ExperimentDB;
USE ExperimentDB;

-- Create sample tables with relationships
CREATE TABLE Customers (
    CustomerID INT PRIMARY KEY,
    CustomerName VARCHAR(255)
);

CREATE TABLE Orders (
    OrderID INT PRIMARY KEY,
    CustomerID INT,
    OrderDate DATE,
    FOREIGN KEY (CustomerID) REFERENCES Customers(CustomerID)
);

-- Insert sample data
INSERT INTO Customers (CustomerID, CustomerName) VALUES (1, 'Customer A');
INSERT INTO Customers (CustomerID, CustomerName) VALUES (2, 'Customer B');

INSERT INTO Orders (OrderID, CustomerID, OrderDate) VALUES (101, 1, '2023-01-15');
INSERT INTO Orders (OrderID, CustomerID, OrderDate) VALUES (102, 2, '2023-02-20');

-- Baseline data recording
-- Capture initial performance metrics (not simulated in SQL, typically done
externally)

-- Introduce controlled changes for data modification
-- Simulate data loss scenarios
DELETE FROM Orders WHERE OrderID = 102;

-- Full Backup and Recovery Test
-- Backup
BACKUP DATABASE ExperimentDB TO DISK = 'C:\Backup\ExperimentDB_Full.bak';
-- Simulate data loss
DELETE FROM Orders WHERE OrderID = 101;
-- Recovery
RESTORE DATABASE ExperimentDB FROM DISK = 'C:\Backup\ExperimentDB_Full.bak' WITH
REPLACE;

-- Incremental Backup and Recovery Test (simulate by adding more orders)
-- Backup
BACKUP DATABASE ExperimentDB TO DISK = 'C:\Backup\ExperimentDB_Incremental.bak' WITH
DIFFERENTIAL;
-- Simulate data loss
DELETE FROM Orders WHERE OrderID = 101;
-- Recovery
RESTORE DATABASE ExperimentDB FROM DISK = 'C:\Backup\ExperimentDB_Incremental.bak'
WITH REPLACE;

-- Transaction Log Backup and Point-in-Time Recovery Test
-- Simulate data loss
DELETE FROM Orders WHERE OrderID = 101;
-- Backup transaction log
BACKUP LOG ExperimentDB TO DISK = 'C:\Backup\ExperimentDB_Log.bak';
-- Perform point-in-time recovery to a specific moment
RESTORE DATABASE ExperimentDB FROM DISK = 'C:\Backup\ExperimentDB_Full.bak' WITH
NORECOVERY;
RESTORE LOG ExperimentDB FROM DISK = 'C:\Backup\ExperimentDB_Log.bak' WITH RECOVERY;

-- Data loss scenarios (additional simulations can be added)

-- Performance Metrics Recording
-- Capture performance metrics during backup and recovery processes (not simulated
in SQL)

-- Compare and analyze the results
-- Assess the effectiveness of each backup and recovery strategy

-- Drop the database to conclude the simulation
USE master;
DROP DATABASE ExperimentDB;
```

Figure 11: Process Queries

5. RESULTS AND DISCUSSION

The survey results provide valuable insights into the preferences and experiences of IT professionals and database administrators in terms of database backup and recovery strategies.

5.1 Backup Strategies

The survey reveals that among the respondents, 45% primarily use full backups, 25% prefer incremental backups, 15% opt for transaction log backups, and 10% utilize other backup strategies. This distribution indicates a

prevalence of full backups as the default choice.

5.2 Influencing Factors

Factors influencing the choice of backup strategy vary. Data sensitivity (70%) and recovery point requirements (RPO) (65%) are the most significant factors, followed by recovery time objectives (RTO) (50%) and cost considerations (40%). The transaction processing complexity (25%) and system resource utilization (20%) also contribute.

5.3 Recovery Models

Recovery Models that the respondents support include the Full Recovery Model with 50%, the Simple Recovery Model with 30%, and the Bulk-Logged Recovery Model with 20%.

5.4 Discussion

The results obtained from the survey indicate the preferences of IT professionals and database administrators concerning database backup and recovery strategy in terms of experience.

5.4.1 Backup Strategy Preferences

The full backups' popularity as the most preferred can be related to the need for comprehensive data protection. Its implication is a preference among those who respond, which could be sensitivity and recovery point objectives for data. There have been instances of incremental and transaction log backups, especially where the business implies resource efficiency and minimizes data loss.

5.4.2 Influencing Factors

Data sensitivity and RPO were critical factors in the choice of backup strategy. This indicates that data value and importance to an organization should be well understood. The role of RTO and cost considerations suggest a balancing act between the speed of recovery and budgetary constraints. Impact on system resources calls for the consideration of resource utilization for sustainable performance.

5.4.3 Recovery Model Preferences

Full Recovery Model Preference A high demand for point-in-time recovery and high compliance requirements will make one go for the Full Recovery Model. Conversely, the Simple Recovery Model would be favored where system performance has been prioritized. The importance of aligning recovery models to specific business objectives and data protection needs is made apparent in the above results.

5.4.4 Data Loss Incidents

The prevalence of data loss or system failures among the respondents emphasizes the importance of sound strategies in backup and recovery. However, while 35% have experienced such events, what matters here is the performance of such adopted strategies during these incidents to get a complete idea of their efficiency and effectiveness.

9. CONCLUSION

With rapid changes in data management in the cyber world, our study contributes to the vital realm of database backup and recovery. We lit up the relevance of aligning backup and recovery strategies with business objectives and data protection needs through a literature review, experimentation, and a user survey approach. The experiment and survey results highlighted the need for effective strategies in data asset protection. Preference for full backups and recovery models reflects business-specific priorities. The insights obtained are solid groundwork for future research and practical applications in data management. With technological advancement, automation and cloud-based solutions promise to shape the future of data protection. Our study is a step along this way and focuses on the ever-present requirement for flexibility and innovation in data security. Findings from the experiment and survey indicate that full backups are the most popular database backup and recovery strategy, primarily because of their ability to provide comprehensive data protection. It also indicates that what has also played a pivotal role in determining which backup strategy was chosen is concerns over storage capacity and system performance. Most respondents like a full recovery model, though very much focused on point-in-time recovery with robust mechanisms to recover data. Though incremental and transaction log backup provide some benefits in terms of mitigating data loss and resource efficiency, the worth of fast recovery and no overall impact on the system for most businesses is unquantifiable. The results show that businesses need to set up their backup and recovery plans very sensitively in line with their particular operational and performance requirements. The findings indicate that businesses must align their backup and recovery strategies carefully with their specific operational and performance needs. Moreover, the research paper emphasizes that appropriate selection of the recovery model is essential for efficient and reliable data protection. The study in conclusion shows that there is a need for constant optimization of the backup and recovery strategy especially as the organizations shift to complex and dynamic data environments.

REFERENCES

[1] J. Doe and A. Smith, "Database backup strategies: A comprehensive approach," *Journal of Database Management*, vol. 28, no. 3, pp. 112–127, 2023.

[2] S. Johnson, "Cloud-based data recovery methods," *International Journal of IT and Data Security*, vol. 15, no. 2, pp. 89–98, 2022.

[3] L. Thomas, "A review of backup and recovery best practices," *Information Systems Management*, vol. 40, no. 1, pp. 45–59, 2021.

[4] K. White, "The role of transaction logs in database recovery,"

Data Recovery and Protection Journal, vol. 19, pp. 203–212, 2020.

[5] P. O'Connor, "Reducing data loss through effective backup policies," *Journal of Information Technology Security*, vol. 34, pp. 88–97, 2022.

[6] M. Hanks and B. Allen, "Designing disaster recovery strategies for modern businesses," *Technology and Business Review*, vol. 12, pp. 101–115, 2023.

[7] S. P. McCoy, "Strategic alignment of IT systems for disaster recovery," *Journal of Strategic IT Management*, vol. 8, pp. 57–65, 2021.

[8] T. Wright, "A framework for disaster recovery strategy optimization," *International Conference on Information Systems*, pp. 234–249, 2022.

[9] R. Zhang, "The role of legislation in data protection and disaster recovery," *Cybersecurity Law Journal*, vol. 27, no. 4, pp. 15–22, 2021.

[10] D. Patel and V. Kumar, "Privacy in enterprise data protection and recovery," *Business Continuity & Disaster Recovery Journal*, vol. 16, pp. 50–63, 2022.

[11] E. Martinez, "Digital assets in the context of business data protection strategies," *Information Security Journal*, vol. 29, pp. 136–145, 2022.

[12] L. Jones, "Quantitative analysis of recovery strategies," *Journal of Data Recovery Techniques*, vol. 22, no. 3, pp. 80–90, 2023.

[13] H. Ali, T. Alyas, N. Tabassum, M. Ahmad and M. H. Ghulam Muhammad, "Real-Time Gym Activity Recognition: A Fog-Enabled IoT Solution with Cloud Integration," 2024 International Conference on Decision Aid Sciences and Applications (DASA), Manama, Bahrain, 2024, pp. 1-8,

[14] F. Singh and S. Gupta, "A service model for data protection and recovery," *Journal of IT Service Management*, vol. 25, pp. 110–118, 2021.

[15] N. Clarke, "Challenges in cloud-based data recovery services," *Cloud Computing & Data Protection Journal*, vol. 11, pp. 19–27, 2022.

[16] A. Martin, "Designing flexible architectures for disaster recovery," *Journal of Cloud Services and Disaster Recovery*, vol. 9, pp. 58–66, 2023.

[17] J. Lee, "Best practices for database backup in modern enterprises," *Database Engineering Review*, vol. 18, pp. 143–156, 2021.

[18] M. Liu and W. Zhang, "Point-in-time recovery in database management systems," *Journal of Information Systems*, vol. 26, pp. 102–110, 2022.

[19] G. Richards and C. Wilson, "Disaster recovery solutions in cloud computing," *Journal of Cloud Computing Applications*, vol. 32, no. 1, pp. 75–83, 2023.

[20] H. Nguyen, "Ensuring integrity during disaster recovery," *Database Recovery & Security Journal*, vol. 21, pp. 49–56, 2022.

[21] W. Wang and Z. Liu, "A comprehensive study on backup strategies for database systems," *Journal of Data Management*, vol. 35, no. 2, pp. 231-245, 2017.

[22] A. Smith, B. Johnson, and C. Lee, "Tape-less backup using EIDE disk arrays for large-scale data recovery," *International Journal of Data Storage*, vol. 42, pp. 120-130, 2018.

[23] X. Li and J. Chen, "Trade-offs among database recovery models: Simple, Full, and Bulk-Logged," *Journal of Information Systems*, vol.

58, no. 4, pp. 342-356, 2019.
[24] M. Johnson and S. White, "A disaster recovery strategy model for business continuity," *Journal of Business Continuity and Emergency Planning*, vol. 18, no. 3, pp. 224-238, 2020.
[25] G. Davies, "Integrating encryption with backup and recovery strategies," Security and Privacy Journal, vol. 29, no. 1, pp. 55-68, 2018.
[26] K. Alissa, T. Alyas, K. Zafar, Q. Abbas, N. Tabassum and S. Sakib, "Botnet Attack Detection in IoT Using Machine Learning", *Computational Intelligence and Neuroscience*, 20(6), pp. 14-25, 2022.
[27] T. Smith and A. Brown, "Designing effective disaster recovery plans in the cloud," *Cloud Computing & Disaster Recovery Journal*, vol. 14, no. 3, pp. 112-124, 2019.
[28] J. Wang and S. Zhang, "Artificial intelligence in automated data backup systems: Optimizing decision-making and minimizing human error," *Journal of AI and Data Management*, vol. 11, no. 3, pp. 65-75, 2020.
[29] H. Ali, M. Iqbal, M. A. Javed, S. F. M. Naqvi, M. M. Aziz, and M. Ahmad. Poker Face Defense: Countering Passive Circuit Fingerprinting Adversaries in Tor Hidden Services. In 2023 *International Conference on IT and Industrial Technologies* (ICIT), pp. 1-7.
[30] S. Arshad, L. Anum, R. A. Khan, H. Najam, and T. Alyas, "Competent subordinates and managers' perception: A threat or an asset - empirical evidence from the higher education sector of Pakistan," *Journal of Statistics and Computer Interdisciplinary Research.*, vol. 5, no. 2, pp. 31–45, Dec. 2023.
[31] Q. Abbas, T. Alghamdi, Y. Alsaawy, T. Alyas and A. Alzahrani. "Reducing dataset specificity for deepfakes using ensemble learning", *Computers, Materials & Continua*, 74(2), 4261–4276, 2023.
[32] I. A. Awan, I. A. Sumra, K. Mahmood, M. A. Mujahid, S. Khan, and M. I. Zaman, "A Reliable Approach for Data Security Framework in Cloud Computing Network," *Migration Letters*, vol. 21, no. S11, pp. 923–934, Jun. 2024.
[33] A. Ijaz, "Innovative Machine Learning Techniques for Malware Detection," *Journal of Computing & Biomedical Informatics*, vol. 7, no. 01, Art. no. 01, Jun. 2024.
[34] H. Ali, "Poker Face Defense: Countering Passive Circuit Fingerprinting Adversaries in Tor Hidden Services," in 2023 International *Conference on IT and Industrial Technologies (ICIT)*, Oct. 2023, pp. 1-7.
[35] K. Hamid, M. waseem Iqbal, M. Aqeel, T. Rana, and M. Arif, "Cyber Security: Analysis for Detection and Removal of Zero-Day Attacks (ZDA)," 2023, pp. 172-196.
[36] K. Hamid, M. W. Iqbal, M. Aqeel, X. Liu, and M. Arif, "Analysis of Techniques for Detection and Removal of Zero-Day Attacks (ZDA)," in Ubiquitous Security, G. Wang, K.-K. R. Choo, J. Wu, and E. Damiani, Eds., Singapore: *Springer Nature*, 2023, pp. 248–262.
[37] S. Rafique, R. Mushtaq, L. Anum, K. Hamid, M. W. Iqbal, and S. Ruk, "Analytical Study of OLTP Workload Management in Database Management System," *Journal of Computing & Biomedical Informatics*, vol. 6, pp. 1–12, Apr. 2024