Security Challenges of Distributed Databases Systems – A Systematic Review of Related Literature

Ferdinand V. Dalisay
Pangasinan State University
fvdalisay@psu.edu.ph

Abstract – *This paper presents a theoretical study on the security challenges of the distributed database system. The paper reflects on the understanding of what are the different security issues and challenges in the past years. The literature review demonstrates the identified solutions to the specific security problems, as well as the lapses of the distributed database system, and examines the state of current research on the topic and points out gaps in the existing literature. The findings of the study revealed that a distributed database system encountered a lot of security issues in the past years, but this technology has been overcome most of the security challenges. However, there are still security issues that still persist in recent years and have yet to be addressed adequately which affect the performance of this technology. These security issues will keep potential researchers as well as distributed database system implementers occupied for a long time.*

Keywords: Security, Databases, distributed databases, Replication Control, Concurrency Control

1. INTRODUCTION

Distributed database technology has gradually replaced centralized databases in various domains over the last decade, including commerce, finance, and even an organization [1]. Thanks to promising database management functionality, a distributed database is now beneficial. It offers a "seamless" interface to data processed through various computer systems [2]. According to (S. Ceri), the technical features that motivate the success and growth of distributed databases have not to deal with their reduced cost but instead with their increased benefits. A distributed database is a collection of databases distributed and stored on multiple computers within a network [3]. Distributed database system functions include distributed query management, distributed transaction processing, distributed metadata management, and enforcing security and integrity across the multiple nodes.

The analysis of distributed database research problems has introduced a range of

"first-generation" commercial products. Distributed database technology is expected to impact data processing the same way that centralized systems did a decade ago. Stonebraker suggests that centralized database managers will be an "antique curiosity" within the next ten years, and most organizations will move toward distributed database managers [20]. Furthermore, with the advancement of distributed database systems, the protection of data stored and shared between sites has become a more pressing concern. An example of this is the recent advance in cloud computing, and distributed web applications have created the need to store a large amount of data in distributed databases that provide high availability and scalability.

Distributed Database Systems (DDBS) are central to many security threats and those present in a centralized database management system. In addition, the expansion of sufficient distributed database security has been complex by the relatively recent opening of the object-oriented

database model. This new replica cannot be ignored. It has been created to address the growing difficulty of the data stored in present database systems. This paper aims to carry out a systematic literature review of evidence about the different security challenges of a distributed database system. This review will focus on answering the following questions;

1) What are the security challenges of distributed database systems?

2) What are the solutions presented to address the security challenges of distributed database systems?

This paper will discuss the challenges and propose solutions that affect the distributed database's performance by reviewing the existing body of research on the topic.

2. METHODOLOGY
2.1 Literature search

In this literature review, the results of functional studies were published in electronic databases such as Google Scholar, Research Gate, and EBSCOhost. A variety and combination of keywords were used in the review, including *Challenges of Distributed Database Systems, Distributed Database Systems Application, Distributed Database Systems security, and Distributed Databases Systems Security Challenges.* In addition, after conducting preliminary searches in electronic databases, a set of criteria is established to select appropriate studies for inclusion in the study.

The paper selection process included 1) literature study evidence regarding DDBS security challenges, 2) published as articles or journals, 3) have an abstract, and 4) the research method is demonstrated clearly. A citation was omitted if 1) it was in academic settings such as book review, 2) it was a book or chapter, or 3) no theoretical data was reported. The authors, the year of publication, the study's goals, the settings, the procedure used (research design, data collection, and analysis techniques), and the key findings were all methodically collected. This method addresses all the questions and identifies solutions by this process of literature search.

2.2 Identifying the security challenges of distributed database systems

Based on journal papers, global surveys, current industry happenings, and showcase trends, the researcher gathered data from various writing sources. After collecting the literature, the researcher sorts it to see if it's essential to finalize the subject's representative literature. The importance of the literature is determined by its meaning, authority, effectiveness, and dependability.

There are two parts in organizing the process of determining the security challenges of DDBS. The articles are analyzed in the first part using the existence and prohibition criteria. The journals without practical evidence related to the security challenges of DDBS are eliminated in the review. The remaining journals with potential are obtained and will be examined further in the following section. In the second part, the reviewed journals' challenges are listed, and the overview of each issue is summarized.

2.3 The solutions presented on the security challenges of DDBS

Additional research into the problems and challenges listed has been completed. The researcher further analyzed the solutions in every security challenge of DDBS. For the security challenges, several solutions have been found. However, these solutions did not resolve all the difficulties in identifying the security challenges of DDBS.

In this part, all of the identified challenges and their corresponding answers are separated into two parts; 1) security challenges that already have a solution, and 2) remained challenges in security that anticipate the resolution. This section also includes further reference searches, in which the researcher investigates the relations between the originally discovered papers and the references made to those papers.

3. RESULTS AND DISCUSSION
3.1 Literature search

Table 1 shows the number of papers in each of the databases that were identified using the key search. The second column in table 1

contains all the results, including the non-scientific writings such as magazine articles. In all, there are 50 journals examined against the inclusion and exclusion criteria. Based on the reading of abstracts, 26 journals are excluded from the review, and the remaining 24 journals with potential are acquired. The remaining full texts journal articles are evaluated to check if it has theoretical evidence. Finally, 11 articles are excluded, and 13 qualified peer-reviewed research papers on security challenges in DDBS were left for review.

Table 1. Results from searched databases

3.2 The security challenges of DDBS

In their studies, the majority of the reviewed papers listed popular distributed database problems, and through analysis, a total of five security challenges that distributed databases have faced in recent years were identified. However, as technology evolves, application areas develop, and familiarity with distributed database technology grows, new security issues emerge.

Table 2 summarizes the recorded results relevant to the distributed database system's security issues.

Security Challenges	Paper
Distribution Design	[8] [6]
Operating Environment	[4] [5]
Replication Control	[4] [6] [7]
Distributed Deadlock Management	[4] [5]
Distributed Concurrency Control	[4] [6] [9]

The Distributed Concurrency Control is the integrity of the database is maintained by specifying the synchronization of access to the distributed database to manage concurrency; different locking techniques uses based on the mutual exclusion of access to data. The Replication Control applies to distributed systems where a database is supposed to be replicated if the whole database or its percentage is copied. The copies are stockpiled at different sites. Having more than one copy of a database, the issue is continuing the copies' communal uniformity, ensuring with all copies are identical schema and data.

Deadlock Handling if the users request the same resources from the database if the resources are obtainable, then the database allows permission for the resources to that user if not available. The user has to wait until the resources have released another user. Sometimes the users

Database	Total number of result	Peer review paper
Research Gate	46	3
OBSCOhost	1,256	4
Google Scholar	16,800	6

do not remove the resources blocked by some other users. To implement distributed database environment, a specific operating system is required as per organizational requirements. The operating system plays a vital role in managing the distributed database because sometimes it supports the distributed database. Furthermore, the central problem area in the distributed database where data is located in numerous locations and the number is used. So, the transparent management of information is essential to maintain the integrity of the distributed database.

3.3 Solutions presented to address the security challenges of DDBS

There have been several researches on the various security issues and difficulties in a distributed database system. Table 3 are listing the solutions to the DDBS's most common security problems.

Table 3. Solutions for the security challenges in DDBS

Solutions	Paper
Locking Techniques	[9] [4]

Fragmentation	[12] [8]
Deadlock Detection Agents	[4] [11]
Summary-schema Model	[10]

Abbas et al. highlight locking as a technique of synchronizing the access through concurrent transactions towards database items that resolve the distributed concurrency control issue [9]. Through fragmentation, a lot of topics in DDBS can be resolved. Fragmentation is when large schema records are divided into independent pieces or parts to increase the speed of distributed query processing [12]. This also enables us to achieve the distribution design and security issues in DDBS [8]. For deadlock handling issues, Krivokapic et al. presented the deadlock detection agents (DDAs). Deadlock resolution strategies determine which transaction(s) are to be aborted in order to resolve the deadlock [11]. Bright and Hurson introduce the summary-schema model, one of the possible solutions to powerful user interfaces, effective distribution of processing, and increased semantic content in distributed design. [7]

Almost all of the identified challenges in DDBS have already been solved; however, some issues and challenges persist in this recent years, such as the integration of distributed databases with distributed operating systems and the data replication control. The distributed DBMSs require modifications in how the distributed O.S. performs their traditional functions (e.g., task scheduling, naming, buffer management). In this context, efforts that include too much of the database functionality inside the operating system kernel or those that modify tightly-closed operating systems are likely to prove unsuccessful [8][4][5].

Besides, some of the operating systems are not supported by distributed databases. On the other side, replication of data is the method of sharing information to make sure data consistency between redundant resources, such as software or hardware components, to improve reliability, fault tolerance, or accessibility [13]. However, many issues affect the design of a replicated database system to maintain its

requirements [4][6][7]. These are the remaining security challenges in a distributed database system that persist in recent years, furthermore, new issues arise with the rapid changing of technology that may cause a distributed database system to unsolved these security issues.

4. CONCLUSION

In this paper, a discussion of the state-of-the-art in distributed database research is conducted. Specifically, the researcher (a) reviewed the different security challenges in implementing a distributed database system; and (b) discussed the identified solutions for the specific security issue and addressed the remaining challenges.

In this paper, the researcher figures out the common security problem in implementing the distributed database system, namely distributed concurrency and replication control, distributed query processing, security issues in implementing the various versions of operating systems, distribution design that led to the security and privacy concern, and issue in detecting deadlock. In addition, these security problems can be further aggravated by the changing nature of the technology on which distributed DBMSs are implemented.

The majority of the security challenges in distributed databases have already been solved. Some of the identified solutions resolve numerous challenges in DDBS because most of the challenges are relevant and somewhat connected to each other. For instance, fragmentation resolves a lot of common issues such as distributed query processing, and distribution design. As well as the summary-schema model which settles the distribution of processing and distributed multidatabase systems issues. This suggests that the distributed database is very flexible in resolving various challenges that make it even popular compared to the traditional database management. In some cases, the provided solution resolves only a particular security issue. In the context of distributed concurrency control, locking technique insights to resolve this issue. Besides, the deadlock handling can be controlled through the use of deadlock detection agents. However, the

replication control and the integration of distributed databases in various operating systems are the security challenges that still persist in recent years. The analysis suggests that these security issues are continuously developing due to the rapid changes of technology so that there is no absolute solution for these security challenges.

5. REFERENCE

[1] KITUTA, K., KANT, S. AND AGARWAL, R., 2019. A systematic review on distributed databases systems and their techniques. Journal of Theoretical and Applied Information Technology, 96(1), pp.236-266.

[2] STONEBRAKER, M., 1989. Future trends in database systems. IEEE Transactions on Knowledge & Data Engineering, (1), pp.33-44.

[3] Thuraisingham, B.: Multilevel Security Issues in Distributed Database Management Systems II. Journal of Computers and Security 10(9), 727–747 (1991)

[4] RANA, M.S., SOHEL, M.K. AND ARMAN, M.S., 2018. Distributed Database Problems Approaches and Solutions-A Study. In International Journal of Machine Learning and Computing (IJMLC).

[5] REDDY, P.K. AND BHALLA, S., 1993. Deadlock prevention in a distributed database system. ACM Sigmod Record, 22(3), pp.40-46.

[6] HIREMATH, D.S. AND KISHOR, S.B., 2016. Distributed database problem areas and approaches. IOSR Journal of Computer Engineering (IOSR-JCE), 2, pp.15-18.

[7] PUPEZESCU, V. AND RĂDESCU, R., 2016, June. The influence of data replication in the knowledge discovery in distributed databases process. In 2016 8th International Conference on Electronics, Computers and Artificial Intelligence (ECAI) (pp. 1-6). IEEE.

[8] OZSU, M.T. AND VALDURIEZ, P., 1991. Distributed database systems: where are we now? Computer, 24(8), pp.68-78.

[9] ABBAS, Q., SHAFIQ, H., AHMAD, I. AND THARANIDHARAN, S., 2016, January. Concurrency control in distributed database system. In 2016 International conference on computer communication and informatics (ICCCI) (pp. 1-4). IEEE

[10] HURSON, A.R. AND BRIGHT, M.W., 1991. Multidatabase systems: An advanced concept in handling distributed data. In Advances in computers (Vol. 32, pp. 149-200). Elsevier.

[11] KRIVOKAPIĆ, N., KEMPER, A. AND GUDES, E., 1999. Deadlock detection in distributed database systems: a new algorithm and a comparative performance analysis. The VLDB Journal, 8(2), pp.79-100.

[12] TARUN, S., 2019. Distributed Database Systems Design Challenges and Countermeasures – A Study. Journal of the Gujarat Research Society, 21(6), pp.875-886.

[13] TIWARI, S.K., SHARMA, A.K. AND SWAROOP, V., 2011. Issues in Replicated data for Distributed RealTime Database Systems. IJCSIT) International Journal of Computer Science and Information Technologies, 2(4), pp.1364-1371.