Available Online at www.ijcsmc.com

International Journal of Computer Science and Mobile Computing

A Monthly Journal of Computer Science and Information Technology

ISSN 2320–088X

IJCSMC, Vol. 3, Issue. 3, March 2014, pg.1067 – 1074

RESEARCH ARTICLE

TO ENHANCE THE SECURITY IN DBMS USING INTEGRATED CRYPTOGRAPHIC ALGORITHMS

[1]Davinderpal Kaur, [2]Kewal Krishn

[1]Lovely Professional University, Phagwara, Punjab, India

[2]Lovely Professional University, Phagwara, Punjab, India

[1] davy.dhillon02@yahoo.com, [2] kewal.krishna@lpu.co.in

***Abstract** - In distributed database the data is physically stored on two or more computer systems, so DDBMS allow to manages or organize the whole database as a single collection of data as result the individuals able to access data from any of the database system because the replication of the data among all the systems in the network. A distributed database may be partitioned and replicated in addition to being distributed across multiple sites. All of is not visible to the users. In this sense, the distributed database technology extends the concept of data independence, which is a middle conception of database management, to situation where data are distributed and replicated over a number of machines connected by a network. In this paper our main focus is to improve the security of the distributed database*

***Keywords** - Distributed database, Security, Data, Replication*

I. INTRODUCTION

Data as an abstract concept can be viewed as the lowest level of abstraction from which information and then knowledge are derived. To store various kind of data database is required. Database is a collection of information organized in such a way that a computer program can quickly select desired pieces of data. The database can be organized in three ways.

1) Field

2) Record

3) Files

A field is a single piece of information. Record is one complete set of fields. File is a collection of records. To access information from a database, database management system is needed. This is a collection of programs that enables person to enter, organize, and select data in a database.

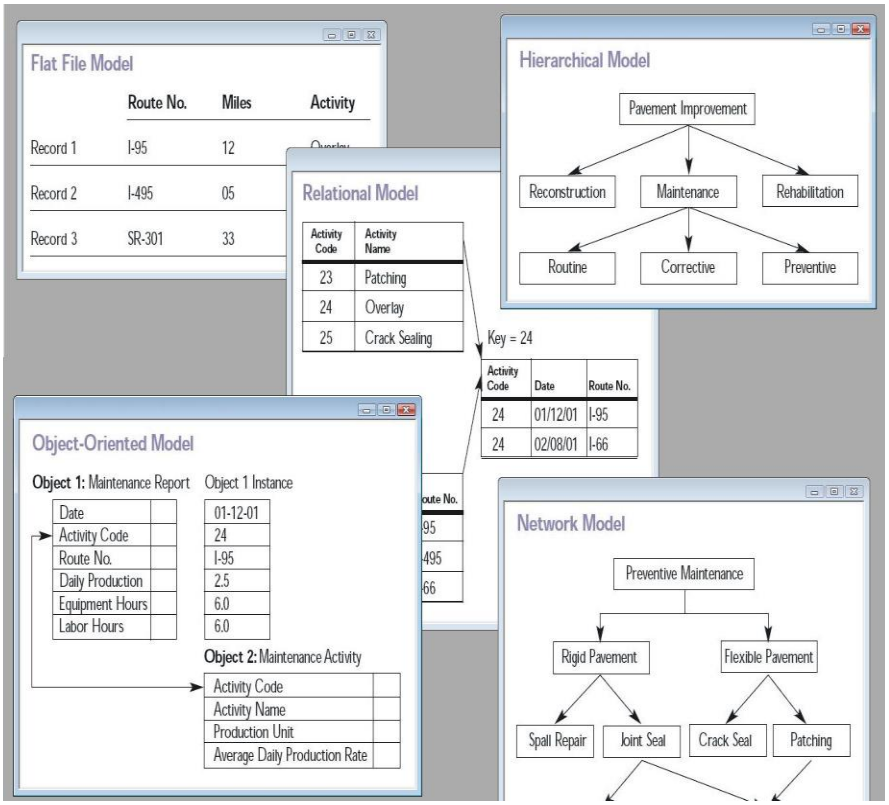

Fig 1: database of various models

A. Distributed Database

Distributed database is defines as a collection of multiple, logically interrelated database distributed on the computer network.[1]. Distributed database is the system in which data is not placed on the single site or location like centralized database, here the data is distributed and replicated over the different physical locations. Distributed database is the software database in which the relevant information is stored over the different database or storage

© 2014, IJCSMC All Rights Reserved

devices that are not directly attached to the same CPU. However, different users can able to access the required data from any of the database system connected by communication networks.

The software that creates and administers the distributed database and provides data to the users is called the Distributed Database Management system (DDBMS).It coordinates the access to the data at various nodes in the distributed network environment. A distributed database management system is a software system that permits the management of a distributed database and makes the distribution transparent to the users.

B. Strategies of Distributed database

Designing a distributed database is very difficult; many technical and organizational issues arise in designing of the distributed environment. There are number of strategies that include in distributed database is discussed below:

C. Data replication

The replication of the data provides the better access to the data from the different locations. User can access the data from any sites because the copy of the data is placed on multiple sites. Data availability is increased, if one site is down then the data is retrieved from another copy of the data. Reliability, fast data response, availability are the advantages of the data replication.

D. Fragmentation

In distributed database the global relation of the data can be split into several non overlapping portions which are called fragments. The splitting operation of global relation into different fragments can be performed by two different types of fragmentation. The sub transaction of the global relation of the data is distributed on the different locations in the computer network.

A. Types of distributed database

Distributed DBMS may be classified as Homogenous and Heterogeneous DDBMS. These are two main type of distributed database:

1) Homogeneous DDBMS

Same DBMS is used at each site of the distributed network. Each database can work independently and support all the functionality of the distributed DBMS. Simplicity, ease of designing, incremental growth is the advantages of Homogenous DDBMS. It is much easier to design a homogenous DDBMS as compared to heterogeneous system because at each site same database is used result easy communication among the databases over the network.

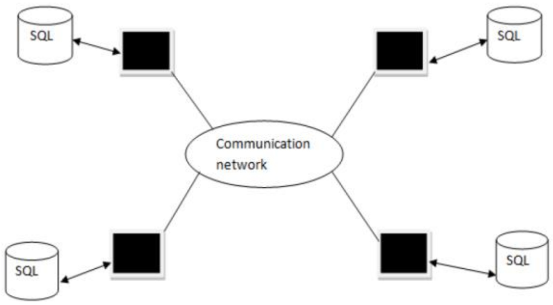

Fig 2: Homogenous distributed database

The figure 2 shows the homogenous structure of the database system where different nodes have the same database that's SQL server Database. The data from each node can be easily access by different nodes because of the homogenous environment of the database system.

2) Heterogeneous DDBMS

It refers to the distributed database system in which at least two different DBMS are used. In other words, we can say that not same database is used in the distributed systems. Each site may run on different DBMS. For one site may use SQL database or other may use ORACLE.

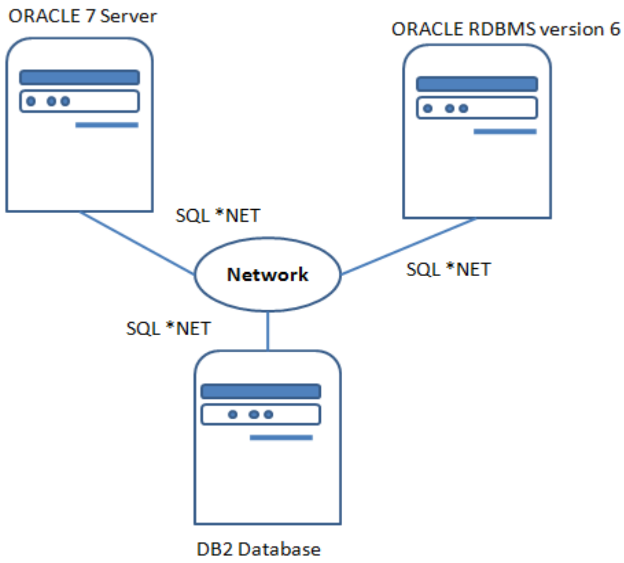

Fig 3: Heterogeneous Distributed systems

The figure 3 shows the heterogeneous structure of the database system where different nodes have the different database that's ORACLE 7 server Database, ORACLE RDBMS 6 version and non –Oracle RDBMS as shown in above figure. These are the different databases that can communicate with each other efficiently.

II. LITERATURE REVIEW

N.Batra et.al [2] survey many variants of concurrency control algorithms in database systems. They classify the different alternatives under locking, time-stamp, optimistic algorithms. The performance of different concurrency control algorithms have been explored extensively for database management systems. Proportional study among different variants of lock algorithm, time stamp ordering and optimistic concurrency control algorithms which have been implemented over last 20 years in distributed, mobile databases have been done based on various chosen parameters like reduced blocking, consistency, load balancing, efficiency, security etc. Depending upon application's requirement and resources available for a system, suitability of a particular variant to be used can be decided for a specific environment.

Lomet David et.al [3] described various concurrency control approaches are use in database that allows concurrent transaction access with high availability and performance. A Multi-versioning technique based on time stamping is used to provide the serializability of the transactions. This paper purpose the new concurrency control approach that enables all serializability to utilize multiple versions to increase the concurrency. The main idea is based on manage the range of time stamps for each transaction that deals with impact of conflicts.TCM timestamp range conflict manager can handle the conflicts occur in the transactions and also overcome the blocking problems faced in S2PL models. TCM is the new approach that provides the concurrent read/write access by multi-version that guarantees serializability.TCM also deals with the distributed transactions by handling both sub transactions and global transaction with timestamp range. Author described the TCM timestamp range conflict manager that provides read/write concurrent access while proving all SQL isolation levels, including serializability.
The different techniques that provide the concurrent access to the global as well as distributed transactions are described. Time stamping is very powerful approach of concurrency control through which various conflicts like RAW, WAW, WAR are to be resolved. And provide the serial execution of the transactions.

U. Aydonat et al. [4] purposed the use of a more relaxed concurrency control algorithm that provides better concurrency. It describe the 2 phase locking protocol used in centralized system that provide the good performance and fast transaction operations and work well with short transactions in which less conflicts occur in the concurrent transactions, but it limit the concurrency in applications with the long running transactions and with high contention .2 Phase Locking Protocol always abort the transaction every time a conflict occur. In this model conflicting access to the shared resources are not allowed. To overcome the limited access of the shared data, the author purposed the use of more relaxed algorithm that is based on the conflict serializability model (CS).This model transaction to complete even when the conflicting access to the shared data are present. Conflict serializability allow the transaction to complete when
they are aborted with the use of 2PL, that's results better performance and leading to more concurrency. The CS model ensures the consistency property hold for each transaction. This given proposal is implemented in software transaction memory STM system using various benchmarks that clearly shows that the 2PL model is less effective on applications with long running transactions and with high frequency of conflicts. Whereas the purposed CS model significantly reduces the abort rates and leads to better performance. The transactions with high abort rate are also provide better execution using the more relaxed model that is Conflict Serializability model and give better performance. Aydonet et.al defined the serializability of the transactions in the STM system ,it concern with enforcing the serializability by proving the CS model using the multi-versioning that further increase concurrency by allowing transactions to read order versions of shared data. It define an adaptive approach that the system switches between 2PL and CS .We can use the 2PL when the abort rate is low and switches to use CS when the abort rate is

high. This model is implemented and evaluated the performance using CS model and multi-versioning to enforce the serializability.

John Daniel et.al [5] defined the partitioning of database for improving the performance of distributed Online transaction processing (OLTP) database. Here the author compare the two low over head concurrency control schemes that allow data partitions to work on the other transactions when network stalls happens. Two schemes are defined first is simple blocking technique based on 2 phase locking and other is speculative approach. The comparison is done that shows that the blocking scheme works well with few partitions of the transactions and the speculative execution can provide better throughput in multi partition transactions. The purposed approach is implemented on TPC-C benchmark .and shows the result that the speculation concurrency increase the throughput as compare to the locking based system that handle the multi –partition transaction is to block until they complete

III. PROPOSED WORK

Today most research and business environment deals with the database. Database is simply the structured collection of data. It is field that concern with the real time applications like aircraft control, network management and stock marketing etc. For reliable and scalable accessing of the data distributed database come into exist that handle or access the data. Distributed database is the system in which data is not placed on the single site or location like centralized database, here the data is distributed and replicated over the different physical locations. When dealing with the distributed database several issues are to be considered. The main issue is the security of the distributed data systems. In the database several transactions can handle at the simultaneously. The goal of security is to prevent the inconsistency and interference among the transactions, so that it can be easily secure.

Discretionary access controls have the limitation that they can be easily circumvented by malicious users. TO illustrate, if the user has been granted the SELECT privilege without the grant option on a particular relation, then she should not be able to grant this privilege to other users. He can easily subvert this intent by making a copy of the relation in question, and as owner can give SELECT privilege on the copy to others. Even if users are trusted not to violate the security in this way, a malicious user can imbed Trojan horses that can do so. Multilevel secure databases enforce mandatory access controls to help eliminate these problems.

In a multilevel system, access controls are based on the security labels associated with each user and each data item. Security labels have two components: a hierarchical component and a set of categories or compartments which could be empty. The security policy requires that a user may have access to a data item if the security label of the user dominates that of the data item. Thus, a user with a SECRET clearance can have access to a relation with CONFIDENTIAL data, but not a relation with TOP SECRET data.

IV. METHODOLOGY

To maintain the security of distributed database systems, we can use the many techniques. We firstly use the relevant concepts underlying the notion of database security and summarize the most well-known techniques for the data base security. We then discuss current challenges for database security and some preliminary approaches that address some of these challenges. In our new algorithm we are going to integrate enhanced RSA and elgamal. As shown in figure we are using enhanced RSA and in elgamal we were use elgamal to Send Phi(n) from alice to bob. It will provide us more security in DBMS.

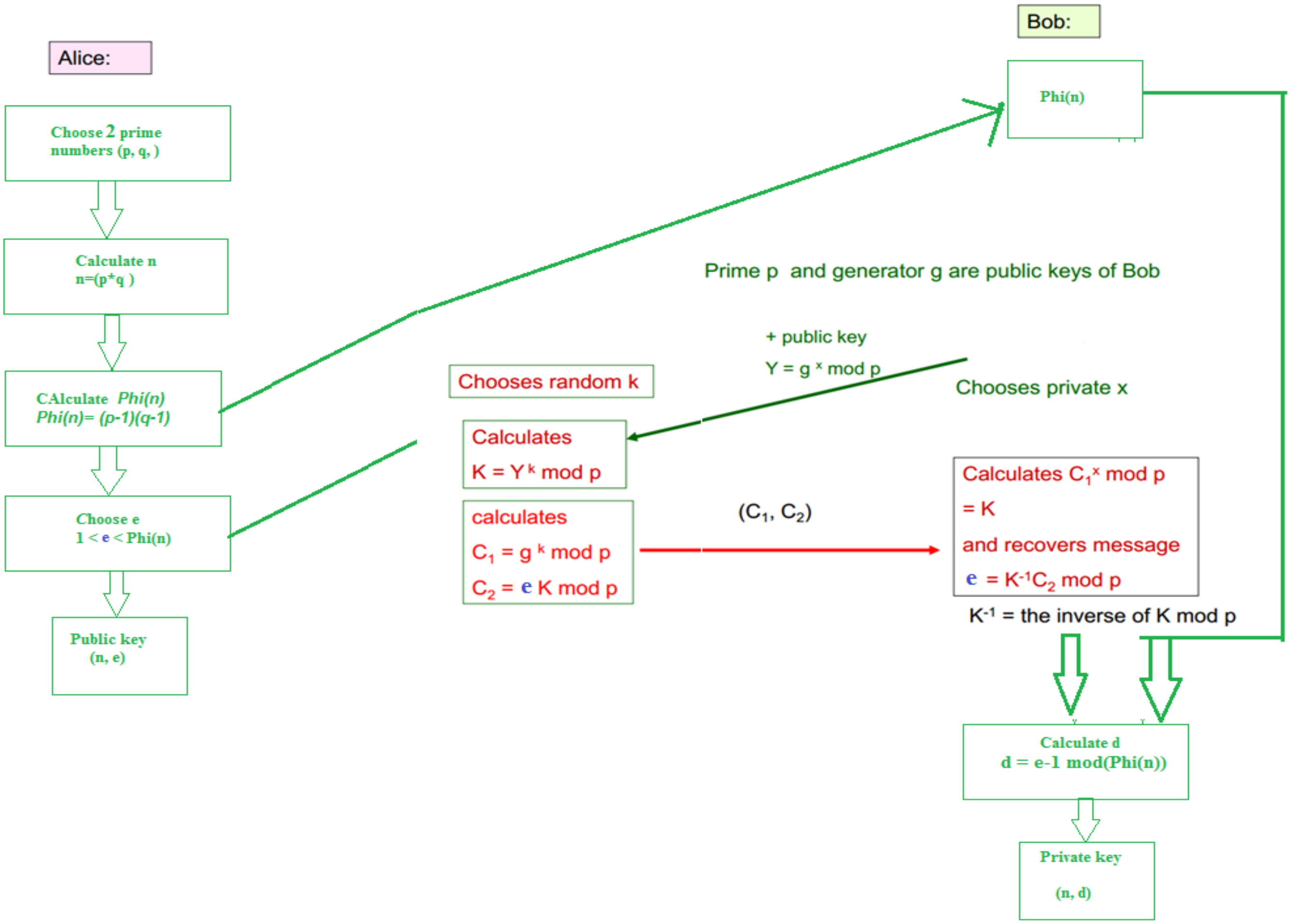

Fig 4: Proposed algorithm

Here in our novel approach we are going to integrate enhanced RSA and elgamal algorithm.
steps of the integrated algorithm are as following:

- User choose two large prime numbers p, q.
- Calculate n, where n = p*q.
- Also, compute the value of Phi (n) = (p-1)(q-1).
- Choose an integer e between 1 and Phi (n) such that gcd (e, Phi (n)) = 1. And use elgamal here to send value of e to receiver.
- Compute d whereby d = e-1 mod(Phi(n)).The public key is (n , e) whereas the private key is (n , d).

© 2014, IJCSMC All Rights Reserved

REFERENCES

[1]. Stefano Ceri and Giuseppe Pelagatta *Distributed Databases Principles & Systems*" 2008.
[2]. Neera Batra, and A. K. Kapil. "*Concurrency Control Algorithms and its Variants: A Survey*." In AIP Conference Proceedings, Vol. 1324, pp. 46-50, 2010.
[3]. Lomet, David, Alan Fekete, Rui Wang, and Peter Ward. "*Multi-Version Concurrency via Timestamp Range Conflict Management*." In *Data Engineering (ICDE), 2012* IEEE 28th International Conference on, pp. 714-725. IEEE, 2012.
[4]. Utku Aydonat and Tarek S.abdelrahman "*Relaxed concurrency control in software transaction memory*" IEEE Transactions on Parallel and DistributedSystems, VOL. 23, No. 7, pp 1312-1325,July 2012.
[5]. Jones, Evan PC, Daniel J. Abadi, and Samuel Madden. "*Low overhead concurrency control for partitioned main memory databases*." In Proceedings of the 2010 international conference on Management of data, pp. 603-614. ACM, 2010

.