Introduction advantages of fragmentation. the degree of concurrency

IntroductionDistributeddatabase is defined as a database which can be stored on multiple computers. Today, in the age of advances in Information Technology, there is important for peopleto gain access of the updated information on hand.

Users have the opportunityto gain access of the information at anywhere and anytime in the network byusing the distributed database. The security issues and concurrency control indistributed database are discussed in paper 1. TheoryBasedon paper 1, the communication and data processing have been improved by usingthe distributed database system. This is because the data on distributeddatabase is spread throughout different computer network sites. Not only forincreasing the speed of data access, but it also provides a local control ofdata for users and a single-point of failure is much less probably to arise. Distributed database is a database which is spread across multiple computersthat are connected via the data communication links. The advantage ofdistributed database is data is distributed, so that the network traffic can bereduced. Moreover, if the network of the company is temporarily broken, thelocal database does not affected and it will remain the works.

Due to the distributeddatabase is stored in multiple computers, so that the work of one branch willnot be affected when there is problems exist in other branch. However, toensure that the information and indexes are not altered will become moredifficult. Besides that, it is not well-organized when there is heavyinteractions occur between sites.

LiteraturesreviewFragmentation, replication and data allocation are presented as the design of distributeddatabase. According to the research of Shin and Irani 2, fragmentation isdefined as a design method which is used to divide the relation into two ormore partitions. Parallelism is one of the advantages of fragmentation. Thedegree of concurrency and parallelism is increased due to the transaction canbe divided into several sub queries by using the fragmentation. However, theoverall performance and integrity control will become slow and difficult tocontrol due to the data are stored at different sites. Fragmentation is dividedinto three types which is horizontal, vertical and hybrid fragmentation. Coroleland Morris described that data replication is refer to the storage of datacopies at different locations and different sites that served by a computernetwork 3.

Maintain the stability of data is becoming the main problem inmanaging the replicated data. There are several advantages of replication whichincluded improved response time, reduced the network traffic and also increasedthe reliability and availability. The process of deciding where to locate the datais known as the data allocation and the algorithm is considered into severalfactors such as performance and data availability goals 4. Data allocationstrategies are classified into centralized data allocation, partitioned dataallocation and replicated data allocation. Result outcomeAccordingto paper 1, it explained a lot of details about the concurrency control andsecurity in distributed database. Concurrency control in distributed databaseis defined as the action of processing the concurrent access to the database.

Distributedtwo-phase locking (2PL) is the most familiar distributed concurrency controlsystem. " read any, write all" is the main approach of 2PL protocol and it isused as the basic concurrency control protocol 5. Each transaction in 2PL haveexecuted in two phase which is growing phase and shrinking phase. Growing phaseis for obtains locks in transaction, while shrinking phase is for releaseslocks. Lock managers in 2PL are spread to all sites and each of them isresponsible to lock the data at that site. Distributed Optimistic protocol is anotherprotocol for concurrency control. It is operated through exchange the certificationinformation. Security is important in distributed database.

It is used to preventthe information and data modified or misused by other people. In this paper, there are four security components is presented which is securityauthentication, authorization, encryption and also access control. Moreover, deadlock is clarified as the major problem that occurs in distributed system. In this research, 2PL algorithm with Timestamps mechanism is found that it iseffectively enough for concurrency control in distributed database. ConclusionInpaper 1, it is presented about the design, concurrency control and securityof distributed database. Security is one of the most important things indistributed database as it is required to ensure that the information and datais operating in a secure environment and integrity.

Nowadays, distributeddatabase is becoming famous in computer science. Hence, we need to understandit and try to find out the solutions to improve the weakness of the distributeddatabase.