Database Management System – cs422 DE

Assignment 10 – Week 14

This assignment is based on lecture 12 (chapter 24 – Distributed Databases)

○ Submit your *own work* on time. No credit will be given if the assignment is submitted after the due date.
○ Note that the completed assignment should be submitted in .doc, .docx, .rtf or .pdf format only.
○ In MCQs, if you think that your answer needs more explanation to get credit then please write it down.
○ You are encouraged to discuss these questions in the Sakai forum.

---

(1) Compare and contrast a DDBMS with distributed processing. Under what circumstances would you choose a DDBMS over distributed processing?

Distributed Database Management System is mainly to make the distribution transparent to the user. The main aim of the transparency of the distributed system is to make appear like of centralized system, unknown to users. But if the data is centralized but users may access the data over the network which is called distributed processing.

When the data ware house is in different locations then it is better to have distributed DBMS in different locations. Data can be kept close to user who can locally access data and create policies of using data.

(2) Compare and contrast a DDBMS with a parallel DBMS. Under what circumstances would you choose a DDBMS over a parallel DBMS?

Distributed Database Management System is mainly to make the distribution transparent to the user. The main aim of the transparency of the distributed system is to make appear like of centralized system, unknown to users. The Single Processor system can no longer meet the growing requirements. There is need of multi-processor which is parallel DBMS. In order to provide multi-processor with common access to a single database, a parallel DBMS must provide shared resource management.

(3) Discuss the advantages and disadvantages of a DDBMS.

Advantage: Reflects organizational structure: If the company is distributed in many locations there is important to have DDBMS.
Improved availability: In centralized DBMS if one system is failure it goes with large loss, But in DDBMS we can have restored of data and data secure.
Improved reliability: As data may be replicated so that it exists at more than one site, the failure of a node or a communication link does not necessarily make the data inaccessible.

Disadvantage:
Complexity: It is very complex to maintain and adopt DDBMS. If the software does not handle data replication it reduces the reliability, availability.
Costly: As compared to Centralized database it is more costly
Complexity of design of Database.

(4) What is the difference between a homogeneous and heterogeneous DDBMS? Under what circumstances would such systems generally arise?

Homogenous DDBMS all site use the same product and they are easy to design and manage. while in Heterogonous system it may run the different DBMS products. and translations are required to allow communication between different DBMSs. To provide DBMS transparency, users must be able to make requests in the language of the DBMS at their local site.

(5) What functionality do you expect in a distributed DBMS?

As compared to centralized DBMS, distributed DBMS must be in advances form and able to provide the data if one system fails. Replication of data should be better so that all systems are in Consistent state. The availability of the data must be insured in any case (failure of a system).

(6) One problem area with DDBMSs is that of distributed database design. Discuss the issues that have to be addressed with distributed database design. Discuss how these issues apply to the global system catalog.

DDBMS is better approach for handling the large volume of data in better way. The processing of the data can be fast and reliable. But the design of such system is complex. The software used in such system but be fully compatible to handle all the processing, resiliency, and perform with large volume of data replication. The DDBMS must be designed in order to provide data to user if one system fails also, and also distributed nature should be hidden to users and should perform like of centralized system.

(7) What are the strategic objectives for the definition and allocation of fragments?

DDBMS consist of single logical database that are split into fragments. One Each fragment is stored in one or more computers under the control of separate database. The computers are connected by the network. Fragments may be replicated, and they are allocated to sites. The links are connected by networks.

(8) Describe alternative schemes for fragmenting a global relation. State how you would check for correctness to ensure that the database does not undergo semantic change during fragmentation.

- Horizontal Fragmentation: partition the relation in tuple level
- Vertical Fragmentation: partition the relation in attribute level
- Derived Horizontal Fragmentation: selection with matched attribute values from two relations

Mixed Fragmentation: Apply VF and HF on same relation.

Checking the correctness:

- Completeness: If a relation R is decomposed into fragments each data item that can be found in R can also be found in one or more of Ri's
- Reconstruction: The re-construction ability of the relation from its fragments ensures that constraints defined on the data in the form of dependencies are preserved
- Disjoint. If a relation R is horizontally decomposed into fragments, data item di is in Rj, it is not in any other fragment.

(9) What layers of transparency should be provided with a DDBMS? Give examples to illustrate your answer. Justify your answer.

As the database systems are scattered in distributed DBMS but the user should be transparent about this. The system should work as if the centralized DBMS. The distributed system should be hidden or should not be transparent to the user. The user should feel of using centralized system, should be reliable, data replication should be done properly and consistency should be maintained so that if one system fails also another should provide the same date , this how the transparency should be hidden.

(10)A DDBMS must ensure that no two sites create a database object with the same name. One solution to this problem is to create a central name server. What are the disadvantages with this approach?Propose an alternative approach that overcomes these disadvantages.

Disadvantages of central name server are

-loss of some local autonomy;

-performance problems, if the central site becomes a bottleneck;

-low availability; if the central site fails, the remaining sites cannot create any new database objects.

The alternative solution can be is to prefix an object with the identifier of the site that created it.