NAME OF STUDENT: .....

Student Number: | | | | | | | | | | Group : |

PLEASE CROSS (X) Lecturer:

Dan Hunter	
Amlan Mukherjee	
Ephias Rhode	

SUBJECT		TIME
INFORMATION SYSTEMS III MODULE B **FULL TIME /PART TIME**	**DATE** 11th June 2012	10:00 AM **DURATION** 120 MINUTES

Cape Peninsula University of Technology

FACULTY OF INFORMATICS AND DESIGN

COURSE: INFORMATION TECHNOLOGY

EXAMINER : MR A MUKHERJEE **MODERATOR** : MR D HUNTER

SPECIAL INSTRUCTIONS:

Answer all the questions.

All question papers must be handed in.

Write answers on space provided in question paper.

Credit will be given to answers which are **brief** and **to the point.**

1.1 REQUIREMENTS : None.

QUESTION 1 (/ 25)	
QUESTION 2 (/ 8)	
QUESTION 3 (/ 12)	
QUESTION 4 (/ 13)	
QUESTION 5 (/ 15)	
QUESTION 6 (/ 12)	
TOTAL (/ 85)	

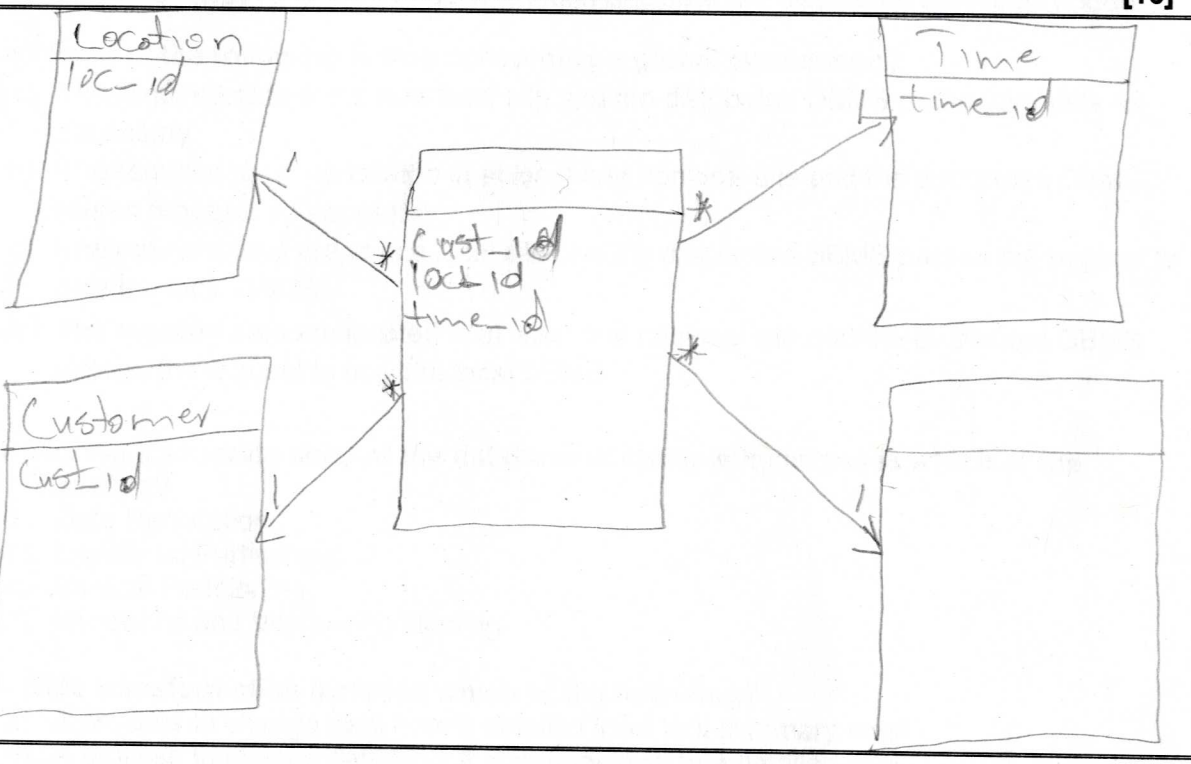
% =

Question 1 [25] As Question 3

1.1> Draw a detailed Star Schema for the following Data Warehouse/Mart situation:
Our large Car Hire Company is keeping track of the most popular cars for hire i.e. **hiring** of various Models of cars by Customers in the country. We need to keep track of daily quantity of cars hired for each Model that is available, for each individual Location (or car hire HUB), for each Customer. A typical query might be:
Find total number of days that a VW Polo Classic TDI 2011or 2012 was hired, for all our locations in Gauteng, for each month of 2011. [Do not give the query just the Star Schema]. **[15]**

1.2> If you needed to query your Car Hire Data Warehouse (which you created as a star schema) for yearly totals frequently, monthly totals often, and daily totals sometimes, explain how your fact table(s) would be structured and the relevant dimension table(s)? You should use an ERD drawing to show the structure. Don't draw a whole Star Schema but all relevant entities and relations. **[10]**

1.1>



1.2>

Question 2 [8]
Please tick the correct option given:

2.1> A distributed database has which of the following advantages over a centralized database?
a) Software cost
b) Software complexity
c) Slow response
d) Modular growth

2.2> A heterogeneous distributed database is which of the following?
a) The same DBMS is used at each location and data are not distributed across all nodes.
b) The same DBMS is used at each location and data are distributed across all nodes.
c) A different DBMS is used at each location and data are not distributed across all nodes.
d) A different DBMS is used at each location and data are distributed across all nodes.

2.3> **Which of the following is true concerning a global transaction?**
a) The required data are at one local site and the distributed DBMS routes requests as necessary.
b) The required data are located in at least one nonlocal site and the distributed DBMS routes requests as necessary.
c) The required data are at one local site and the distributed DBMS passes the request to only the local DBMS.
d) The required data are located in at least one nonlocal site and the distributed DBMS passes the request to only the local DBMS.

2.4> Storing a separate copy of the database at multiple locations is which of the following?
a) Data Replication
b) Horizontal Partitioning
c) Vertical Partitioning
d) Horizontal and Vertical Partitioning

2.5> Data transformation includes which of the following?
a) A process to change data from a detailed level to a summary level
b) A process to change data from a summary level to a detailed level
c) Joining data from one source into various sources of data
d) Separating data from one source into various sources of data

2.6> An operational system is which of the following?
a) A system that is used to run the business in real time and is based on historical data.
b) A system that is used to run the business in real time and is based on current data.
c) A system that is used to support decision making and is based on current data.
d) A system that is used to support decision making and is based on historical data

2.7> Fact tables are normally which of the following?
a) Completely denormalized
b) Partially denormalized
c) Completely normalized
d) Partially normalized

2.8> A star schema has what type of relationship between a dimension and fact table?

a) Many-to-many
b) One-to-one
c) One-to-many
d) All of the above.

Question 3 [12] A5 Question 1

Please choose the correct option provided and fill in the blank with the corresponding alphabet.

3.1 To speed up access in MOLAP data cubes are stored in memory which is called a G.

3.2 _____ is a data modelling technique used to map multidimensional decision support data into a relational database.

3.3 F are numeric measurements that represent a specific business aspect or activity.

3.4 A small, single-subject data warehouse subset that provides decision support for a small group of people is called a D.

3.5 A _____ is an arrangement of computerized tools to assist managerial decision making.

3.6 OLAP uses the data warehouse and the _____.

3.7 When a particular data item has more than one copy in a distributed database it is called _____.

3.8 _____ is the protocol to recover data in case of a failure in a distributed environment.

3.9 _____ is the lowest level of transparency in a distributed database.

3.10 The coordination of simultaneous execution of transactions in a multiuser database is known as I.

3.11 _____ is a disadvantage of 2-Phase locking protocol.

3.12 _____ can be considered as an optimized approach for locking in a distributed transaction.

A. Operational Database	B. Location Transparency	C. Fragmentation Transparency	D. 3.4 Data Mart	E. Majority Locking
F. 3.3 Facts	G. 3.1 Cube Cache	H. Cascading Rollback	I. 3.10 Concurrency Control	J. DSS
K. OLAP	L. 2Phase Commit Protocol	M. Star Schema	N. Primary copy locking	O. 2 Phase locking Protocol
P. Replication	Q. OLTP	R. Local Mapping Transparency	S. Fragmentation	T. Lock Granularity

Question 4 [13] As Question 2

Please read the following statements carefully and critically comment on them. You may agree or disagree with the statements. Please furnish logic supporting your decision.

4.1> " MOLAP is more dynamic compared to a ROLAP"

Wrong

[3]

4.2> "Fragmentation Transparency implies extensive use of the Distributed Data Catalog (DDC)"

Correct
DBMS has to know about how fragmentation is done and where the fragments are.

[3]

4.3> " Locking can be a problem in a distributed database"

Since the data can be replicated and scattered in different server situated in different physical location it can create a problem to locks each of them.

[4]

4.4> " A DSS should have support for VLDB and Client – Server Architecture"

True
Data warehouse can grow up to 300 terabyes as well.

[3]

Question 5 [15] As Question 4

5.1> You are the designer of a distributed system. Unfortunately, some of your team members do not understand when to implement a distributed database and when to implement a distributed processing system. Explain in brief the difference between these two to your team members and inform them in what different situations you would implement them.

In a distributed processing a database's logical processing is shore among two or more physically independent sites that are connected through a network. For example, the data input/output, data selection, and data validation might be performed on one computer.
A distributed database, on the other hand, stores a logically related database over two or more physically independent sites. The sites are connected via a computer network.
Distributed processing does not necessarily require a distributed database, but a distributed database requires distributed processing [4]

5.2> Discuss in brief cascading rollback and suggest a solution.

[2+2 =4]

5.3> In a 2PC what if participant crashes during the request phase before writing anything to log?

[2]

5.4> What if the coordinator crashes during Phase 2 (before sending the decision) and does not wake up in 2PC? Give any two possible solutions.

[1+4 =5]

Question 6 [12] As Question

6.1> Explain why you would recommend to the data modeller to use a star schema in the design of a data warehouse.

[6]

6.2> Write down three practical problems and challenges you are likely to face when designing a data warehouse? Please note that your data sources are both the transactional database and diverse external sources.

1.

2.

3.

[2+2+2 = 6]

********** END OF TEST **********