



Undergraduate Programme
Department of Computer Science and Information Technology
Semester 1, Year 2017

FINAL COURSE EXAMINATION PAPER

ITC251: DATA WAREHOUSING

Important Instructions to Students:

1. Read the questions, instructions carefully and answer the questions.
2. You are allowed an **extra 10 minutes reading time** (during this time, you are not allowed to write with pen) and **180 minutes (3 Hours) to complete this paper**.
3. Silent, non-programmable calculators are allowed, but not supplied.
4. Write all your answers in the answer booklet provided. If you need extra sheets, please ask from the supervisor. Write your name, course code and student identification number in the answer booklet and in the attendance sheet.
5. **This Final Course Examination Paper is worth 100 marks contributes 50% of the total course assessments.**
6. This paper contains one (1) section A. All questions are compulsory to answer.
7. Write all your answers legibly. The marker has the right to award zero for answers that are not legible.
8. This is a **Closed Book Examination**. The University views plagiarism, cheating and dishonest practice as serious offences and if found, offenders are penalized.

Examination Paper Content (Format):

<u>Section</u>	<u>Question</u>	<u>Content</u>	<u>Marks</u>	<u>Suggested Time</u>
A	1-35	Short Answer Questions	100	180 minutes
		TOTAL	100	180 Minutes

SECTION A:**100 MARKS***All questions in this section are compulsory*

1. What do you mean by enterprise data warehouse?
[2 Marks]
2. Why do enterprises really need data warehouses?
[2 Marks]
3. What do we mean by strategic information? For a commercial bank, name four (4) types of strategic business objectives.
[3 Marks]
4. Why does information gap exist in all organisations?
[2 Marks]
5. List and explain the following defining features of a data warehouse. Use examples to explain your answer.
A. separate,
B. subject-oriented,
C. integrated,
D. time-variant, and
E. nonvolatile.
[10 Marks]
6. List and briefly explain the following report and applications:
A. Ad hoc Reports.
B. Special Extract Programs.
C. Small applications.
[3 Marks]
7. Compare and contrast between operational and informational systems based on the following factors:
A. Data Content
B. Data Structure
C. Access Frequency
D. Access Type
E. Usage
F. Response Time
G. Users
[3.5 Marks]

8. Explain data granularity and how it is applicable to the data warehouse.

[2 Marks]

9. Explain why a data warehouse is an environment and not a product

[2 Marks]

10. List and explain five (5) differences between a data warehouse and a data mart.

[2.5 Marks]

11. Explain the top-down approach to build a data warehouse. Discuss four (4) advantages and four (4) disadvantages of using this approach.

[5 Marks]

12. What are the essential differences between defining requirements for operational systems and for data warehouses?

[2 Marks]

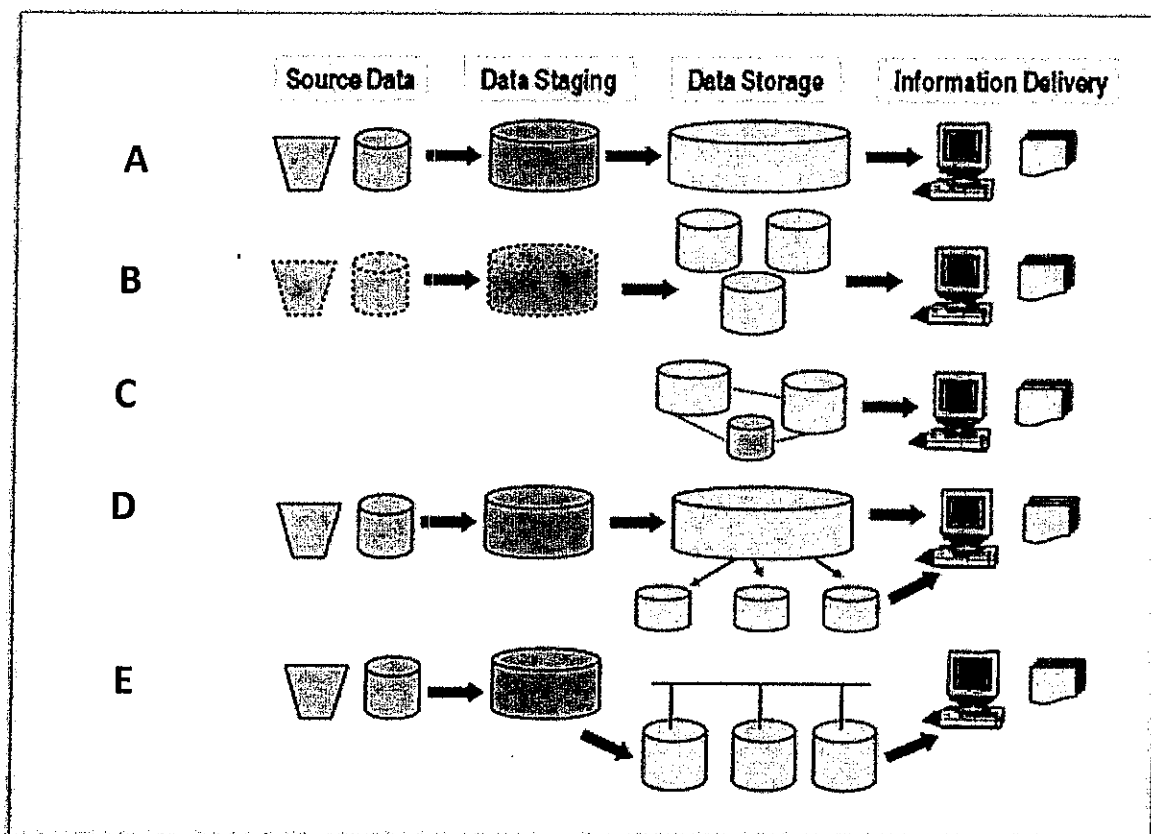


Figure 1: Data Warehouse architectural types
[Source: Rainardi, V.(2008)]

13. Figure 1 shows the five (5) different types of data warehouse architectures labelled by letter A -E.
Identify the five (5) architectures and label each letter by its corresponding name of architecture.
[2.5 Marks]
14. What are the essential differences between defining requirements for operational systems and for data warehouses?
[2 Marks]
15. Source data coming into the data warehouse may be grouped into the following four (4) categories.
A. Production Data
B. Internal Data
C. Archived Data
D. External Data
Explain each category with the help of an example.
[4 Marks]
16. Why do you need a separate place or component to perform the data preparation?
[2 Marks]
17. When we implement an operational system, we are likely to pick up data from different sources, move the data into the new operational system database, and run data conversions. Why can't this method work for a data warehouse?
[2 Marks]
18. How is data transformation for a data warehouse more involved than for an operational system?
Explain at least four (4) factors.
[2 Marks]
19. List and explain three (3) advantages of using real-time data warehousing.
[3 Marks]
20. Explain the difference between structured, semi-structured and unstructured data with the use of an example.
[3 Marks]
21. List the five (5) steps for gathering functional and non-functional requirements.
[2.5 Marks]
22. Why is it important to verify risks and make sure they are migrated when conducting the data feasibility study?
[2 Marks]

23. List and explain the three (3) types of SCD with the help of an example.

[3 Marks]

24. List and explain the three (3) types of SCD. How will the following record change with each type of SCD.

Table 1: A record

Customer Key	Name	City
1001	Tina	Lautoka

[3 Marks]

25. Explain the following two (2) types of SQL Server licensing models:

E. Per-processor license:

F. Server + CAL license:

[2 Marks]

26. Explain the three (3) fundamental principles to consider when performing data extraction.

[3 Marks]

27. What are the four (4) main methods for extracting data relational databases

[2 Marks]

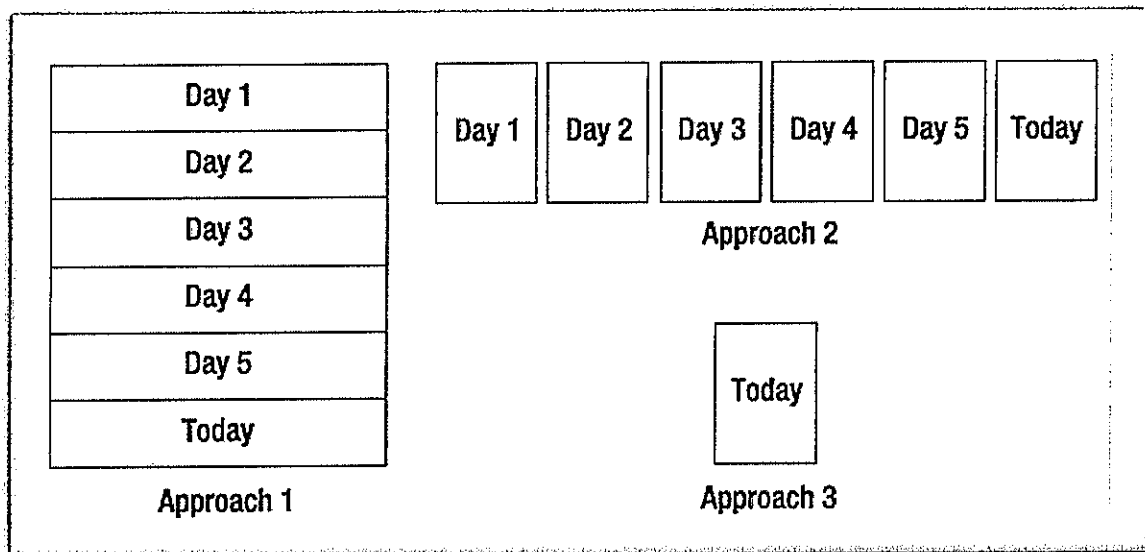


Figure 2: Different approaches for stage loading

[Source: Rainardi, V. (2008)]

28. Figure 2 illustrates the three different approaches of how stage tables are structured. Explain these three (3) approaches.

[3 Marks]

29. What are data firewalls? How are they similar to network firewalls? How are they different from network firewalls?

[3 Marks]

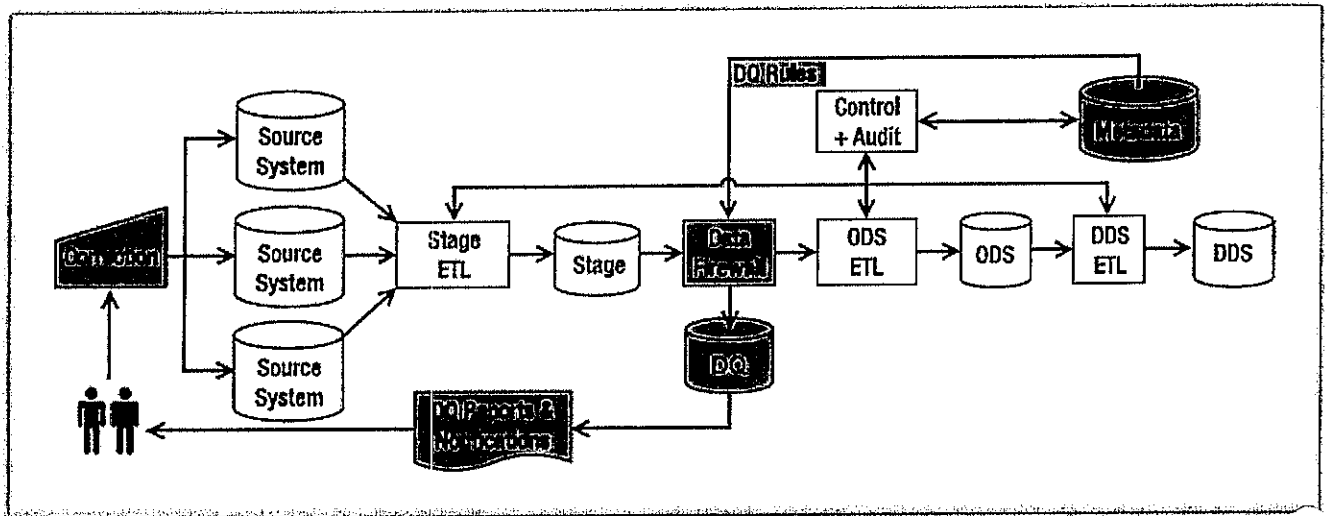


Figure 3: Data quality components in the data warehouse architecture

[Source: Rainardi, V. (2008)]

30. Figure 4 shows the data quality components in the data warehouse architecture. Provide a detailed explanation of the data quality process using Figure 3?

[4 Marks]

31. List and explain the three (3) types of matching logic in MS SQL Server?

[3 Marks]

32. What is metadata in the data warehousing context?

[1 Marks]

33. List four (4) types of metadata in the data warehousing context?

[2 Marks]

34. What is the primary reason for having metadata in a data warehouse?

[1 Marks]

35. Figure 4 shows the system architecture for the Amadeus entertainment data warehouse. Study it and answer the questions that follow.

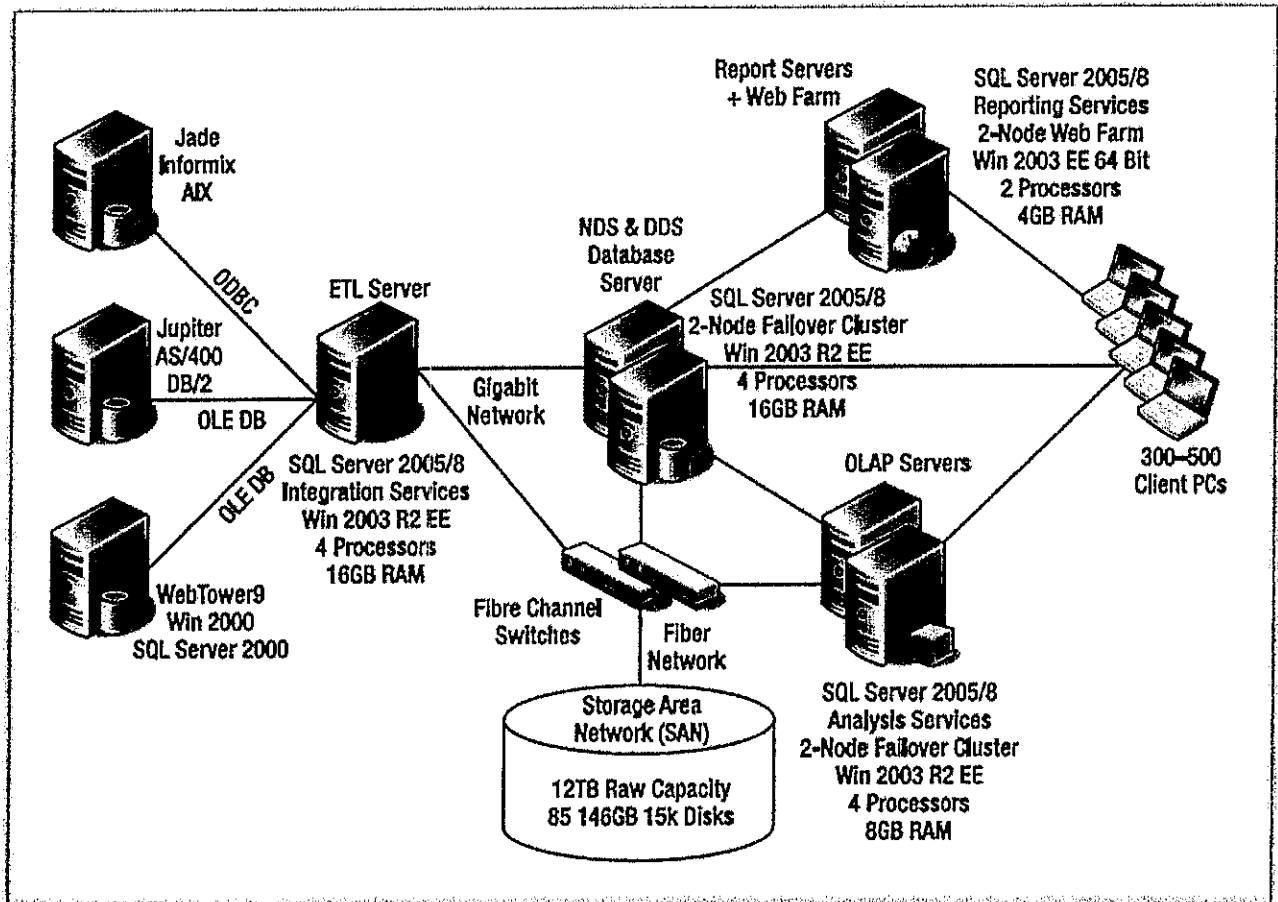


Figure 4: System architecture for the Amadeus Entertainment data warehouse
[Source: Rainardi, V. (2008)]

A. Explain failover cluster.

[1 Marks]

B. How many failover clusters are there in the above diagram?

[1 Marks]

C. Explain network load balancing.

[1 Marks]

D. Which SQL Server component is not a cluster-aware application and why?

[2 Marks]

E. Why are 64-bit server SQL Server installed and not 32bit?

[1 Marks]

THE END

