Managing Relational and Non-Relational Data

Duration: 90 minutes

Prof. António Azevedo 7 June 2019

Prof. João Loureiro Exam 1st period

Prof. Hugo Nunes Version A

Student Name: Student Nr.:

GROUP I

Grading: 7,0

(ATTENTION: each wrong answer discounts its value divided by the number of options)

1 Below is the output of ‘Orders’ table. We want to find the total sum (total order) of

each customer. Which t-SQL statement should be used?

Table output including headers:

OrderID OrderDate OrderPrice Customer

1 2018/11/23 1000 Harry

2 2018/12/04 1100 Nancy

3 2019/01/23 550 Harry

4 2018/12/23 500 Peter

( ) SELECT Customer, SUM(OrderPrice) FROM Orders

GROUP BY OrderPrice

(X ) SELECT Customer, SUM(OrderPrice) FROM Orders

GROUP BY Customer

( ) SELECT Customer, SUM(OrderPrice) FROM Orders

GROUP BY Orders

( ) SELECT OrderID, SUM(OrderPrice) FROM Orders

GROUP BY OrderID

2 What is the expected dataset that will be returned after executing t-SQL query below?

Note: Please assume that table ‘Product’ in scope of ‘Production’ schema already

exists and has one data row. Also, that one of the table columns is named ListPrice

having int as column data type.

DECLARE @Discount INT = 0;

BEGIN TRY

UPDATE Production.Product

SET ListPrice = ListPrice / @Discount

END TRY

BEGIN CATCH

PRINT ERROR\_MESSAGE();

THROW 50001, 'An error occurred', 0;

END CATCH;

( ) (0 rows affected)

( ) (0 rows affected)

Msg 50001, Level 16, State 0, Line 8

An error occurred

(X ) (0 rows affected)

Divide by zero error encountered.

Msg 50001, Level 16, State 0, Line 8

An error occurred

( ) (1 rows affected)

Divide by zero error encountered.

Msg 50001, Level 16, State 0, Line 8

An error occurred

3 Which of the following statement is false?

( ) SQL Server Database must involve at least two files (data file and transaction log

file)

( X) SQL Server Database recovery model cannot be modified after database has

been created

( ) SQL Server Database can contain more than one transaction log file

( ) SQL Server Database must have the PRIMARY filegroup

3

4 Below is the output of ‘persons’ table. What is the expected dataset that will be

returned after executing t-SQL query below?

Table output including headers:

CustomerID FirstName LastName

1 João NULL

2 P. Teixeira

3 Marta Gomez

Query:

SELECT FirstName + ' ' + LastName

FROM persons

ORDER BY CustomerID desc

( X) Marta Gomez

P. Teixeira

NULL

( ) NULL

P. Teixeira

Marta Gomez

( ) Marta Gomez

João

P. Teixeira

( ) João

Marta Gomez

P. Teixeira

5 Which of the following statement is false?

A table can have one of the following index configurations:

( ) A non-clustered index

( ) A clustered index and many non-clustered indexes

( ) A clustered index

(X ) Many clustered indexes and many non-clustered indexes

6 Below is the output of ‘SalesOrderDetail’ table. What is the expected dataset that will

be returned after executing t-SQL query below?

Table output including headers:

SalesOrderID UnitPrice OrderQty LineTotal

1 4 6 1

2 50 7 3

3 25 5 2

Query:

SELECT SalesOrderID,

SUM(UnitPrice \* OrderQty) AS TotalPrice

FROM SalesOrderDetail

GROUP BY SalesOrderID

HAVING SUM(UnitPrice \* OrderQty) > 100 AND LineTotal > 2

( ) 2 350

( ) 2 350

3 125

( ) 3 125

2 350

( X) Msg 8121, Level 16, State 1, Line 5

Column 'SalesOrderDetail.LineTotal' is invalid in the HAVING clause because it is

not contained in either an aggregate function or the GROUP BY clause.

7 Which of the following SQL Server data types cannot be implicit (automatically) converted by SQL Server engine?

( ) float to int

(X ) nvarchar to binary

( ) smallint to int

( ) int to smallint

8 Which of the following statement is false?

( ) CTEs are named table expressions defined in a query

( ) CTEs are similar to derived tables in scope and naming requirements

( X) CTEs cannot be used in a CREATE VIEW statement, as part of the view’s SELECT query

( ) Unlike derived tables, CTEs supports multiple references and recursion

9 Assuming table named ‘Employee’ is populated with 10 million data rows and

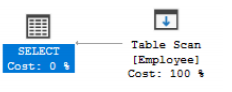
SocialSecurityNumber column is set as unique. Which execution plan is expected to be

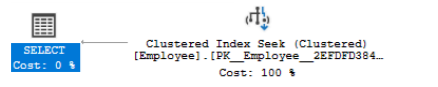
more efficient from performance point of view for this particular t-SQL query?

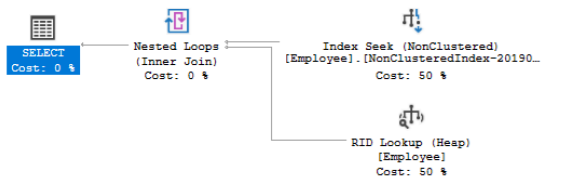
Query:

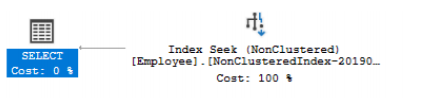
select \* from Employee

where SocialSecurityNumber = 1121212

( ) 

(X) ) 

( ) 

( ) 

10 Indicate which one of the following SQL Server Isolation Levels might allow dirty reads?

( ) Repeatable Read

( X) Read uncommited

( ) Read Committed

( ) SERIALIZABLE

11 \_\_\_\_\_\_\_\_\_\_\_\_ marks the end of a successful implicit or explicit transaction

( ) ROLLBACK TRANSACTION

(X ) COMMIT TRANSACTION

( ) IF

( ) All the above

12 Which of the following statements is false?

( ) Subqueries are nested queries. Basically, queries within queries.

( ) CROSS APPLY applies the right table expression to each row in left table

( ) Views may be referenced in a SELECT statement just like a table

( X) Scalar subquery returns multiple rows to outer query

13 Which of the following statements is false?

( X) UNION ALL removes duplicates during query processing

( ) Outer joins return all rows from one table and any matching rows from second table

( ) Self join compares rows in same table to each other

( ) Cross join combines each row from first table with each row from second table

14 Which SQL statement is not a DML (Data Manipulation Language) statement?

( ) UPDATE

(X ) CREATE

( ) DELETE

( ) INSERT

GROUP II

Grading: 3,0

1) Describe what is a SQL Server blocking scenario versus SQL Server dead lock scenario and how you

would resolve both issues.

A SQL Server blocking happens when a user locks a resource, and another user tries to use it. Usually blocking is solved rapidly and do not need to redefine the architecture of the application. With dead locks a user locks a resource, another user blocks a second resource and tries to access the resource from the first user. The first user then tries to access the resource blocked by the second user. This is an architectural issue and should be solved by either changing the application logic. In the meanwhile, the dba should understand where the deadlock is and unlock the cheapest block, by killing the process.

2) What are the differences between a primary key and a foreign key, and why both are important to

maintaining a relational database structure?

Primary keys have the purpose of identifying uniquely a record. Foreign keys have the purpose of creating relationships between entities. Foreign keys can guarantee the consistency of the data as they reinforce the constraint of having to have the value in the primary key when inserting or updating the data of the second entity.

GROUP III

Grading: 6,0

(ATTENTION: each wrong answer discounts its value divided by the number of options)

1 What factor\factors should you consider when setting a consistency level (select all

applicable)?

(X ) How fast should write operations return a success status after saving data to

persistent storage?

(X ) Is it acceptable for two users to look up a set of columns by the same row ID

and receive different data?

( ) If your application runs across multiple data centers and one of the data centers

fails, must the remaining functioning data centers have the latest data?

( ) What is the amount of storage needed so that you build your data center

accordingly?

2 Which of the following statement is true?

( ) NoSQL database guarantee Consistency, Availability or Partition Tolerance – you

can’t guarantee more than one at the same time.

( ) NoSQL have strict adherence to ACID properties.

( ) In NoSQL Consistency is traded in favor of Availability.

(X ) In RDBMS Consistency is traded in favor of Availability.

3 Select all applicable reasons for choosing a key-value database for your application.

( ) There is a need for variable attributes.

( X) The problem domain requires a relatively simple data model.

( ) Send a banner ad to an area on a web page for millions of users at the same

time.

(X ) Find references for a specific person in different scanned books.

4 Select all applicable use cases for document databases.

(X ) Product Catalog.

( ) Caching data from relational databases to improve performance.

( X) Managing highly related data.

( ) Fraud Detection.

5 Which statement is true (select all applicable)?

( X) In a Column-family database a row is composed of columns that are grouped by

a row key.

( ) Column-family databases are appropriate to modeling data warehouses and

data marts.

(X ) Column-family databases are appropriate to work with Big Data and scale

volumes of data.

( X) Column family databases are appropriate when a large number of servers are

required to meet expected workloads.

6 What is an Unweighted Graph (select all applicable)?

( ) Graphs which Vertexes don’t have a weight.

( ) Graphs which Edges don’t have a weight.

(X ) Graphs were both Edges and Vertexes don’t have weights.

( ) Graphs where Edges have a direction.

7 Select all types of Graph Algorithms?

(X ) Path finding

(X ) Centrality

(X ) Community detection

( ) Recurrence

8 You have a key-value entry with a reference to a file in blob storage, you need to

modify the name of the file and ensure that the key-value entry is updated

accordingly.

Which pattern should you use?)

() Compound key pattern.

(X ) Denormalization pattern.

( ) Inter-partition secondary index.

( ) Eventually consistent transactions pattern.

9 When designing Azure Storage Tables which is the least read-efficient option?

( ) $filter= PartitionKey eq 'NovaIMS' and RowKey eq 75

( ) $filter=PartitionKey eq 'NovaIMS' and RowKey ge 10 and RowKey lt 80

( ) $filter=PartitionKey eq 'NovaIMS' and Unit eq 'edsa'

( X) $filter=Unit eq 'edsa'

10 Which of these object types are not part of a collection in an Azure Cosmos SQL API

(select all applicable)?

( X) Attachment

( ) Document

( ) Stored Procedure

( ) Permission

11 Which of these stages belong to Map Reduce (select all applicable)?

(X ) Map

( X) Reduce

( X) Shuffle

( ) Redistribute

12 Which of the following are motivating factors for database designers and other IT

professionals to develop and use NoSQL databases (select all applicable)?

( X) Elastic Scaling

(X ) Flexible Models0

( X) Large amounts of data

( ) Maturity

GROUP IV

Grading: 4,0

1) How does vertical partitioning differ from horizontal partitioning?

2) Name the four main NoSQL database types and their strengths for the following scenarios:

Key-value Webpage cache,

Graph - social networks,

Document - product catalog,

Column oriented IoT telemetry.

3) Name two types of formats for storing data in a document database.

4) The E in BASE stands for eventually consistent. What does that mean?