



ACTIVIDAD

Gap fill activity about the previous reading.

Se alcanza la aprobación con el 70%

Fill in the blanks with the appropriate terms from the reading:

1. Data ______, also known as horizontal partitioning, involves dividing a large dataset into smaller, more manageable subsets or partitions. Each partition contains a subset of the data, often based on specific criteria such as a range of values, geographical location, or other relevant attributes.

2. Implementing data partitioning offers a range of benefits, including enhanced _______. By distributing data across multiple partitions, data partitioning allows for parallel processing, which can significantly improve system performance. Each partition can be processed independently, resulting in faster query execution times and reduced latency.

3. Selecting the appropriate ______ is crucial for the success of data partitioning. The partitioning key should be carefully chosen based on the characteristics of the dataset and the specific requirements of the application. Factors such as selectivity, query performance, and data distribution should be considered.

4. Data ______ is one challenge in data partitioning, referring to the uneven distribution of data across partitions. It can lead to hotspots, where certain partitions receive a disproportionately high load of queries or updates, while others remain underutilized. Mitigating data skew involves techniques such as dynamic partitioning and sub-partitioning.

5. Several modern database management systems (DBMS) provide built-in support for data partitioning, making it easier to implement and manage partitioned datasets. For example, Oracle offers a robust partitioning feature that allows for efficient data organization and management. Microsoft SQL Server and PostgreSQL also provide partitioning features to enable efficient data storage and processing. These DBMS with partitioning capabilities are essential tools for handling ______ in various applications, providing scalable and efficient processing capabilities for large-scale data.

