

CMU 15-445 - Database Systems Lecture 4 - Database Storage 2

Database Storage 2

Data Representation

This section describes how the DBMS stores the values inside the tuples in memory. Tuples' data is essentially just byte arrays. It is up to the DBMS to know how to interpret those bytes to derive the values for attributes. A data representation scheme is how a DBMS stores the bytes for a value. There are four main types that can be stored in tuples: integers, variable precision numbers, fixed point precision numbers, variable length values, and dates/times.

Integers, Floats, Reals

These are usually just stored in the same way C/C++ store them, as specified by the IEEE-754 standard.

Fixed Point Precision Numbers

The DBMS defines how it will handle fixed point precision numbers, and implements an API to access them itself. It does not rely on CPU instructions, since they have rounding errors. Instead, the entire precision of the number is stored in the representation, along with additional metadata telling the DBMS how to interpret the value.

These types are used when rounding errors are unacceptable, for example in scientific or financial applications. However, handling these types is much slower than floating point numbers.

Variable Length Data

These are data types that consist of a varying amount of data that needs to be stored. For example, `TEXT`, `VARCHAR`, `BLOB`, `VARBINARY`. Has a header that keeps track of the length of the string to make it easy to jump to the next value. Most DBMSs do not

allow a tuple to exceed the size of a single page, so they solve this issue by writing the value on an overflow page and have the tuple contain a reference to that page.

Some systems will let you store these large values in an external file, and then the tuple will contain a pointer to that file. For example, if our database is storing photo information, we can store the photos in the external files rather than having them take up large amounts of space in the DBMS. One downside of this is that the DBMS cannot manipulate the contents of this file.

Datetime Data

These are usually just represented as an integer equal to the number of milli or micro seconds passed since the UNIX epoch.

Workload Types

Depending on the type of work you want to do with your data, systems can be divided into two types - Online Transaction Processing (OLTP), and Online Analytical Processing (OLAP).

OLTP

OLTP basically means you want transactions on your database to be fast. It is usually used when your application is read heavy, and you want to collect potentially large amounts of data from your frontend efficiently.

Characteristics of OLTP systems are - 1. Fast, short running operations. 2. Queries usually only operate on a single entity at a time (hence joins are rare). 3. Application is read-heavy. 4. Usually collect more data than they need to analyze.

OLAP

OLAP means that you want joins and other complex operations involving multiple parts of your database to be fast, usually because you want to perform analytics.

Characteristics of OLAP systems are - 1. Long running, complex queries accessing multiple columns, tuples, and tables. 2. Long scans through entire tables. 3. Analytical and exploratory queries. 4. Usually analyze and read more data than they write.

Storage Models/Methods

The relational model does not specify that you should store your tuples contiguously. There are two types of storage models - Decomposition storage model (column

storage) and N-Ary storage model (row storage).

NSM

NSM stores all the columns of a single tuple contiguously in the same page. This approach is ideal for OLTP workloads where transactions tend to operate only on an individual entity and insert heavy workloads. It is ideal because it takes only one fetch to be able to get all of the attributes for a single tuple.

Its advantage is fast inserts and updates, and it is good for queries that need the entire tuple. Its disadvantage is queries that use only a small portion/number of columns of the tuple.

Therefore, row stores are usually not used for OLAP systems.

DSM

DSM stores a single attribute or column for all tuples contiguously in a single page.