Lecture 3

Hybrid-Range Partitioning Strategy

Controlled placement of data based on the processing capability of the underlying hardware & the system workload, so we want to make sure that the overhead of parallelism is absolute minimum, such that the query becomes expedited by using multiple computers, we also want to maximize degree of parallelism to balance the workload more evenly.

Terminology:

1024 Bytes is one kilobyte ; 1024 KB is one Megabyte

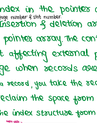
1024 Megabytes is one Gigabyte ; 1024 GB is one terabyte.

1024 Terabytes is one Petabyte ; 1024 PT is one Exabyte

1024 XB is one Zetabyte.

DBMS Manufactures → 1000 KB

Operating Systems → 1024 KB

Indexed Heap / Physical organization & Records in Blocks

Each page consists of an array of pointers, each pointer points to a record within the block. A record is located by providing its block numbers & index in the pointer array. This combination is called TID & RID. Insertion & deletion are easy, accomplished by manipulating the pointer array. The contents of a block may be reorganized without affecting external pointers pointing to a record. RID does not change when records are moved around within a block. When deleting a record, you take the record out & set the slot number to be null; we'll reclaim the space from the disk page, & now reorganize & reconstruct the index structure from scratch.