

ever, can result in vastly different steps by the DBMS to answer the query. How the DBMS thinks about using a file organization (e.g., what index and how and when it uses a hashing algorithm), you can design better, more efficient queries.

The three families of file organizations cover most of the file organizations you will have at your disposal as you design physical files and databases. More complex structures can be built using the data structures outlined in Appendix A, but you are unlikely to be able to use these with a database management system.

Table 5-3 summarizes the comparative features of sequential, indexed, and hashed file organizations. You should review this table and study Figure 5-7 to see which comparative feature is true.

Clustering Files

Some database management systems allow adjacent secondary memory to contain rows from several tables. For example, in Oracle, rows from one or more related tables that are often joined together can be stored so that they are adjacent data blocks (the smallest storage units). A cluster is defined by the tables or columns by which the tables are usually joined. For example, a Clustered Index

TABLE 5-3 Comparative Features of Different File Organizations

Factor	File Organization		
	Sequential	Indexed	Hashed
Storage space	No wasted space	No wasted space for data but extra space for index	Extra space may be required for addition and deletion after the initial set is loaded
Sequential retrieval on primary key	Very fast	Moderately fast	Impractical, unless index
Random retrieval on primary key	Impractical	Moderately fast	Very fast
Multiple-key retrieval	Possible but requires scanning whole file	Very fast with multiple indexes	Not possible unless index
Deleting records	Can create wasted space or require reorganizing	If space can be dynamically allocated, this is easy but requires maintenance of indexes	Very easy
Adding new records	Requires rewriting a file	If space can be dynamically allocated, this is easy but requires maintenance of indexes	Very easy, but multiple same address required
Updating records	Usually requires rewriting a file	Easy but requires maintenance of indexes	Very easy