Evaluation of Relational Operations: Other Techniques

Chapter 12, Part B

Using an Index for Selections

- Cost depends on # qualifying tuples, and clustering.
  - Cost of finding qualifying data entries (typically small) plus cost of retrieving records (could be large w/o clustering).
  - In example, assuming uniform distribution of names, about 10% of tuples qualify (100 pages, 10000 tuples). With a clustered index, cost is little more than 100 I/Os; if unclustered, up to 10000 I/Os!

- Important refinement for unclustered indexes:
  1. Find qualifying data entries.
  2. Sort the rid's of the data records to be retrieved.
  3. Fetch rid's in order. This ensures that each data page is looked at just once (though # of such pages likely to be higher than with clustering).

Two Approaches to General Selections

- First approach: Find the most selective access path, retrieve tuples using it, and apply any remaining terms that don't match the index:
  - Most selective access path: An index or file scan that we estimate will require the fewest page I/Os.
  - Terms that match this index reduce the number of tuples retrieved; other terms are used to discard some retrieved tuples, but do not affect number of tuples/pages fetched.
  - Consider day<8/9/94 AND bid=5 AND sid=3. A B+ tree index on day can be used; then, bid=5 and sid=3 must be checked for each retrieved tuple. Similarly, a hash index on <bid, sid> could be used; day<8/9/94 must then be checked.
Intersection of Rids

- **Second approach** if we have 2 or more matching indexes that use Alternatives (2) or (3) for data entries:
  - Get sets of rids of data records using each matching index.
  - Then **intersect** these sets of rids (we’ll discuss intersection soon)
  - Retrieve the records and apply any remaining terms.
  - Consider \( \text{day}<8/9/94 \text{ AND bid}=5 \text{ AND sid}=3 \). If we have a B+ tree index on \( \text{day} \) and an index on \( \text{sid} \), both using Alternative (2), we can retrieve rids of records satisfying \( \text{day}<8/9/94 \) using the first, rids of recons satisfying \( \text{sid}=3 \) using the second, intersect, retrieve records and check \( \text{bid}=5 \).
Discussion of Projection

- Sort-based approach is the standard; better handling of skew and result is sorted.
- If an index on the relation contains all wanted attributes in its search key, can do **index-only** scan.
  - Apply projection techniques to data entries (much smaller!)
- If an ordered (i.e., tree) index contains all wanted attributes as **prefix** of search key, can do even better:
  - Retrieve data entries in order (index-only scan), discard unwanted fields, compare adjacent tuples to check for duplicates.

Set Operations

- Intersection and cross-product special cases of join.
- Union (Distinct) and Except similar; we’ll do union.
- Sorting based approach to union:
  - Sort both relations (on combination of all attributes).
  - Scan sorted relations and merge them.
  - **Alternative**: Merge runs from Pass 0 for both relations.
- Hash based approach to union:
  - Partition R and S using hash function \( h \).
  - For each S-partition, build in-memory hash table (using \( h2 \)), scan corr. R-partition and add tuples to table while discarding duplicates.

Aggregate Operations (AVG, MIN, etc.)

- Without grouping:
  - In general, requires scanning the relation.
  - Given index whose search key includes all attributes in the **SELECT or WHERE** clauses, can do index-only scan.
- With grouping:
  - Sort on group-by attributes, then scan relation and compute aggregate for each group. (Can improve upon this by combining sorting and aggregate computation.)
  - Similar approach based on hashing on group-by attributes.
  - Given tree index whose search key includes all attributes in **SELECT, WHERE and GROUP BY** clauses, can do index-only scan; if group-by attributes form prefix of search key, can retrieve data entries/tuples in group-by order.
**Impact of Buffering**

- If several operations are executing concurrently, estimating the number of available buffer pages is guesswork.
- Repeated access patterns interact with buffer replacement policy.
  - e.g., Inner relation is scanned repeatedly in Simple Nested Loop Join. With enough buffer pages to hold inner, replacement policy does not matter. Otherwise, MRU is best, LRU is worst (*sequential flooding*).
  - Does replacement policy matter for Block Nested Loops?
  - What about Index Nested Loops? Sort-Merge Join?

---

**Summary**

- A virtue of relational DBMSs: *queries are composed of a few basic operators*; the implementation of these operators can be carefully tuned (and it is important to do this!).
- Many alternative implementation techniques for each operator; no universally superior technique for most operators.
- Must consider available alternatives for each operation in a query and choose best one based on system statistics, etc. This is part of the broader task of optimizing a query composed of several ops.