CS 186 - Fall 2024
Exam Prep Section 10
Parallel Query Processing

1 Types of Parallelism

For each of the following scenarios, state whether it is an example of:

- Inter-query parallelism
- Intra-query, inter-operator parallelism
- Intra-query, intra-operator parallelism
- No parallelism

1. A query with a selection, followed by a projection, followed by a join, runs on a single machine with one thread.

   Answer: No parallelism. It might look like a pipeline, but at any given point in time there is only one thing happening, since there is only one thread.

2. Same as before, but there is a second machine and a second query, running independently of the first machine and the first query.

   Answer: Inter-query parallelism.

3. A query with a selection, followed by a projection, runs on a single machine with multiple threads; one thread is given to the selection and one thread is given to the projection.

   Answer: Intra-query, inter-operator parallelism.

4. We have a single machine, and it runs recursive hash partitioning (for external hashing) with one thread.

   Answer: No parallelism, because there is only one machine and one thread. Don't confuse this with parallel hashing!

5. We have a multi-machine database, and we are running a join over it. For the join, we are running parallel sort-merge join.

   Answer: Intra-query, intra-operator parallelism. We have a single query and a single operator, but that single operator is going to do multiple things at the same time (across different machines).

2 Partitioning for Parallelism

1. Suppose we have a table of size 50,000 KB, and our database has 10 machines. Each machine has 100 pages of buffer, and a page is 4 KB.

   We would like to perform parallel sorting on this table, so first, we perfectly range partition the data. Then on each machine, we run standard external sorting.

   How many passes does this external sort on each machine take?

   Answer: 2 passes.
   After range partitioning, each table will have 5,000 KB of data, or 1,250 pages. With 100 pages of buffer, this will take 2 passes to sort.

2. Suppose we were doing parallel hash join. The first step is to partition the data across the machines, and we usually use hash partitioning to do this.

   Would range partitioning also work? What about round-robin partitioning?

   Answer: Range partitioning also works, because items with the same key still end up on the same machine as required. Round-robin partitioning does not do that, so it does not work.

3. Suppose we have a table of 1200 rows, perfectly range-partitioned across 3 machines in order. We just bought a 4th machine for our database, and we want to run parallel sorting using all 4 machines.

   The first step in parallel sorting is to repartition the data across all 4 machines, using range partitioning. (The new machine will get the last range.)

   For each of the first 3 machines, how many rows will it send across the network during the repartitioning? (You can assume the new ranges are also perfectly uniform.)

   Answer: 100 rows, 200 rows, and 300 rows.
   The original partitions were 3 ranges of 400 rows each; the new 4 ranges will have 300 rows each.
   The first machine held the first 400 rows originally, and now only needs to hold the first 300. It will send the remaining 100 rows over the network (to machine 2).
   The second machine held rows 401-800 initially, but now needs to hold rows 301-600. It will send rows 601-800 (200 rows) to machine 3.
   The third machine held rows 801-1200 initially, and similarly needs to send rows 901-1200 (300 rows) to machine 4.

3 Parallel Query Processing

Suppose we have 4 machines, each with 10 buffer pages. Machine 1 has a Students table which consists of 100 pages. Each page is 1 KB, and it takes 1 second to send 1 KB of data across the network to another machine.

1. How long would it take to send the data over the network after we uniformly range partition the 100 pages? Assume that we can send data to multiple machines at the same time.
   Answer: 25 seconds.
   After we uniformly partition our data, Machine 1 will send 25 pages to Machines 2, 3, and 4. It will take 25 seconds to finish sending these pages to each machine if we send the pages to each machine at the same time.

2. Next, imagine that there is another table, Classes, which is 10 pages. Using just one machine, how long would a BNLJ take if each disk access (read or write) takes 0.5 seconds?
   Answer: 105 seconds. ,
   BNLJ will require 10 + ceil(10/8) * 100 = 210 I/Os, which will take 105 seconds.

3. Now assume that the Students table has already been uniformly range partitioned across the four machines, but Classes is only on Machine 1. How long would a broadcast join take if we perform BNLJ on each machine? Do not worry about the cost of combining the output of the machines.
   Answer: 40 seconds.
   First, we must broadcast the Classes table to each machine, which will take 10 seconds (since we send the table to each machine at the same time). Next, we will perform BNLJ on each machine, which requires 10 + ceil(10/8) * 25 = 60 I/Os, or 30 seconds. In total, the time required to send the data over the network and perform the join will be 40 seconds.

4. Which algorithm performs better?
   Answer: Broadcasting the Classes table and performing a parallel BNLJ runs faster than just using one machine, even with the additional time required to send the table over the network.

5. Knowing that the Students table was range partitioned, how can we improve the performance of the join even further?
   Since we know the Students table was range partitioned, we can also range partition the Classes table on the same column and only send the corresponding partitions to each machine. This should lower the network cost and decrease the number of disk I/Os.