

Project 0 Implementation Overview

Scan

Scan operator casts *scannable* to *rowstore* and read tuples from *rowstore*.

Filter

Filter operator reads input tuples and then filter them with the provided *predicate* function.

Project

Project operator reads input tuples and then apply the provided *evaluator* function to select the desired columns.

Aggregate

Aggregate operator first groups input tuples into a hash table using the provided *groupby* key. Then for each group, for each aggregate call, reduce each group using each aggregate call to obtain a tuple of aggregated values for each group in the form of *(groupkey, aggregatedVal_1, aggregatedVal_2 ..., aggregatedVal_n)*.

Join

Join operator groups input tuples from left table input into a hash table where key is the provided *getleftkey*. Then for each right tuple, get the right key and search the hash table to find left tuples with the matching key. Finally concatenate all matched left tuples with this right tuple.

Sort

Sort reads input tuples into a priority queue, with the ordering provided by an *Ordering[Tuple]* object. If required output is from the top *n*th tuple to the top *m*th tuple, then top *m* tuples are dequeued and then the top *n-1* tuples are dropped. For example, if after sorting the 101th to the 200th tuple are required, then we will dequeue 200 tuples from the priority and drop the first 100 tuples.