**Vectorization in Map Reduce**

A framework for writing applications that process large amounts of data

**Enable Vectorization**

Vectorization allows Hive to process a batch of rows together as a unit instead of processing one row at a time. Each batch consists of a column vector which is usually an array of primitive types. Operations are performed on the entire column vector, which improves the instruction pipelines and cache usage.  Files must be stored in ORC format in order to use vectorization.

To enable vectorization,

set hive. vectorized.execution.enabled=true;​

Vectorization is disabled by default.