

Attachment:

NEC accelerates machine learning for vector computers

The new technology features the following.

1. Accelerates sparse matrix computing through the development of a method suitable for vector computers
NEC developed a hybrid format using sparse matrix in which processing is appropriately executed by column or by row depending on the number of non-zero portions. This enables the high-speed processing of machine learning without decreasing the processing efficiency of vector computers.

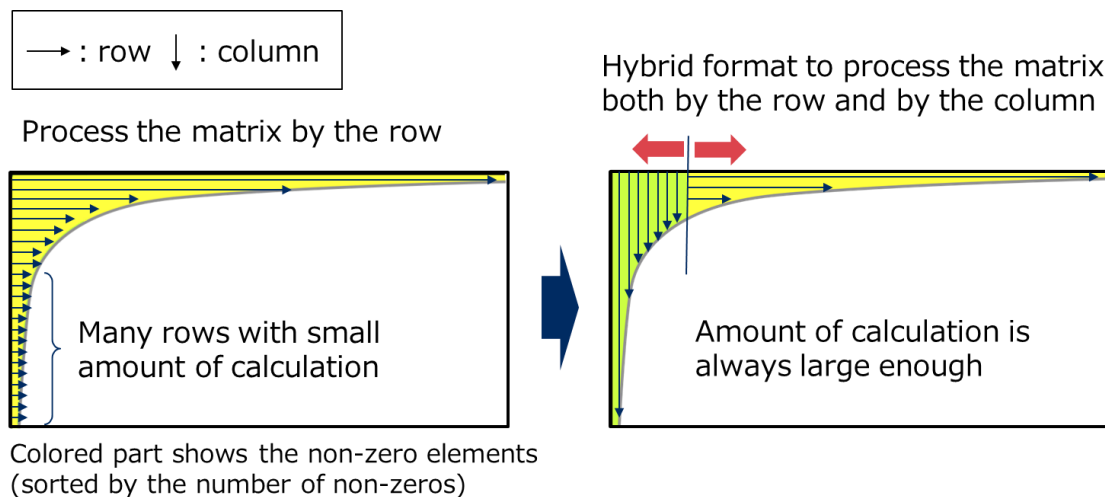


Figure 1: Data processing method that speeds up sparse matrix computing

2. Reduces communication volume in machine learning for sparse matrix

In a sparse matrix that is used as input data in machine learning, columns occur that have no value in any row, after the division of the data for distributed processing. In machine learning processing, because the processing result is updated

depending on the values in columns, the portions corresponding to the columns that have no value are not updated. By removing these non-updated portions from communication, the reduction of communication volume is achieved.

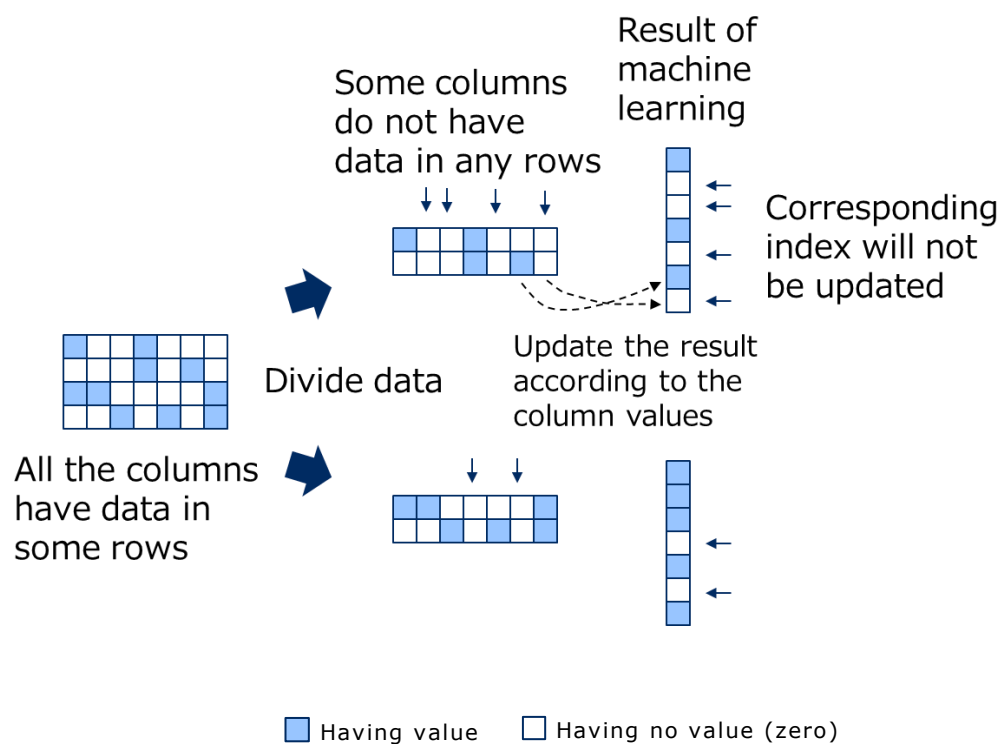


Figure 2: Data communication method that achieves a reduction in communication volume

3. Develops middleware incorporating the above technology
 By implementing this middleware using C++ and MPI, efficient processing can be executed. In addition, as is the case with Spark, the middleware has been designed so that parallel processing straddling multiple processors can be easily described. Moreover, the middleware can be launched

from Spark and Python in the same format as the provided machine learning library.

As for the execution time spent on machine learning, we compared a case in which Spark is executed in a cluster where multiple servers are connected and a case in which this middleware is executed in SX-ACE, a vector computer produced by NEC. It was found that the execution time of the latter is more than 50 times faster than that of the former.

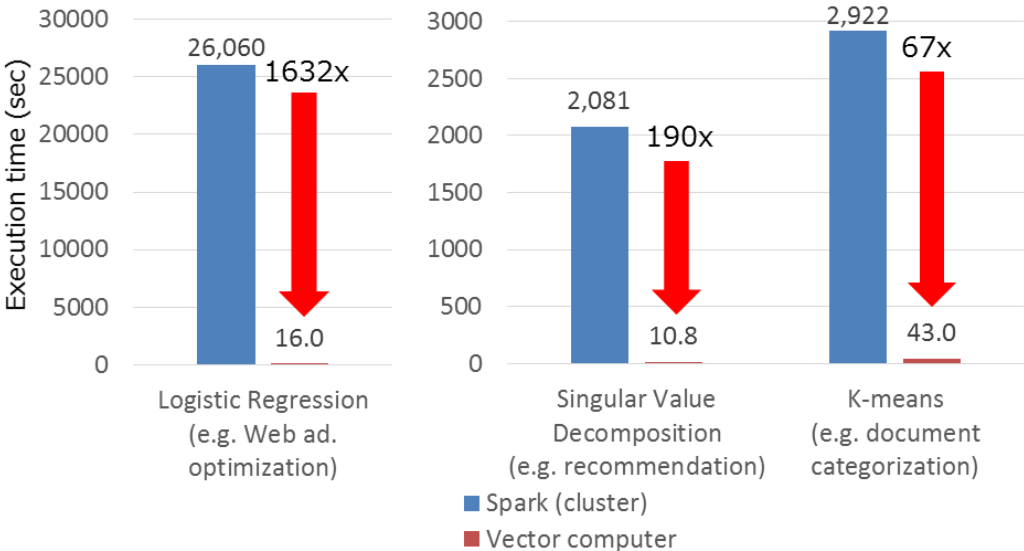


Figure 3: Comparison of data processing speed at the same number of cores (64 cores)