MapReduce is an important paradigm for solving big and costly problems when calculating or analyzing big data problems today.

Almost every problem that can be divided into small pieces and processed independent can be distributed and calculated in parallel by following the MapReduce paradigm.

The following shows an exemplary list of problems that can be efficiently solved by MapReduce:

- Factorization of (big) Integers

- Matrix factorization

- Fourier transformations

- Genetic analysis

- Search engine indexing

- Users' behavior analytics

- Group text documents into topically related groups

MapReduce processing is computed by large clusters. Today it is more a matter of being interested in the results, than having the money to do buy an own cluster. Almost all providers of cloud computing are having resources or hardware optimized solutions for doing MapReduce.

What do you think: Which are the greatest benefits of calculating big data problems through MapReduce on inexpensive cluster environments? Do you think this ability will be used for good or bad things in the future?