# Hama Overview

**Hama** is a parallel matrix computational package, which provides an library of matrix operations for the large-scale processing development environment and Map/Reduce framework for the large-scale Numerical Analysis and Data Mining, which need the intensive computation power of matrix inversion, e.g. linear regression, PCA, SVM and etc. It will be also useful for many scientific applications, e.g. physics computations, linear algebra, computational fluid dynamics, statistics, graphic rendering and many more. Currently, several shared-memory based parallel matrix solutions can provide a scalable and high performance matrix operations, but matrix resources can not be scalable in the term of complexity. The **Hama** approach proposes the use of 2-dimensional Row and Column(Qualifier) and Time space and multi-dimensional Columnfamilies of Hbase, which is able to store large sparse and various type of matrices (e.g. Triangular Matrix, 3D Matrix, and etc.) and utilize the 2D blocked algorithm. In addition, auto-partitioned sparsity sub-structure will be efficiently managed and serviced by Hbase. Row and Column operations can be done in linear-time, where several algorithms such as structured Gaussian elimination and iterative methods run in O(the number of non-zero elements in the matrix / number of mappers) time on Hadoop Map/Reduce.