

# MapReduce MPI

The purpose of this assignment is for you to learn more about

- Map Reduce
- How to write simple problems in Map Reduce MPI

## Installing MapReduce-MPI on mamba

Download the archive to mamba (probably through your git repository). Uncompress it with `tar zcvf mrmpi.tar.gz`. go to the src directory. And compile the library with `CXX=mpicxx CC=mpicc LD=mpicxx make -e linux`.

Compile you own MapReduce MPI application by passing

- `-I path/to/mrmpi/src` at compile time
- and `path/to/mrmpi/src/libmrmpi_linux.a` at link time (make sure it is the last parameter of the linker).

To summarize, I compiled with: `mpicxx -O3 -I /users/esaule/prog/mrmpi/src -c -o foomrmpi.o foomrmpi.cpp` and linked with `mpicxx -o foomrmpi foomrmpi.o /users/esaule/prog/mrmpi/src/libmrmpi_linux.a`

## 1 Word Count

**Question:** write a MapReduce-MPI program that takes a list of filenames and count the number of time each word appears in all the files. Output the result in multiple files if needed (usually one file per rank).

Find some texts from project gutenber on canvas.

## 2 Sorting

**Question:** Adapt the previous program to output the words in lexicographical order (order they would appear in a dictionary).

## 3 Histogram

**Question:** Adapt the previous program to generate a table of how many words appear  $x$  times.