

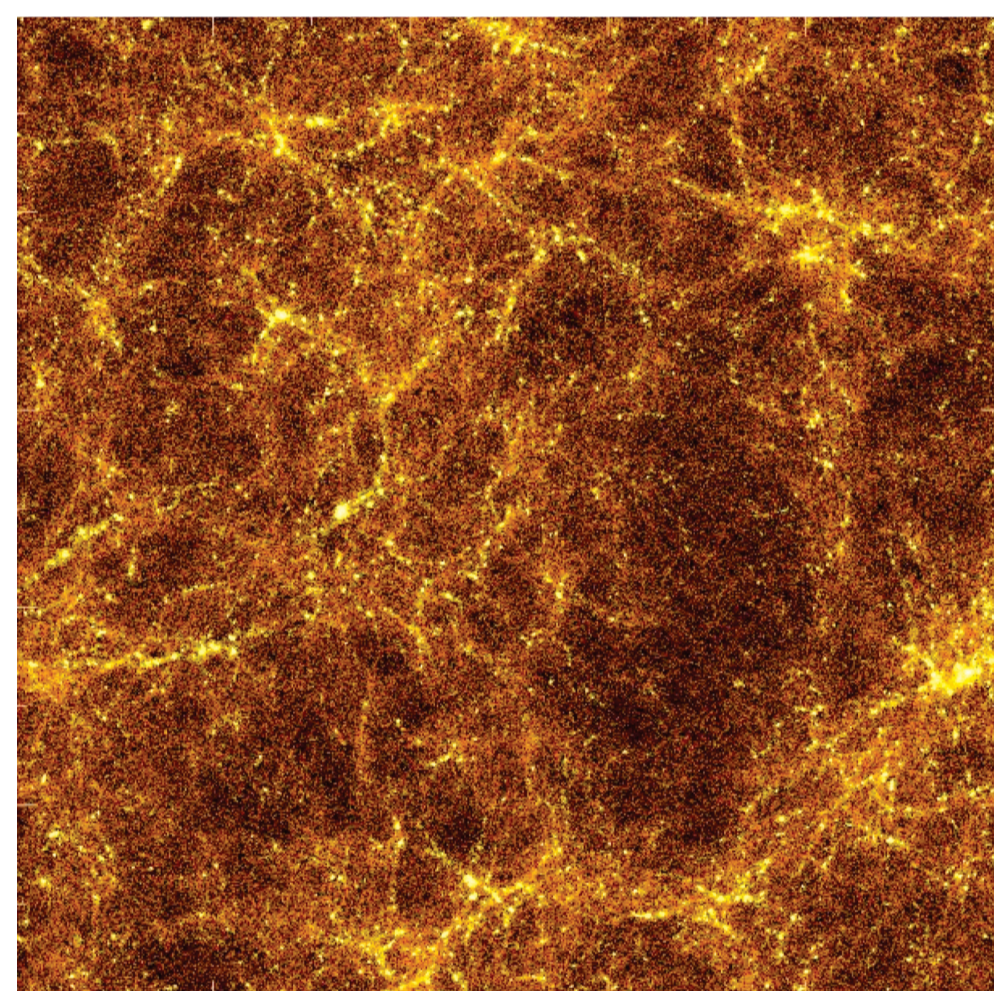
# MPWide, the light-weight communication library



MPWide is a light-weight communication library which connects two applications, each of them running with the locally recommended MPI implementation. We originally developed it to manage the long-distance message passing in the CosmoGrid project, where cosmological N-body simulations run on grids of supercomputers connected by high performance optical networks.

## Motivation

To take full advantage of the network light paths in CosmoGrid, we need a message passing library that supports the ability to use customized communication settings (e.g. custom number of streams, window sizes) for individual network links among the sites. The supercomputers we use vary both in hardware architectures and software setup.



## Application

In the CosmoGrid simulation, we simulate dark matter particles using a parallel tree/particle-mesh N-body integrator, TreePM. This requires relatively little communication between different sites after each timestep. The integrator can be run as a single MPI application, or as two separately launched MPI applications on different supercomputers.

## Benchmarks

We tested MPWide in a production environment, during a CosmoGrid run. In this run, we used the Huygens supercomputer in Amsterdam and the Cray supercomputer in Tokyo. In this run, the calculation time dominated the overall performance, with the communication time constituting about one eighth of the total execution time.

