The main goal of the MIDAS’09 project was to create an efficient collaboration of physicists, applied mathematicians and computer scientists to address challenges faced by one of the most active areas of modern cosmology, i.e., studies of the Cosmic Microwave (CMB) anisotropies. CMB is a unique source of information about the Universe, its composition and early evolution (Big Bang), as well as fundamental laws of physics. This information needs to be extracted from observational data sets. This is very challenging as they keep growing in volume at the Moore’s rate and are already reaching up to hundreds of Tera bytes; have the cosmological signals diluted over billions of measurements, which need to be therefore analyzed simultaneously (no embarrassingly parallel solutions); require the largest supercomputing platforms for their analysis due to needed memory and processing power; require high performance, scalable numerical algorithms to render the expected precision and reliability.

The MIDAS’09 collaboration proposed to address such challenges by developing techniques and algorithms for massively parallel computational platforms for some of the most important stages of the CMB data analysis and to turn them into efficient and scalable, numerical libraries.

PROJECT OBJECTIVES

We have targeted two specific sets of the data analysis problems with particularly broad and profound potential impact on the CMB data analysis. These are:

- Spherical harmonic transforms;
- Map-making operations.

Spherical harmonic transforms (SHT) are among the most common operations performed as part of the statistical characterization of the CMB, as well as many other, e.g., geophysical, oceanographical, data sets. CMB requirements are however particularly stringent due to their unprecedented resolutions, very large data volumes and therefore a large number of FLOPs.

- We have developed new parallel implementations of the SHT for standard, many-core and hybrid (CPU/GPU) computational platforms (Hupau et al 2012, Szydlarski et al 2013);
- We have demonstrated that they allow for efficient parallelization of the transforms up to $O(10^7)$ of cores (and $O(10^5)$ of CPU/GPU units) as required by the modern CMB data analysis, removing some of the previous limitations;
- We have produced a library called $S^4$HAT (Scalable, Spherical Harmonic Transforms) (Fabbian et al 2012).

Map-making operations:

Map-making is an essential analysis step on which the raw data collected by an experiment are turned into estimates of the actual sky signals. This is done by iteratively solving huge linear system of equations. The major challenges are due to the data volumes, need for intensive interprocessor communication, and CPU time required for convergence.

- We have proposed novel two-level preconditioner for solving the map-making equation, improving the time-to-solution by a factor of up to 6 (Grigori et al 2012, 2013);
- We have developed new algorithms for a customized reduce operation of the irregularly distributed data, which in our application superseded standard MPI_Allreduce call by a factor of up to 60 (Cargemel et al 2013);
- We have studied communication avoiding solutions to the CMB map-making problem (Sharify et al 2012);
- We have implemented public, massively parallel, middle-level numerical library for CMB data analysis, MIDAPACK.

CONCLUSIONS

MIDAS’09 has brought CMB data analysis closer to requirements of current, cutting-edge supercomputer platforms, producing very encouraging results. A number of MIDAS’09 produced tools have been already incorporated in user codes (XPOL, XPURE, Lens$^4$HAT). Others are being implemented within the data analysis pipelines of current CMB experiments. More work ahead.