Enhancing low-rank updates performance for sparse direct solvers

Grégoire Pichon, Bora Uçar
Contact : gregoire.pichon@univ-lyon1.fr

Laboratoire LIP, équipe ROMA, Lyon (http ://www.ens-lyon.fr/LIP/ROMA)

Description

Solving sparse linear systems is a key operation in various academic and industrial applications. A classical example is solving partial differential equations for 2D or 3D meshes. Sparse direct solvers are widely used due to their numerical properties that allow them to solve most problems. However, they suffer from their high time and memory complexities. In order to enhance sparse direct solvers, low-rank compression techniques have been recently introduced. Those methods consist of compressing some blocks of the matrix in order to reduce the memory and the cost of the operations associated to. Block Low-Rank (BLR) compression has demonstrated all its potential for MUMPS[1] et PaStiX[2] sparse direct solvers. This approach being somewhat new, various research directions are followed to try to enhance the behaviour of sparse direct solvers when using low-rank compression.

In this internship, we propose to study and enhance one of the basis kernels used in the factorization process : the update of a block. When low-rank compression is used, this operation implies blocks which sizes and properties differ. This is a burden for the efficiency of solvers as it represents the most costly operation.

We propose to follow three direction to enhance the update process :

— Accumulate several updates together before recompression
— Compress as a single low-rank representation several elementary blocks to gain in terms of data granularity
— Represent a bloc as a sparse part and a low-rank update part

This work will be realized in the context of the PaStiX[3] sparse direct solvers, that can solve systems made of millions of unknowns on top of parallel architectures.

References

[1] T. Mary, "Block Low-Rank multifrontal solvers : complexity, performance, and scalability", Ph.D. dissertation, Toulouse University, Toulouse, France, Nov. 2017

[2] G. Pichon, "On the use of low-rank arithmetic to reduce the complexity of parallel sparse linear solvers based on direct factorization techniques", Ph.D. dissertation, Université de Bordeaux, Talence, France, Nov. 2018.

[3] P. Hénon, P. Ramet, and J. Roman, "PaStiX : A High-Performance Parallel Direct Solver for Sparse Symmetric Definite Systems," Parallel Computing, vol. 28, no. 2, pp. 301–321, Jan. 2002.

Environnment

— This internship will take place in the LIP laboratory, located at the ENS Lyon.

— It will be possible to continue this intership with a PhD, around low-rank compression and sparse direct solvers
— For any information, please contact us by mail : gregoire.pichon@univ-lyon1.fr and bora.ucar@ens-lyon.fr
— There will be a collaboration with Mathieu Faverge and Pierre Ramet from Inria Bordeaux Sud-Ouest

Skills

— C programming
— Linear Algebra