

Automatic selection of GPU sparse triangular solvers based on energy consumption

Tuesday, 5 November 2019 11:00 (20 minutes)

Preconditioned Krylov-subspace methods to solve general sparse linear systems are the computational bottleneck of the solution of many science and engineering problems. In particular, it is the preconditioner application on each iteration of the solver, the stage that concentrate the most of the processing time. Many times this stage implies the solution of a number of sparse triangular linear systems, which has motivated their study by the NLA and HPC community.

The parallelization of this procedure in accelerators such as GPUs is widely addressed in the literature. Up to recently, the dominant approach to tackle this operation was the the `level-set` strategy, which relies on preprocessing the sparse matrix to determine sets of independent unknowns that can be solved for in parallel. The most established and mature example of this approach is the routine distributed with the cuSparse library of NVIDIA. In recent work some authors have proposed a number of routines based on the `self-scheduling` paradigm. In this paradigm the execution schedule is decided dynamically as threads have to wait until their data dependencies are resolved by other threads. The experimental results comparing both approaches are not conclusive, and indicate that both paradigms are well suited for certain matrices and not for others.

In previous work we have used machine learning models to attempt to predict which will be the best performing variant of the routine for a given sparse matrix. Now we are interested in incorporating the dimension of energy consumption, studying the relation between runtime and power consumption of each variant and using machine learning to decide which solver to use considering both runtime and enregy.

Primary authors: MARICHAL, Raúl (Universidad de la República); PASTORINI, Marcos (Universidad de la República); DUFRECHOU, Ernesto (Universidad de la República); EZZATTI, Pablo (HCL, INCO, FING, UDELAR); QUINTANA ORTÍ, Enrique S. (Universitat Politècnica de València)

Presenter: MARICHAL, Raúl (Universidad de la República)

Session Classification: Day I