

PLASMA

PLASMA (Parallel Linear Algebra Software for Multicore Architectures) is a dense linear algebra package at the forefront of multicore computing. PLASMA is designed to deliver the highest possible performance from a system with multiple sockets of multicore processors. PLASMA achieves this objective by combining state-of-the-art solutions in parallel algorithms, scheduling and software engineering. Currently, PLASMA offers an array of routines for solving linear systems of equations and least square problems, and progress is being made on eigenvalue and singular value problems as well as multi-GPU acceleration.

MIXED-PRECISION ALGORITHMS

PLASMA includes mixed-precision routines to exploit the speed advantage of single precision arithmetic, while delivering double precision accuracy. This is achieved by using single precision factorization and double precision iterative refinement.

TALL AND SKINNY FACTORIZATION

PLASMA delivers a fast routine for QR factorization of tall (and narrow) matrices. This routine exposes a very high level of parallelism by splitting the matrix along the vertical dimension and employing tree reduction patterns. A routine for LQ factorization is coming early 2011.

IN-PLACE LAYOUT CONVERSION

PLASMA contains parallel and cache efficient routines for column-major to row-major conversion and for conversion from LAPACK layout to layout by tiles (PLASMA's native matrix layout).

COMPLETE SET OF BLAS 3

PLASMA delivers the full set of multi-threaded Level 3 BLAS routines, which often outperform multi-threaded vendor implementations, especially on a large number of cores (32 and above).

EXPLICIT MATRIX INVERSION

PLASMA provides a highly parallel routine for explicitly forming an inverse of a symmetric positive definite matrix using Cholesky factorization. The implementation relies on pipelining and dynamic scheduling of the different stages of this computation. A routine for non-symmetric matrices is coming early 2011.

PLASMA 2.3.0 FUNCTIONALITY

- Linear Systems and Least Squares: LU, Cholesky, QR, LQ
- Mixed-Precision Linear Systems: LU, Cholesky, QR
- Explicit Matrix Inversion: Cholesky
- Tall and Skinny Matrix Factorization: QR
- Complete Set of Level 3 BLAS
- In-Place Matrix Layout Conversions

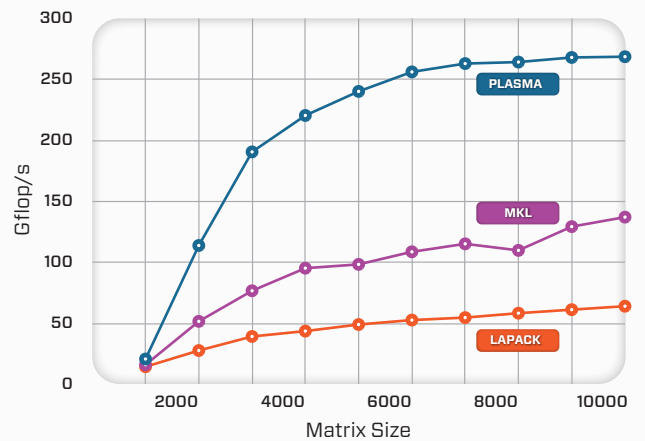
PLASMA 2.3.0 TECHNOLOGY

- Thread Safety
- QUARK Dynamic Scheduler
- Linux, MS Windows, Mac OS, AIX

CURRENT RESEARCH

- GPU Acceleration in CUDA and OpenCL (Including Multiple GPU Acceleration)
- Full Support for Nvidia Fermi Architecture
- Singular Value Problems
- Eigenvalue Problems

QR factorization in double precision



48-core, 2.8 GHz AMD Istanbul System
Double precision peak is 538 Gflop/s

A COLLABORATION OF



WITH SUPPORT FROM



SPONSORED BY

