
COLLEGE OF EDUCATION
ICL
MathWorks

What Next?

- Sven's notes
- Workshop report
- Follow on meeting?
  - When

Sca/LAPACK survey

- UTK/UCB/UCD NSF/DOE effort
- Survey
  - https://www.surveymonkey.com/r/2016DenseLinearAlgebra
- Results
  - https://www.surveymonkey.com/results/SM-J68KNV8Q/

2016 Dense Linear Algebra Software Packages Survey

Summary

About the Survey

The 2016 Dense Linear Algebra Software Packages was administered from January 1st 2016 to April 12 2016. 222 respondents answered the survey. The survey was advertised directly to the Linear Algebra community via our LAPACK/ScaLAPACK forum, NA Digest and we also directly contacted vendors and linear algebra experts. The respondents were for 74% researcher or scientist, for 25% were Principal Investigators and 25% Software maintainer or System administrator.

The goal of the survey was to get the Linear algebra community opinion and provide input on dense linear algebra software packages, in particular LAPACK, ScaLAPACK, PLASMA and MAGMA. The ultimate purpose of the survey was to improve these libraries to benefit our user community. The survey would allow the team to prioritize the many possible improvements that could be done.We also ask input from users accessing these libraries via 3rd party interfaces, for example MATLAB, Intel's MKL, NumPy, AMD's ACML and others.

The survey was composed of six parts:

1. A general section about user's applications and their needs.
2. Specific questions about your LAPACK uses
3. Specific questions about your ScaLAPACK uses,
4. Specific questions about your PLASMA uses,.
5. Specific questions about your MAGMA uses.
6. An open section for any additional comments

Survey link: https://www.surveymonkey.com/r/2016DenseLinearAlgebra
Survey response: https://www.surveymonkey.com/results/SM-J68KNV8Q/summary

Batched

Interface of various batched GEMMs versus GEMM from Sarah's slide

Argument	Description	BLAS sgemm	magma_sgemm_batched	NVidia cublasSgemmBatched	UTK sgemm_batch	Intel MKL sgemm_batch
HANDLE	handle to the cuBLAS library context	--	--	cublasHandle_t	--	--
TRANSA	op(A)	char	char	char	char *	char *
TRANSB	op(B)	char	char	char	char *	char *
M	rows of op(A)/C	int	int	int	int *	int *
N	columns of op(B)/C	int	int	int	int *	int *
K	columns of op(A)/rows of op(B)	int	int	int	int *	int *
ALPHA	alpha	float	float	float *	float *	float *
A	input matrix	float *	float **	float **	float **	float **
LDA	leading dimension of A	int	int	int	int *	int *
B	input matrix	float *	float **	float **	float **	float **
LDB	leading dimension of B	int	int	int	int *	int *
BETA	beta	int	float	float *	float *	float *
C	input/output matrix	float *	float **	float **	float **	float **
LDC	leading dimension of C	int	int	int	int *	int *
BATCHCOUNT	number of matrices	--	int	int	int	--
QUEUE	queue to execute in	--	magma_queue_t	--	--	--
BATCH_OPTS	style for batched (fixed or variable)	--	--	--	enum	--
INFO	error handling	--	--	--	int *	--
GROUP_COUNT	number of groups	--	--	--	--	int
GROUP_SIZES	number of matrices in each group	--	--	--	--	int *

Reproducibility

- Jim's - 7 design goals
- Bill's follow on
  - Data distributions critical
- Mike's - Replication
  - "carrot and stick"

Reduced Precision

- RGEMM_xx and CGEMM_xx
  - Or some variant
  - IGEMM_xx

With Apologies to Gary Larson…