SIAM Conference on
Parallel Processing for Scientific Computing
Sir Francis Drake Hotel
San Francisco, California
February 22-24, 2006

Developments with LAPACK and ScaLAPACK on Today's and Tomorrow's Systems

Jack Dongarra
University of Tennessee
and
Oak Ridge National Laboratory

Also hear Jim Demmel's talk at 2:30 today MS47 Carmel room

Participants

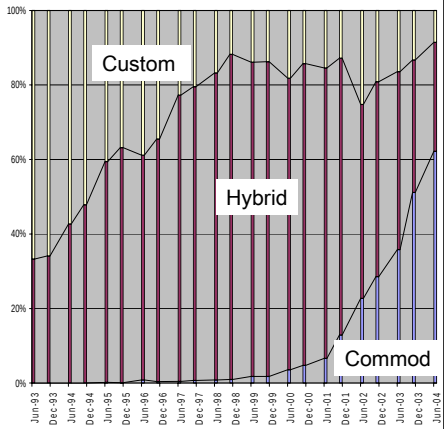
- U Tennessee, Knoxville
  - Jack Dongarra, Julien Langou, Julie Langou, Piotr Luszczek, Jakub Kurzak, Stan Tomov, Remi Delmas, Peng Du
- UC Berkeley:
  - Jim Demmel, Ming Gu, W. Kahan, Beresford Parlett, Xiaoye Li, Osni Marques, Christof Voemel, David Bindel, Yozo Hida, Jason Riedy, Jianlin Xia, Jiang Zhu, undergrads…
- Other Academic Institutions
  - UT Austin, UC Davis, Florida IT, U Kansas, U Maryland, North Carolina SU, San Jose SU, UC Santa Barbara
  - TU Berlin, FU Hagen, U Madrid, U Manchester, U Umeå, U Wuppertal, U Zagreb
- Research Institutions
  - CERFACS, LBL
- Industrial Partners
  - Cray, HP, Intel, MathWorks, NAG, SGI, Microsoft

Architecture/Systems Continuum

Tightly Coupled

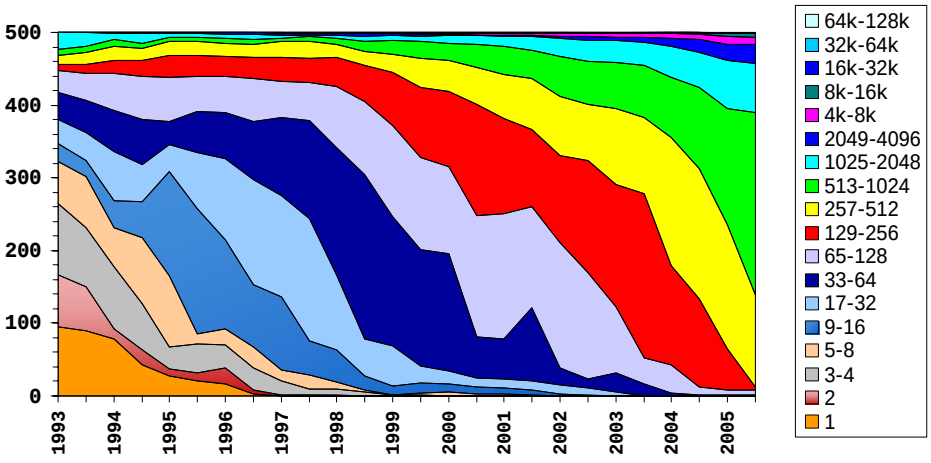
- Custom processor with custom interconnect
  - Cray X1
  - NEC SX-8
  - IBM Regatta
  - IBM Blue Gene/L
- Commodity processor with custom interconnect
  - SGI Altix
    - Intel Itanium 2
  - Cray XT3, XD1
    - AMD Opteron
- Commodity processor with commodity interconnect
  - Clusters
    - Pentium, Itanium, Opteron, Alpha
    - GigE, Infiniband, Myrinet, Quadrics
  - NEC TX7
  - IBM eServer
  - Dawning

Loosely Coupled



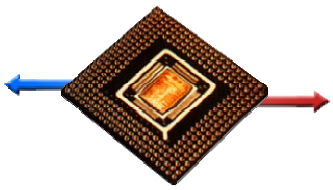
33 4

Parallelism in the Top500



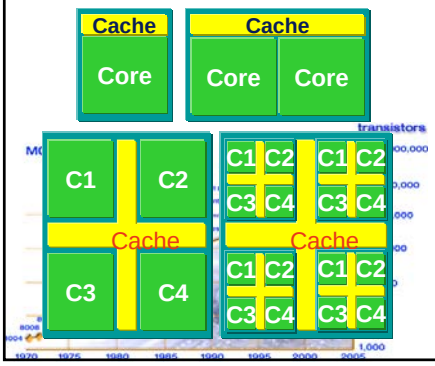
33 5

Increasing CPU Performance: A Delicate Balancing Act



Lower Voltage

Increase Clock Rate & Transistor Density



We have seen increasing number of gates on a chip and increasing clock speed.

Heat becoming an unmanageable problem, Intel Processors > 100 Watts

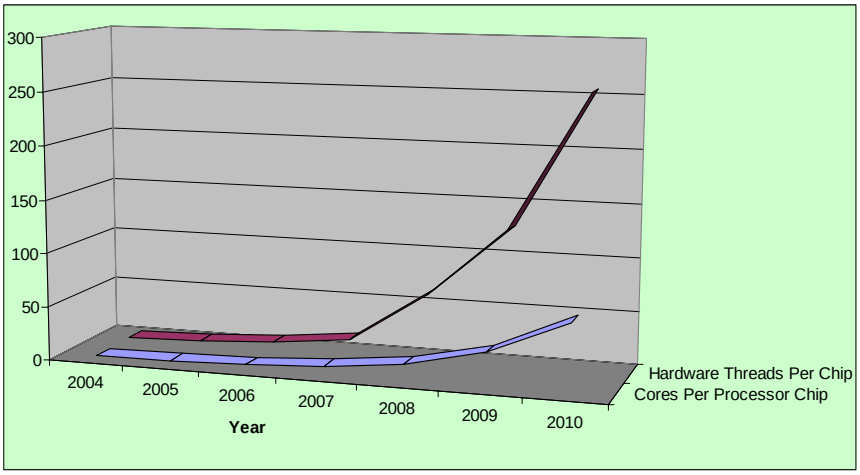
We will not see the dramatic increases in clock speeds in the future.

However, the number of gates on a chip will continue to increase.

6

CPU Desktop Trends – Change is Coming

- Relative processing power will continue to double every 18 months
- 256 logical processors per chip in late 2010



7

Commodity Processor Trends

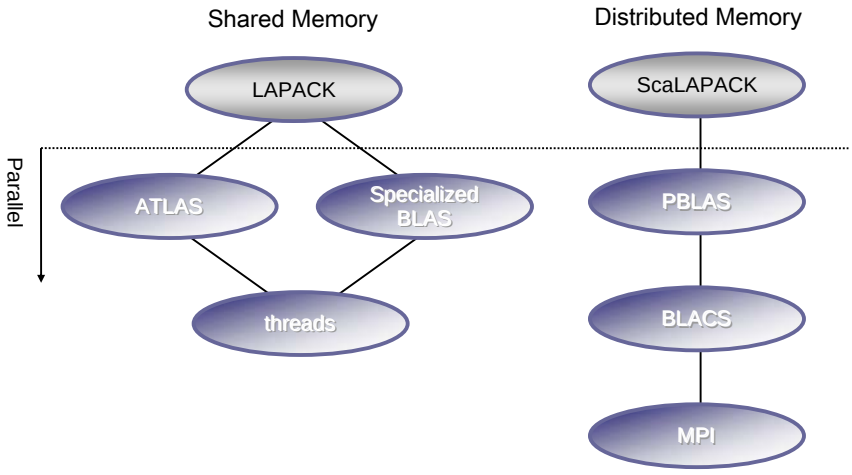
Bandwidth/Latency is the Critical Issue, not FLOPS

Got Bandwidth?

	Annual increase	Typical value in 2006
Single-chip floating-point performance	59%	4 GFLOP/s
Front-side bus bandwidth	23%	1 GWord/s = 0.25 word/flop
DRAM latency	(5.5%)	70 ns = 280 FP ops = 70 loads

33 Source: *Getting Up to Speed: The Future of Supercomputing*, National Research Council, 222 pages, 2004, National Academies Press, Washington DC, ISBN 0-309-09502-6. 8

Parallelism in LAPACK / ScaLAPACK



33

LAPACK and ScaLAPACK Futures

- Widely used dense and banded linear algebra libraries
  - Used in vendor libraries from Cray, Fujitsu, HP, IBM, Intel, NEC, SGI
  - In Matlab (thanks to tuning…), NAG, PETSc,…
  - over 56M web hits at www.netlib.org
    - LAPACK, ScaLAPACK, CLAPACK, LAPACK95
- NSF grant for new, improved releases
  - Joint with Jim Demmel, many others
  - Community effort (academic and industry)
- Next major release scheduled in 2007
- See Jim Demmel's talk at 2:30 today MS47 Carmel room

Goals (highlights)

- Putting more of LAPACK into ScaLAPACK
  - Lots of routines not yet parallelized
- New functionality
  - Ex: Updating/downdating of factorizations
- Improving ease of use
  - Life after F77?, Binding to other languages
  - Callable from Matlab
- Automatic Performance Tuning
  - Over 1300 calls to ILAENV() to get tuning parameters
- New Algorithms
  - Some faster, some more accurate, some new

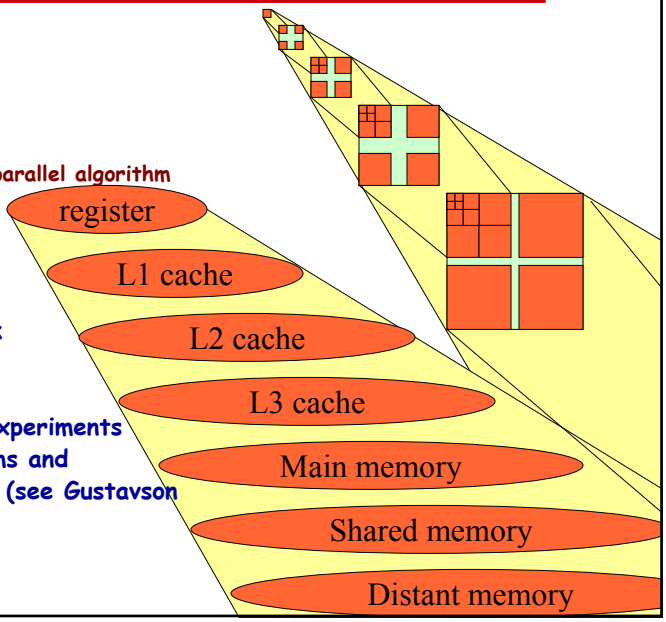
Faster: λ's and σ's

- **Nonsymmetric eigenproblem**
  - Incorporate SIAM Prize winning work of Byers / Braman / Mathias on faster HQR
  - Up to 10x faster for large enough problems
- **Symmetric eigenproblem and SVD**
  - Reduce from dense to narrow band
    - Incorporate work of Bischof/Lang, Howell/Fulton
    - Move work from BLAS2 to BLAS3
  - Narrow band (tri/bidiagonal) problem
    - Incorporate MRRR algorithm of Parlett/Dhillon
    - Voemel, Marques, Willems

33 12

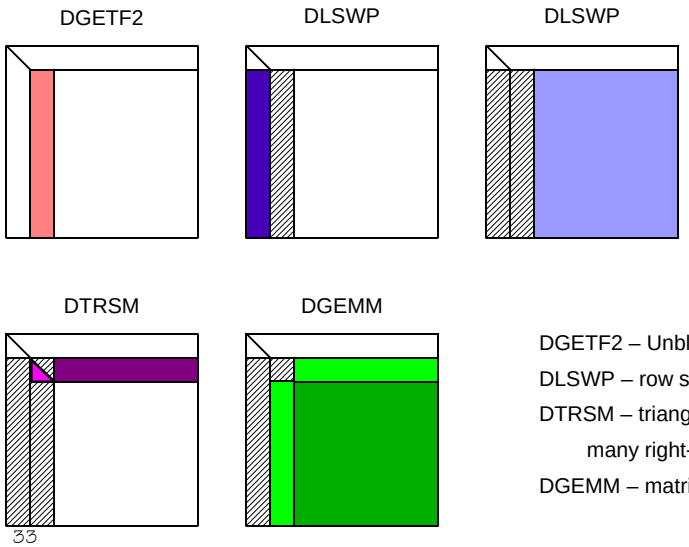
Recursive/Fractal architectures
Recursive/Fractal data layout & Recursive algorithms

- Enable:
  - Register blocking
  - L1cache bocking
  - L2 cache blocking
  - Natural layout for parallel algorithm
- Close to the 2D block cyclic distribution
- Proven efficient by experiments on recursive algorithms and recursive data layout (see Gustavson et al.)



33

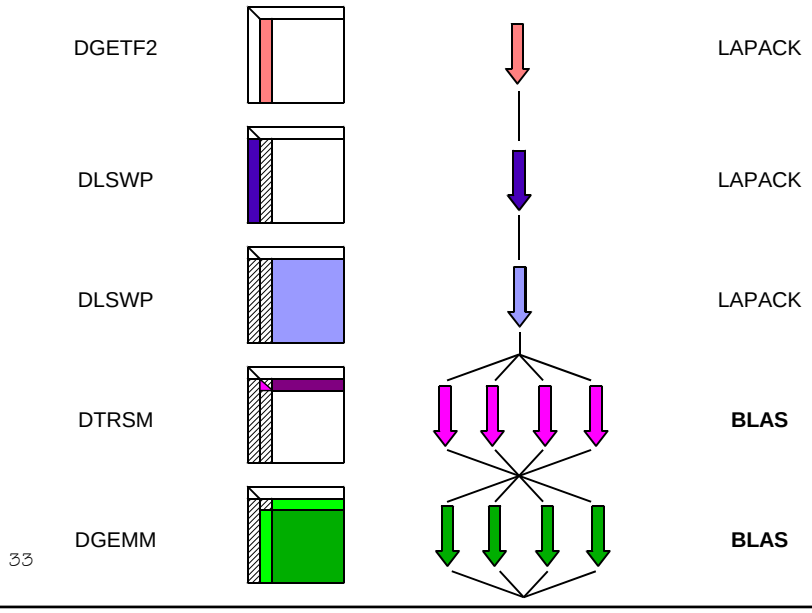
Right-Looking LU factorization (LAPACK)



DGETF2 – Unblocked LU
DLSWP – row swaps
DTRSM – triangular solve with many right-hand sides
DGEMM – matrix-matrix multiply

33 14

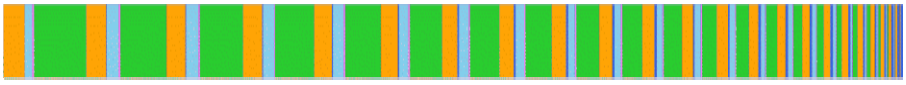
Steps in the LAPACK LU



33 15

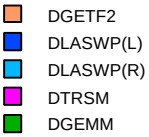
LU Timing Profile

LAPACK + BLAS threads



Time for each component

1D decomposition and SGI Origin



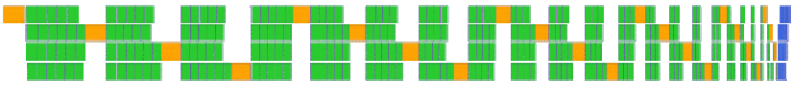
33

LU Timing Profile

LAPACK + BLAS threads

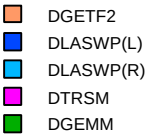
Time for each component

Threads – no lookahead



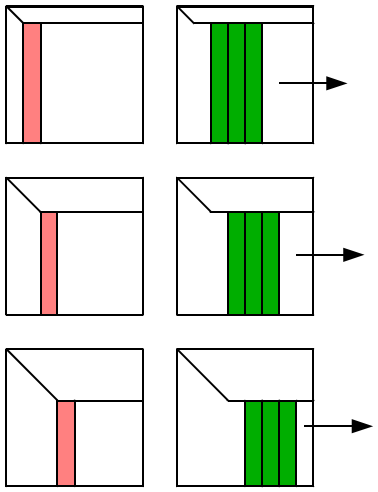
In this case the performance difference comes from parallelizing row exchanges (DLASWP) and threads in the LU algorithm.

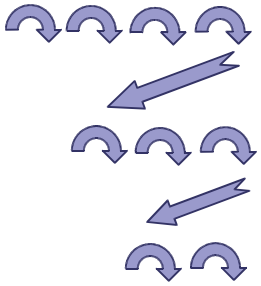
1D decomposition and SGI Origin



33

Right-Looking LU Factorization

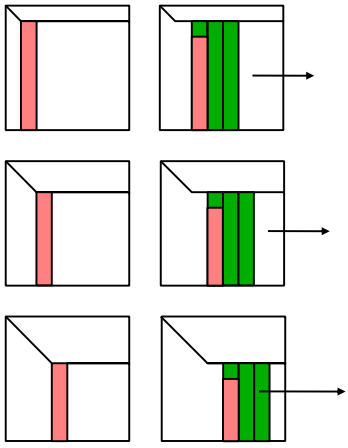


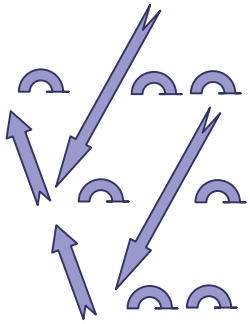


33

18

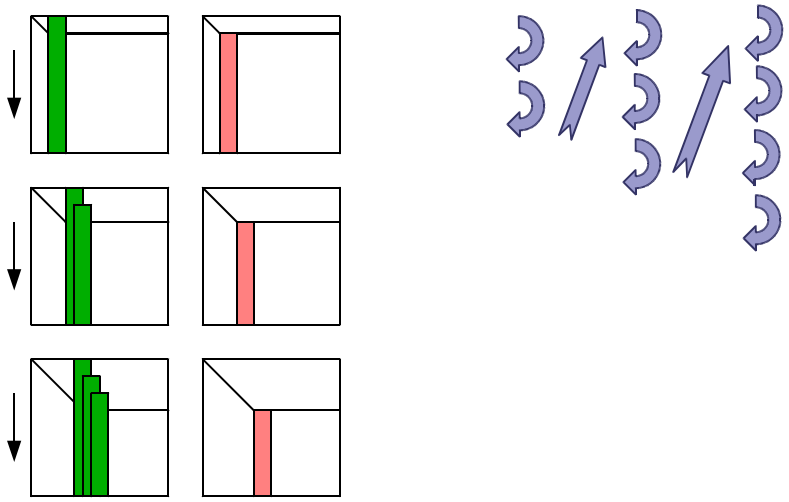
Right-Looking LU with a Lookahead





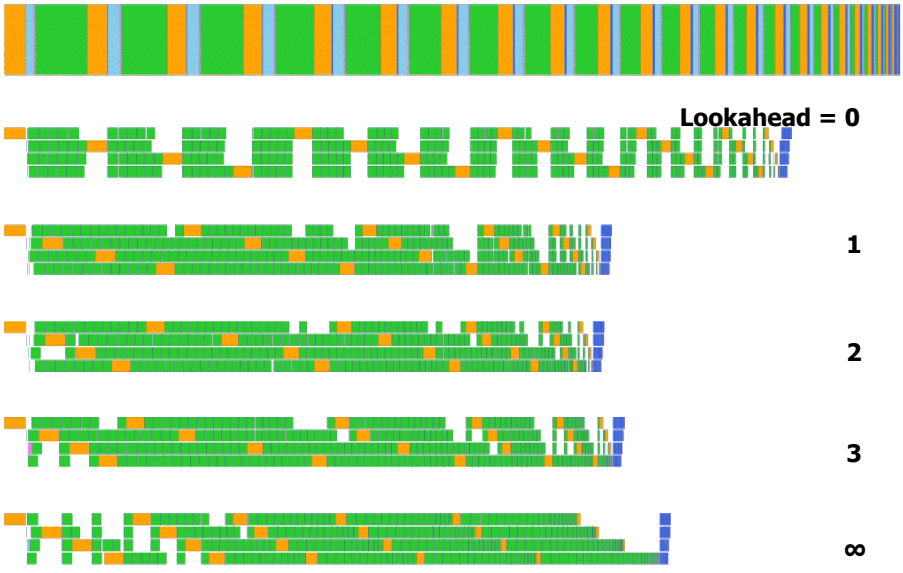
33

Left-Looking LU factorization



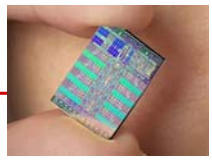
33 20

Pivot Rearrangement and Lookahead
4 Processor runs

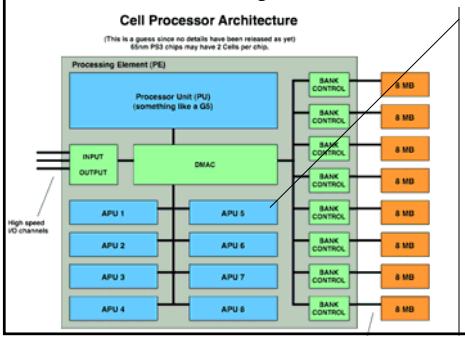


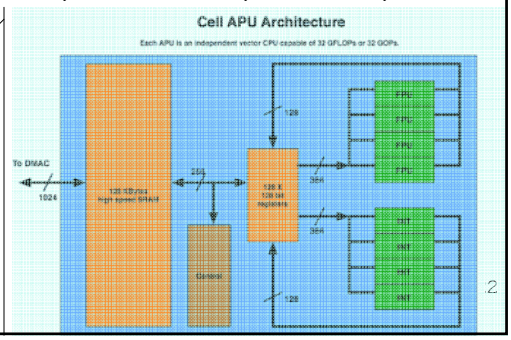
21

Things to Watch: PlayStation 3



- The PlayStation 3's CPU based on a chip codenamed "Cell"
- Each Cell contains 8 APUs.
  - An APU is a self contained vector processor which acts independently from the others.
  - 4 floating point units capable of a total of 32 Gflop/s (8 Gflop/s each)
  - 256 Gflop/s peak! 32 bit floating point; 64 bit floating point at 25 Gflop/s.
  - IEEE format, but only rounds toward zero in 32 bit, overflow set to largest
    - According to IBM, the SPE's double precision unit is fully IEEE854 compliant.





2

32 and 64 Bit Floating Point Arithmetic

- Use 32 bit floating point whenever possible and resort to 64 bit floating point when needed to refine solution.
- Iterative refinement for dense systems can work this way.

Solve $Ax = b$ in lower precision,
save the factorization ($L*U = A*P$); $O(n^3)$
Compute in higher precision $r = b - A*x$; $O(n^2)$
Requires the original data A (stored in high precision)
Solve $Az = r$; using the lower precision factorization; $O(n^2)$
Update solution $x_+ = x + z$ using high precision; $O(n)$
Iterate until converged.

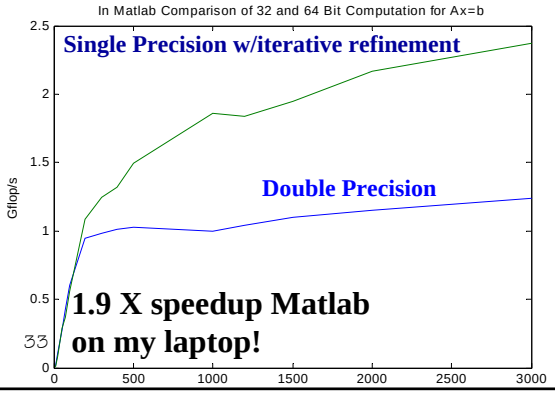
Requires extra storage, total is 1.5 times normal;
$O(n^3)$ work is done in lower precision
$O(n^2)$ work is done in high precision
In the best case doubles number of digits per iteration
Problem if the matrix is ill-conditioned in sp; $O(10^8)$

33

23

Another Look at Iterative Refinement

- On Cell processor, single precision is at 256 Gflop/s and double precision is at 25 Gflop/s.
- On a Pentium; using SSE2, single precision can perform 4 floating point operations per cycle and in double precision 2 floating point operations per cycle.
- Reduced memory traffic (factor on sp data)



Cluster w/3.2 GHz Xeons w/ScaLAPACK

#procs	n	Speedup	#steps
2	2000	1.52	4
2	4000	1.60	5
2	6000	1.66	4
2	8000	1.65	5
4	4000	1.66	4
4	8000	1.78	6
4	12000	1.69	6
4	16000	1.69	5
8	8000	1.64	5
8	16000	1.78	6
8	24000	1.83	5
16	16000	1.92	5
16	32000	1.77	18
32	32000	1.84	12

33

Refinement Technique Using Single/Double Precision

- **Linear Systems**
  - LU (dense and sparse)
  - Cholesky
  - QR Factorization
- **Eigenvalue**
  - Symmetric eigenvalue problem
  - SVD
  - Same idea as with dense systems,
    - Reduce to tridiagonal/bi-diagonal in lower precision, retain original data and improve with iterative technique using the lower precision to solve systems and use higher precision to calculate residual with original data.
    - $O(n^2)$ per value/vector
- **Iterative Linear System**
  - Relaxed GMRES
  - Inner/outer scheme

33 LAPACK Working Note in progress 25

Summary

- Better / Faster Numerics
  - MRRR sym λ & SVD
  - HQR, QZ, reductions, packed
- Expanded Content
  - ScaLAPACK mirror LAPACK
- Extended precision version
  - Variable precision, user controlled
- Callable from Matlab
  - From Matlab invoke LAPACK routine
- Recursive data structures
  - For Performance
- Automate Performance Tuning
- Improve ease of use
- Better Maintenance and Support
- Involve the Community
  - Open source effort

33 26

Collaborators / Support

- U Tennessee, Knoxville
  - Julien Langou, Julie Langou, Piotr Luszczek, Jakub Kurzak, Stan Tomov, Remi Delmas, Peng Du
- UC Berkeley:
  - Jim Demmel, Ming Gu, W. Kahan, Beresford Parlett, Xiaoye Li, Osni Marques, Christof Voemel, David Bindel, Yozo Hida, Jason Riedy, Jianlin Xia, Jiang Zhu, undergrads…
- Other Academic Institutions
  - UT Austin, UC Davis, Florida IT, U Kansas, U Maryland, North Carolina SU, San Jose SU, UC Santa Barbara, TU Berlin, FU Hagen, U Madrid, U Manchester, U Umeå,U Wuppertal, U Zagreb
- Research Institutions
  - CERFACS, LBL
- Industrial Partners
  - Cray, HP, Intel, IBM, MathWorks, NAG, SGI, Microsoft

33