Argonne NATIONAL LABORATORY

Accurate Numerical Simulations of Chemical Phenomena Involved in Energy Production and Storage Using MADNESS

Álvaro Vázquez-Mayagoitia (ALCF), Jeff Hammond (ALCF), Nick Romero (ALCF), Lee Killough ([ALCF]), Kevin Stock (ALCF,OSU), Robert Harrison (ORNL->SBU), Scott Thornton (UTK), George Fann (ORNL), Ed Valeev (VT), Justus Calvin (VT), and the rest of the MADNESS team.

MADNESS: Multiresolution Adaptive Numerical Environment for Scientific Simulation

✓A general purpose numerical framework for reliable and fast scientific simulations:
Chemistry, nuclear physics, atomic physics, material science, climate, fusion, …

MADNESS IS NOT A QUANTUM CHEMISTRY CODE!!!

✓A general purpose parallel programming environment designed for the petascale and beyond
– Standard C++ with concepts from Cilk, Charm++ and Chapel
– Compatible by design with existing applications
– Runs on Titan/OLCF, Mira/ALCF, …, Macbook, …, Amazon EC2

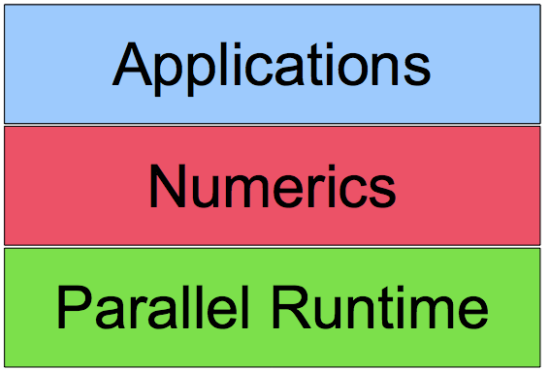
World: parallel runtime used by MADNESS but also TiledArray (VT group) – implements thread pool and queue, atomics, active-messages, futures, distributed containers.

Tensor: Core numerical functionality for multidimensional arrays and contractions – mTxm is the low-level kernel.

MRA: Multiresolution analysis – higher-level mathematical operations

Apps: where the PDEs live – this project focuses on moldft, which is the molecular density functional code.

The moldft source is ~3000 lines…



E.g., with guaranteed precision of 1e-6 form a numerical representation of a Gaussian in the cube $[-20,20]^3$, solve Poisson's equation, and plot the resulting potential (all running in parallel with threads+MPI)

Let

$$\Omega = [-20, 20]^3$$

$$\epsilon = 1e-6$$

$$g = x \rightarrow \exp\left(-\left(x_0^2 + x_1^2 + x_2^2\right)\right) * \pi^{-1.5}$$

In

$$f = \mathcal{F}\, g$$

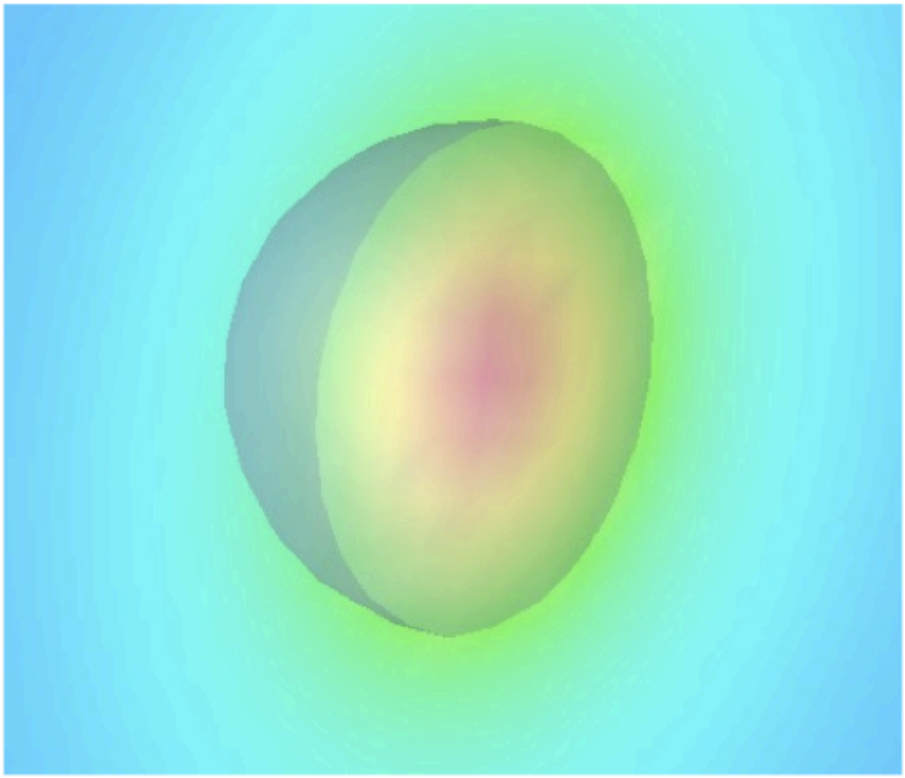
$$u = \nabla^{-2}\left(-4 * \pi * f\right)$$

print "norm of f", $\langle f \rangle$, "energy", $\langle f | u \rangle * 0.5$

plot u

End

output: norm of f 1.00000000e+00 energy 3.98920526e-01



There are only two lines doing real work. First the Gaussian (g) is projected into the adaptive basis to the default precision. Second, the Green's function is applied. The exact results are norm=1.0 and energy=0.3989422804.

Quantum Mechanics

We can write in an integral form the solution for an electronic wavefunction.

$$\hat{H}\varphi(r) = (\tfrac{1}{2}\nabla^2 + \hat{V})\varphi(r) = \varepsilon\varphi(r)$$

Integral equation from:

$$\varphi(r) = -2(\nabla^2 + 2\varepsilon)^{-1}\hat{V}\varphi(r)$$

– Higher accuracy achievable in integral form
– Correct asymptotics by design
– Computationally efficient: LSR and local refinement
– MRA provides fast algorithms with guaranteed precision

Matrix representation leads to dense eigenvalue problem

$$\Psi(r_1, r_2, \ldots) = \frac{1}{\sqrt{N!}} \det\left|\varphi_1(r_1)\varphi_2(r_2)\ldots\right|$$

$$\hat{H}\varphi(r) = (\tfrac{1}{2}\nabla^2 + \hat{V})\varphi(r) = \varepsilon\varphi(r)$$

$$\int dr \varphi_i^*(r)\hat{H}\varphi_j(r) = H_{ij} \qquad \int dr \varphi_i^*(r)\varphi_j(r) = S_{ij}$$

$$(\tilde{H} - \varepsilon\tilde{I})\tilde{S} = 0$$

size of the problem $\{\varphi\}$ by $\{\varphi\}$

MADNESS solves a very small basis set problem for the initial guess, so we still need still, but not like MPQC or other traditional codes. On the other hand, getting all the orbitals is critical for QMBT methods, as in MPQC, NWChem, AQ/CTF.

MADNESS: Summary of developments for ESP

- 2010: Initial port to Blue Gene/P.
- 2010 – present: Raf Schietekat ports Intel TBB to PowerPC then Blue Gene/P then Blue Gene/Q.
- 2011 – present: Alvaro develops molecular DFT code.
- 2012 – 2013: Alvaro modernizes dense linear algebra with Eigen3 and Elemental.
- Summer 2012: Kevin Stock developed highly optimized mTxm kernel for Blue Gene/Q. QPX vector intrinsic code is auto-generated just as for SSE, AVX, MIC and BGP (Summer 2010 at ALCF).
- 2012 – present: George Fann scaling work for the nuclear physics code on Mira.

Improving numerical library usage in MADNESS

NAME	Purpose
GESVD	Singular value decomposition (SVD)
GESV	Solution to: $Ax = B$
GELSS	Minimum norm solution to $\|B-Ax\|$
SYGV	Eigenvectors and eigen values $Ax=\lambda Bx$
SYEV	Compute all eigenvalues of a symmetric mat A
POTRF	Cholesky factorization

Eigen3 - Serial Dense Linear Algebra

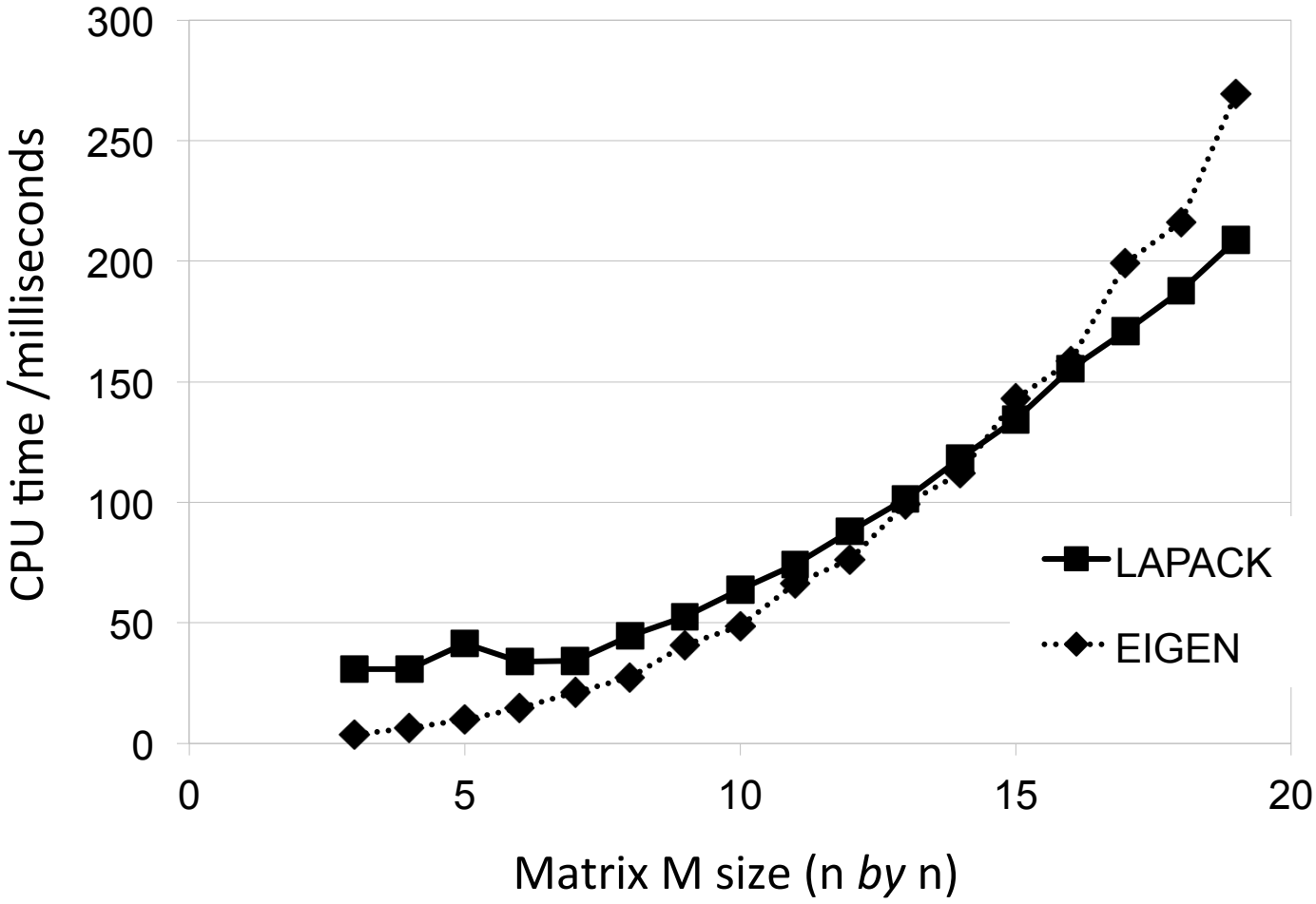
- Pure C++ template library – headers only, no binaries.
- Included in all Linux distributions.
- Open source license: LGPL3+
- Users: Google, Qt-KDE, Shogun, Avogadro, FEM codes, …
- Compilers: – GCC, MSVC, Intel ICC, Clang/LLVM
- Processors: – x86/x86_64, ARM, PowerPC
- Better for small operations (compile-time dimensioning in some cases).
- Works around LAPACK thread-safety problems.

```
typedef Matrix<double,Dynamic,Dynamic> MatrixX;
MatrixX A, B, X;
// init A and B
// solve for A.X=B using LU decomposition
X = A.lu().solve(B);
```

Eigen3 - Serial Dense Linear Algebra

Singular Value Decomposition (SVD)

$$M = U\Sigma V^*$$

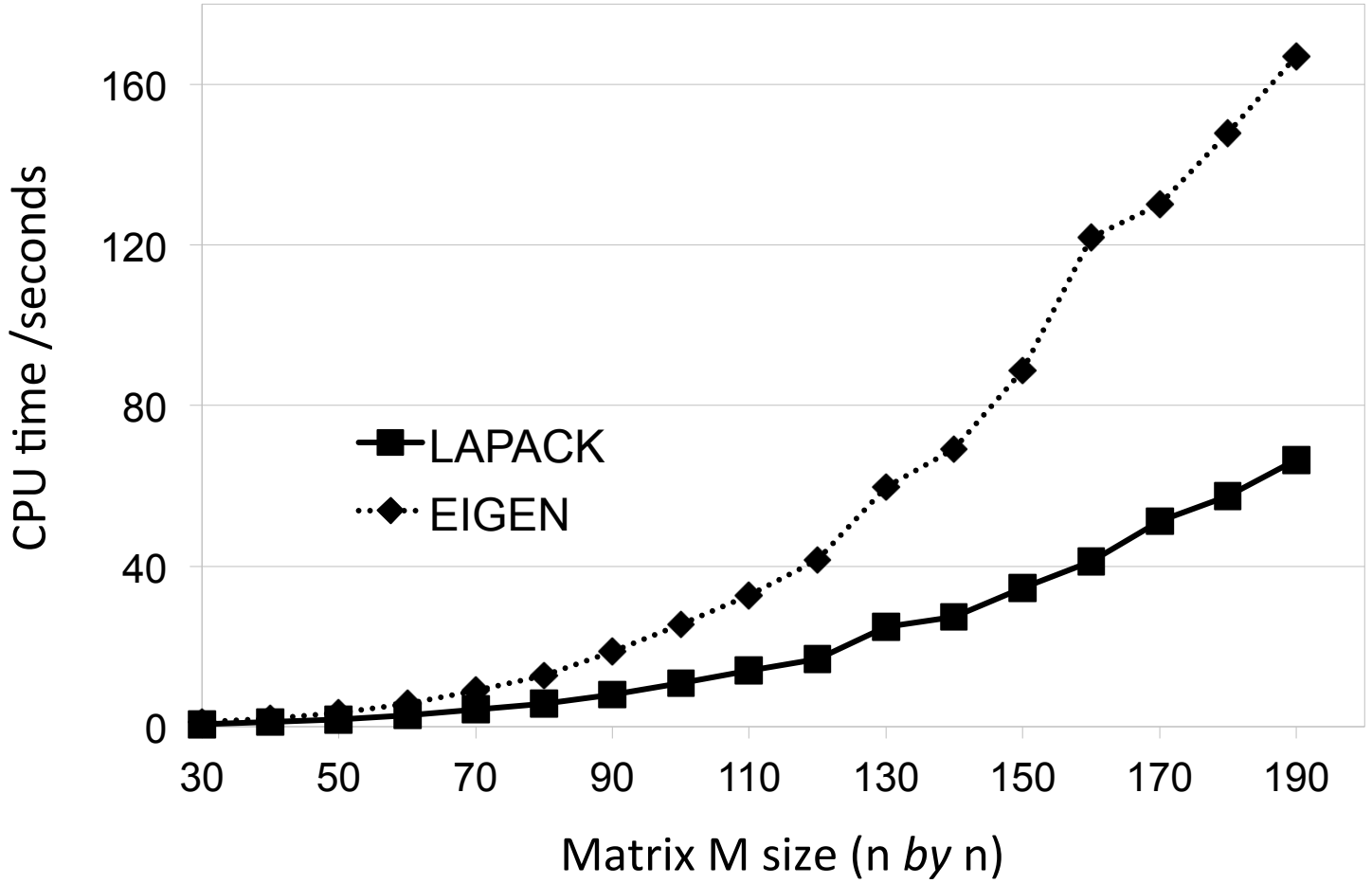


*Serial execution

Eigen3 - Serial Dense Linear Algebra

Singular Value Decomposition (SVD)

$$M = U\Sigma V^*$$



*Serial execution

Elemental- Parallel Dense Linear Algebra

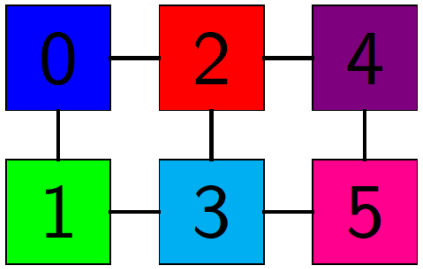
- Developed by Jack Poulson (Georgia Tech) with collaborators at the University of Texas, RWTH, Argonne, …
- Modern C++ design with opaque data containers – very easy to use interface compared to ScaLAPACK.
- Designed to make heavy use of MPI collectives rather than BLACS (based upon p2p); topology-aware on Blue Gene.
- Thread-safe and multithreaded using OpenMP in matrix operations and Pthreads in RWTH's MRRR tridiagonal solver.
- Open-source license: new BSD.
- Used by *MADNESS*, *Qbox*, *MPQC*, PETSc, PSP, Clique, etc.
- NSF-funded collaboration between Elemental and chemists.

Elemental- Parallel Dense Linear Algebra

The set processes is mapped in a 2D Grid.

e.g.

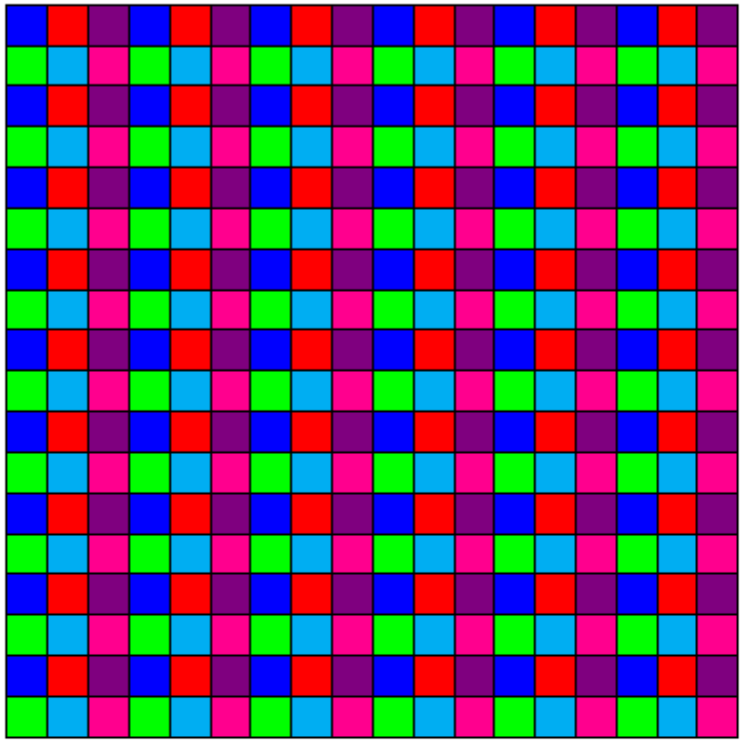
Grid 2 *by* 3



6 MPI processes

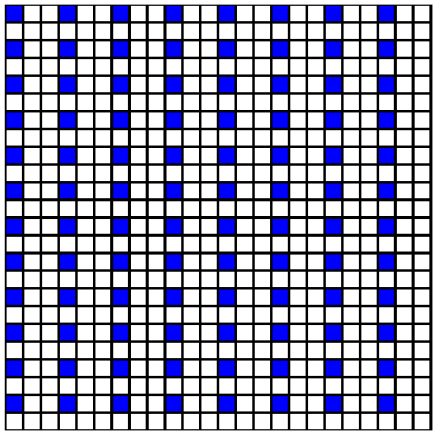
The data is distributed in the grid of processes.

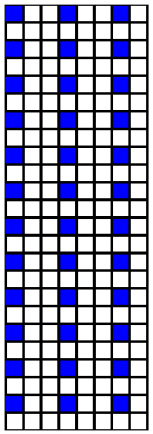
DistMatrix 18 *by* 18

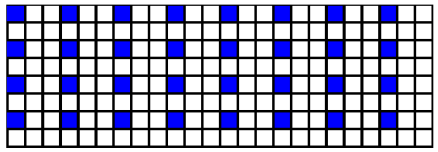


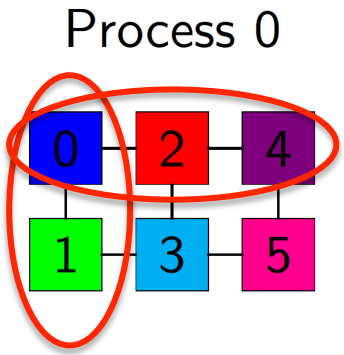
Elemental- Parallel Dense Linear Algebra

Example: Matrix Matrix Multiplication



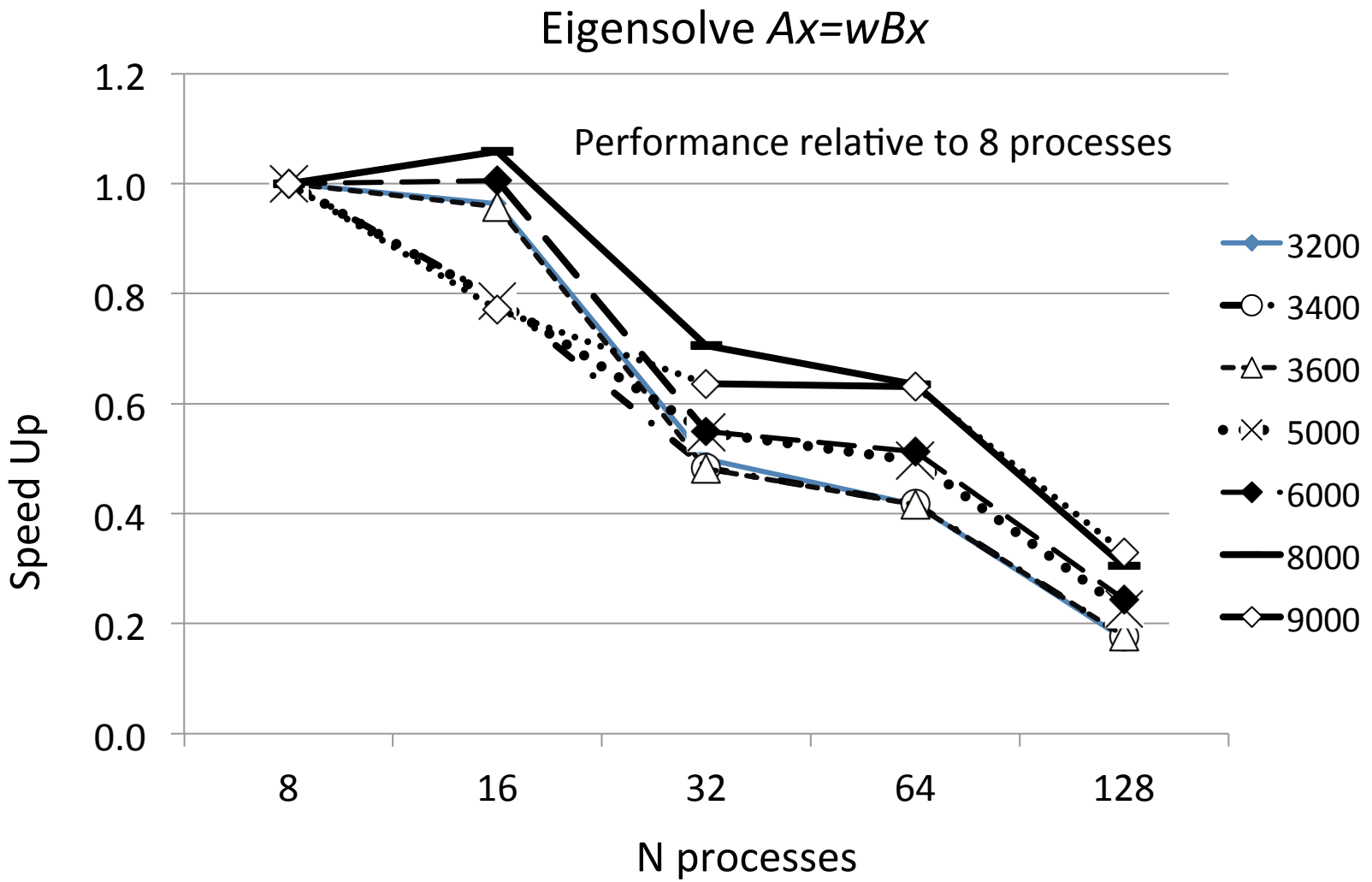






AllGather operations are performed by rows, then by columns of Grid

Elemental- Parallel Dense Linear Algebra



1 MPI process = 1 node

Elemental- Parallel Dense Linear Algebra

	Elemental	Eigen3	LAPACK
Water*70 (Full SCF) 16 nodes	1272.4 6.7%	1387.7 -2.2%	1357.3
Water*213 (maxiter=2) 256 nodes	2753.5 15.2%	3546.9 -10.6%	3171.4
eigen solution time per iteration	1.5 s		~700 s

- Eigen3 has a similar performance than LAPACK
- The potential of Elemental is clear when we deal with big systems (>1K size matrices).

Summary: Dense Linear Algebra

- Interfacing **Eigen3** or **Elemental** to a C++ code is not difficult.
- **Eigen3** is an excellent competitor of LAPACK for *small* matrices in BG/Q, Eigen3 could improved with specific tuning for PowerPC. (Future versions will be multi-threaded).
- **Elemental** in combination with LAPACK has shown an outstanding performance Blue Gene/Q.
- **Elemental** is a modern replacement for ScaLAPACK. Quantum chemistry codes, particularly those that use hybrid programming models (i.e. MPI+Threads) or use modern programming languages like C++ will benefit.

Other contributions of this ESP

- Dozens of PMRs filed for XLC; motivated fundamental changes in the IBM XL OpenMP runtime; both MPQC and MADNESS found thread safety issues in MPI.
- Port of Intel TBB to Blue Gene (**BlueTBB**) is motivated by MADNESS but broadly applicable (e.g. Deal.II).
- MPQC has been ported but not a lot of work yet due to Curt Janssen moving to Google; MPQC ported almost immediately on VEAS@IBM; VT group is working on additional developments.
- Ongoing improvements to Elemental driven by Blue Gene and MADNESS usage.
- **Aquarius** (AQ) and **Cyclops Tensor Framework** (CTF) have been developed as part of this allocation since they are like-minded efforts that provide unique features (QMBT).

Density Functional Theory

DFT is an exact theory, and the energy of the ground state is given precisely if the correct functional is known and minimized.
Electron density is a fundamental variable.

- ✓ DFT can compute energies of chemical systems

$$E = F[\rho_0] - \int dr \rho_o(r) \upsilon_{ext}(r) + \tfrac{1}{2} \sum_{AB} \frac{Z_A Z_B}{R_{AB}}$$

Exchange Correlation Functional

$$E_{ele} = T_0[\rho_0] + \int dr \frac{\rho_o(r')}{|r - r'|} \rho_o(r) + V_{xc}[\rho_0] + \int dr \rho_o(r) \upsilon_{ext}(r)$$

$$\upsilon_{coul}[\rho] = \frac{\delta E_{Coul}[\rho]}{\delta \rho} = \frac{\rho(r')}{|r - r'|} \qquad \upsilon_{coul}[\rho] = \frac{\delta E_{XC}[\rho]}{\delta \rho}$$