Large Scale Plane Wave Pseudopotential Density Functional Theory Calculations on GPU Clusters

Long Wang[1] , Weile Jia[1], Xuebin Chi[1], Yue Wu[2], Weiguo Gao[2], Lin-Wang, Wang[3]

1. Supercomputing Center , Chinese Computer Network Information Center, Chinese Academy of Science
2. School of Mathematical Sciences, Fudan University
3. Material Science Division, Lawrence Berkeley National Laboratory

This work is accepted by Supercomputing Conference 2011 .

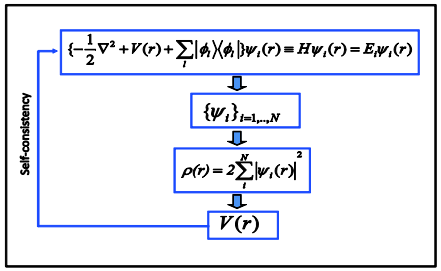
Abstract

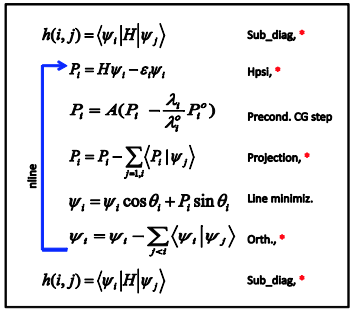
In this poster, we present our implementation of the density functional theory (DFT) plane wave pseudo-potential (PWP) calculation on GPU clusters. This GPU version is developed based on a CPU DFT-PWP code: PEtot. Our test indicates that the GPU version can have a ~10 times speed-up over the CPU version. An analysis of the speed-up and the scaling on the number of CPU/GPU computing units(up to 256) are presented. The success of our speed-up relies on a hybrid reciprocal-space and band-index parallelization scheme. As far as we know, this is the first GPU DFT-PWP code scalable to large number of CPU/GPU computing units.

DFT on GPU clusters

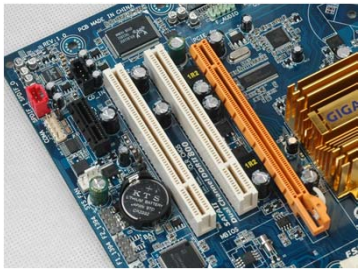
Schrodinger's equation: $[-\frac{1}{2}\nabla^2 + V_{tot}(r)]\psi_i(r) = \varepsilon_i\psi_i(r)$

Basic flow chart and AB-CG method of DFT calculation:

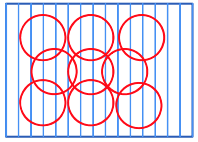


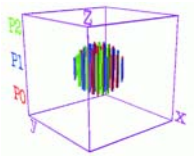


CPU parallelization scheme



Problem on GPU





MPI communication and PCI-E data transfer overwhelme the computation.

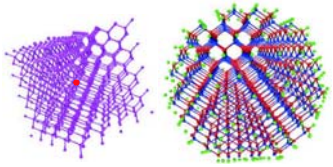
Hybrid parallelization by transposing the wave functions from the band-index parallelization to G-space parallelization.

Totally, GPU PEtot features are:

- **Support both G-space and real-space calculation**
- **Data Structure**
  Transpose wave-function from (16,1) to (1,16)
- **Sphere-to-cube FFT**
  Mapping sphere data to cube data in GPU and do cube FFT
- **MPI communication and GPU computation overlaps**
  Non-block MPI_Alltoall to pipline MPI and GPU computation

Results

Testing systems: left, 512 atoms GaAs bulk system with one As replaced by N;right, a 933 atom CdSe quantum dot passivated by surface H atoms(blue: Cd, red, Se, green H).
~10 times faster than CPU PEtot



If MPI_Alltoall, pzheev and zpotrf can be significantly speed up by the CULA library, then it might be possible to speed up the DFT PWP GPU code by another factor of 10 by scaling the calculation to thousands of CPU/GPU computing units.

Computing units	16	32	64	128	256	256
systems	512-GaAs	512-GaAs	512-GaAs	512-GaAs	512-GaAs	933-CdSe
PEtot (CPU)	842	450	255	152	104	495
PEtot (GPU)	87	49	27	23	17	56
Speed-up (PEtot)	9.7x	9.2x	9.4x	7x	6.1x	8.8x
Total flops (Tflops)	0.59	1.05	1.91	2.24	3.03	5.92
Efficiency	7.1%	6.3%	5.7%	3.3%	2.3%	4.4%

Computing units	16	32	64	128	256	256
systems	512-GaAs	512-GaAs	512-GaAs	512-GaAs	512-GaAs	933-CdSe
Hpsi	7.83	4.58	2.73	1.84	1.52	4.93
MPI_alltoallx2	3.02	2.12	1.49	1.20	1.17	2.05
FFT	2.40	1.23	0.64	0.36	0.19	1.50
Nonlocal proj.	2.00	1.02	0.50	0.23	0.14	1.24
Hpsi_SU	5.1x	4.7x	4.9x	4.0x	3.3x	6.23x
Sub_diag	4.32	2.88	2.32	2.11	2.21	4.63
zheev	1.30	1.28	1.42	1.41	1.76	3.06
Sub_diag_SU	11.7x	9.3x	6.5x	5.5x	4.6x	8.3x
Projection	3.13	1.65	0.91	0.71	0.43	1.49
Projection_SU	15.3x	14.6x	13.4x	8.8x	7.8x	12.7x
Orth.	4.81	2.67	1.55	1.14	0.80	2.77
Orth_SU	16.1x	15.6x	13.7x	12.9x	12.8x	16.44x
Fortran_loopNew	2.15	1.13	0.63	0.38	0.25	0.87
Zpotrf	0.25	0.25	0.25	0.25	0.25	0.62

The Scalability of CPU/GPU PEtot, AB-CG parts scalability