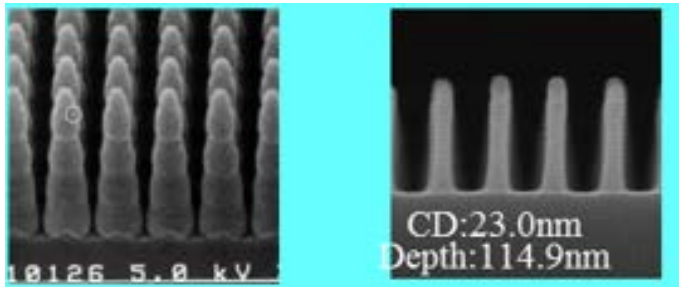
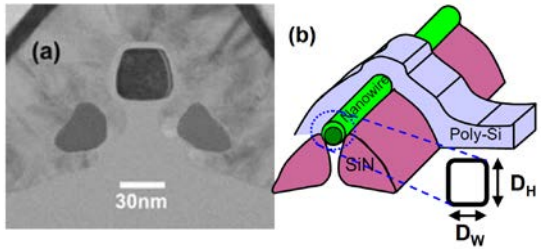
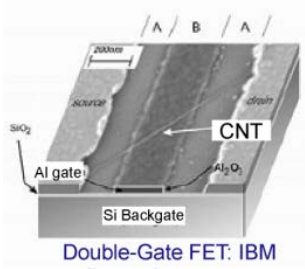
Why Large Scale DFT Calculations?



3Dim Si NAND Memory
Tohoku Univ



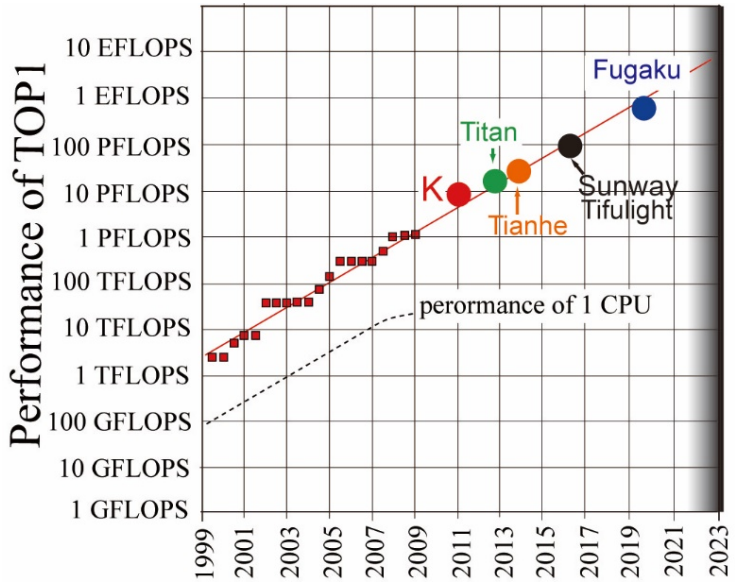
Si-Nanowire FET
TokyoIT, IBM, ,,,



FET with CNT
IBM

Nano-Scale Structure (*Nano-Shape*) Modifies Wavefunctions and then Material Properties

Computics: Fusion of Materials Science and Computer Science



✓ Saturation of performance of a single compute processor

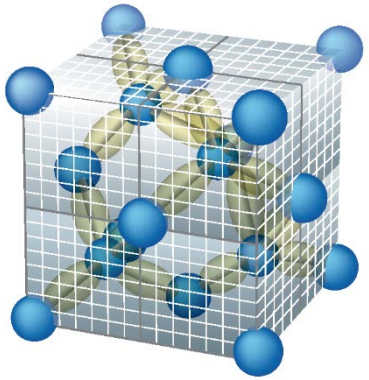
→ Multi-core massively parallel architecture with hardware accelerators

Collaboration between Computational Physics/Chemistry & Computer Science

➢ Developments of Mathematical Methodology and Algorithms Suitable to Current Hard wares

http://computics-material.jp/index.html

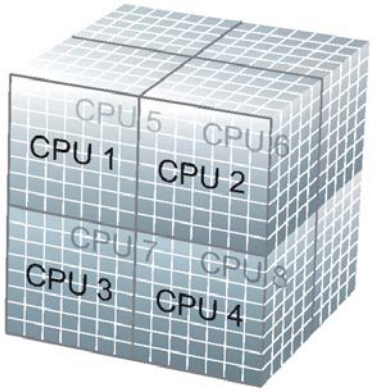
Advantages of Real-Space DFT (RSDFT) in Parallel Computing



Real space

Electron-state space

⟷ CPU space



On multi-core parallel machines:

- Huge number of 3D mesh points divided into moderate-size cells
- Each cell treated by a single node or core
- Also, electron states (orbitals) are grouped and each group is treated by a single node or a core : Hybrid parallelization
- MPI for CPU and OpenMP for core

- Almost free from FFT, reducing communication burden
  ⇒ high efficiency
- Flexible boundary condition to wave-functions
  ⇒ targets expanded including charged objects

High Performance Computing (HPC): Maximum Use of BLAS3 or Hit the Cache!

BLAS: Basic Linear Algebra Subprograms

Solving Kohn-Sham Equation

$$\left[-\frac{1}{2}\nabla^2 + v_{\text{eff}}(\mathbf{r}_k ; n(\mathbf{r}_k)) \right] \varphi_i(\mathbf{r}_k) = \varepsilon_i \varphi_i(\mathbf{r}_k)$$

Getting SCF by iterative computations

1. Conjugate gradient method
   $O(N^2)$
2. Ortho-normalization by Gram-Schmidt method $O(N^3)$
3. Density and Potential update
   $O(N^2)$

$O(N^3)$ operations, e.g., Gram-Schmidt should be accelerated!!

Linear Algebraic Calculation

1. Inner product of vectors

$$A = \sum_{j=1}^{N} a_j b_j$$ BLAS Level 1

Computing = 2N, Data = 2N + 1

2. Matrix times vector

$$A_i = \sum_{j=1}^{N} a_{ij} b_j$$ BLAS Level 2

Computing = $2N^2$, Data = $N^2 + 2N$

3. Matrix times matrix

$$A_{ik} = \sum_{j=1}^{N} a_{ij} b_{jk}$$ BLAS Level 3

Computing = $2N^3$, Data = $3N^2$

High ratio of (computing / data) in BLAS3 has an advantage in large-scale computations!!

Gram-Schmidt Ortho-normalization: Maximum use of BLAS Level 3 $O(N^3)$ Computation

$$\varphi_1 = \psi_1$$

$$\varphi_2 = \psi_2 - \varphi_1 \langle \varphi_1 | \psi_2 \rangle$$

$$\varphi_3 = \psi_3 - \varphi_1 \langle \varphi_1 | \psi_3 \rangle - \varphi_2 \langle \varphi_2 | \psi_3 \rangle$$

$$\varphi_4 = \psi_4 - \varphi_1 \langle \varphi_1 | \psi_4 \rangle - \varphi_2 \langle \varphi_2 | \psi_4 \rangle - \varphi_3 \langle \varphi_3 | \psi_4 \rangle$$

$$\varphi_5 = \psi_5 - \varphi_1 \langle \varphi_1 | \psi_5 \rangle - \varphi_2 \langle \varphi_2 | \psi_5 \rangle - \varphi_3 \langle \varphi_3 | \psi_5 \rangle - \varphi_4 \langle \varphi_4 | \psi_5 \rangle$$

$$\varphi_6 = \psi_6 - \varphi_1 \langle \varphi_1 | \psi_6 \rangle - \varphi_2 \langle \varphi_2 | \psi_6 \rangle - \varphi_3 \langle \varphi_3 | \psi_6 \rangle - \varphi_4 \langle \varphi_4 | \psi_6 \rangle - \varphi_5 \langle \varphi_5 | \psi_6 \rangle$$

.....

On K computer, 96.6 % efficiency using DGEMM to the peak performance

Most calculations can be performed as Matrix x Matrix operations!

$$[\varphi_3, \varphi_4, \varphi_5, \varphi_6, \ldots] = [\psi_3, \psi_4, \psi_5, \psi_6, \ldots] - [\varphi_1, \varphi_2] \begin{bmatrix} {}^{\dagger}\varphi_1 \\ {}^{\dagger}\varphi_2 \end{bmatrix} [\psi_3, \psi_4, \psi_5, \psi_6, \ldots]$$

RSDFT on K Computer at Kobe

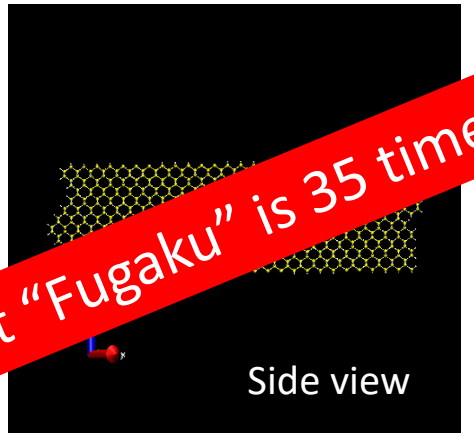
source

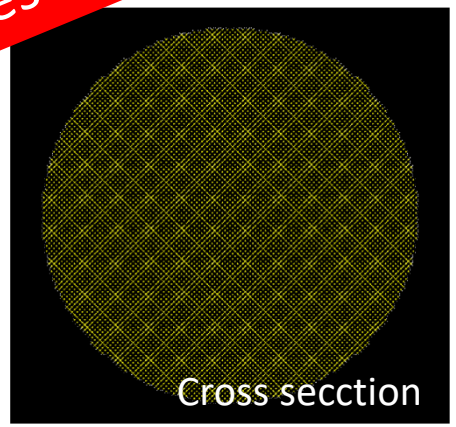
Performance at "Fugaku" is 35 times superior to "K"

Benchmark Data for Si nanowires

- 107,292-atom LDA/GGA electronic-structure calculations using 82,944 nodes
  - 5.48 PFLOPS = 51.7% efficiency to the peak performance
  - Single iteration in SCF = 2,900 sec

Gordon Bell Prize 2011



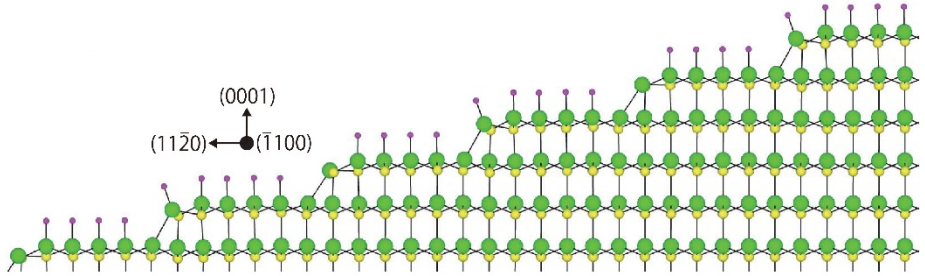


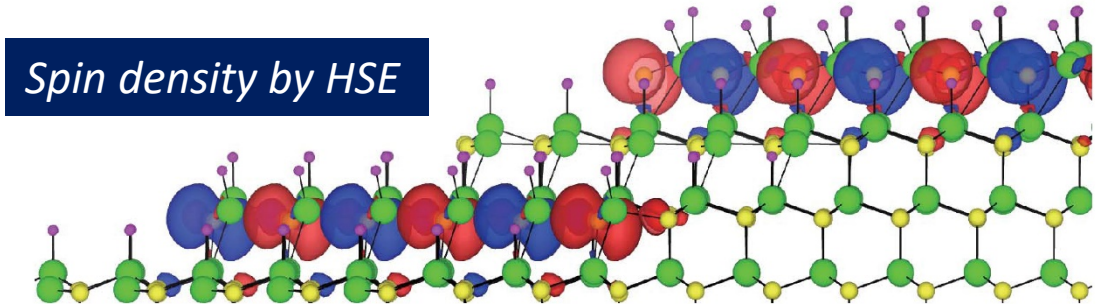
LDA/GGA Run for 10,000-atom Systems

- Getting SCFields = one day using ～10^2 nodes with 34% - 70% efficiency
- Good strong scaling (78% - 88%) for 4,608 – 15,552-atom systems using 768 – 12,288 nodes

HSE Calculations for 1000-atom systems

- Anti-ferromagnetic spin polarization along step edges of nanofacets in SiC





Sawada, Iwata, Oshiyama, PRB 93, 235421 (2016)