12th Symposium on Discovery, Fusion, Creation of New Knowledge
by Multidisciplinary Computational Sciences
CCS, Univ. Tsukuba
Oct. 6, 2020

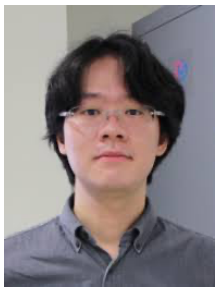
Performance Evaluation of Electron Dynamics Simulation in Supercomputer Fugaku

Ab initio Simulation in Current Optics Research

Kazuhiro YABANA

Center for Computational Sciences
University of Tsukuba

HPC - Condensed Matter Physics Collaboration



Yuta Hirokawa
CCS

Prof. Boku
CCS

HPC

Atsushi Yamada
CCS

Shunsuke Yamada
CCS

Mitsuharu Uemoto
Kobe Univ.

K. Yabana
CCS

Physics

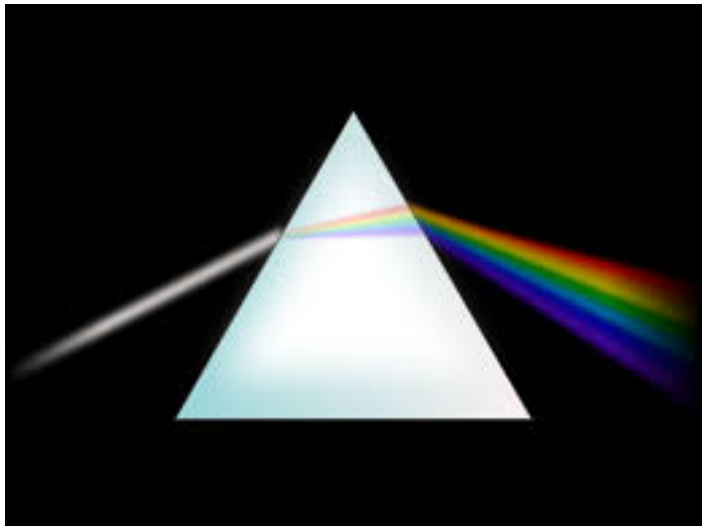
Contents

Contents

Light-Matter interaction in Physics

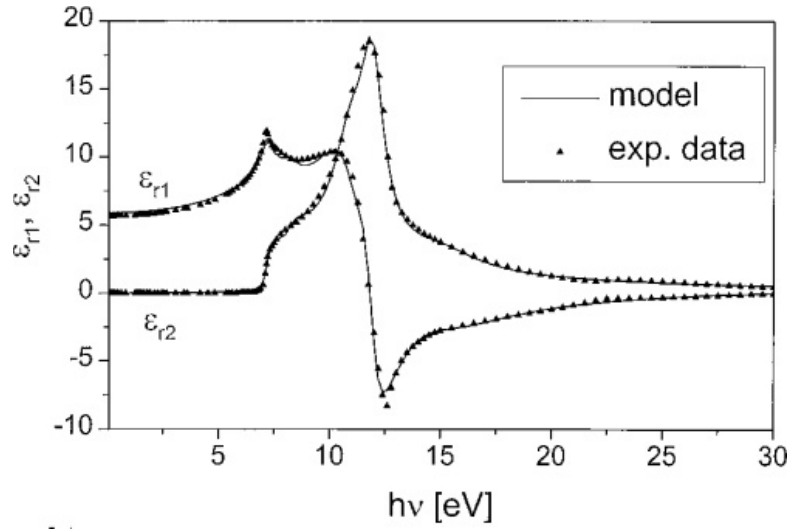
Light propagation
(Optics, Electromagnetism)

$$n = \sqrt{\varepsilon}$$



Dielectric function
(Condensed Matter Physics)

$$\varepsilon(\omega)$$



frequency

Computations in Optical Science So Far

ELECTROMAGNETISM (EM)

Light propagation solving Maxwell's equations

$$\nabla \cdot \boldsymbol{B} = 0$$
$$\nabla \times \boldsymbol{E} + \frac{\partial \boldsymbol{B}}{\partial t} = \boldsymbol{0}$$
$$\nabla \cdot \boldsymbol{D} = \rho$$
$$\nabla \times \boldsymbol{H} - \frac{\partial \boldsymbol{D}}{\partial t} = \boldsymbol{j}$$

e.g, FDTD (Finite-difference time-domain method)

QUANTUM MECHANICS (QM)

First-principles quantum mechanics calculations of optical constants

$$\varepsilon_r = 1 + \frac{2Ne^2}{\varepsilon_0 \hbar} \sum_j \frac{\omega_{j0} \left| \langle 0 | x | j \rangle \right|^2}{\omega_{j0}^2 - (\omega + i\gamma)^2}$$

e.g, GW theory, (TD)DFT

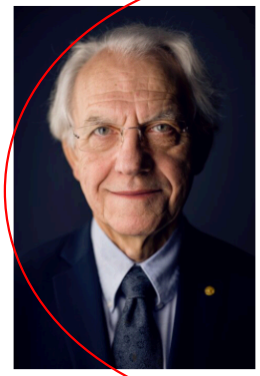
These traditional approaches are insufficient in current optics frontier.

One of directions in current optics:

Generate and Utilize Intense and Ultrashort Laser Pulse (Chirped Pulse Amplification)

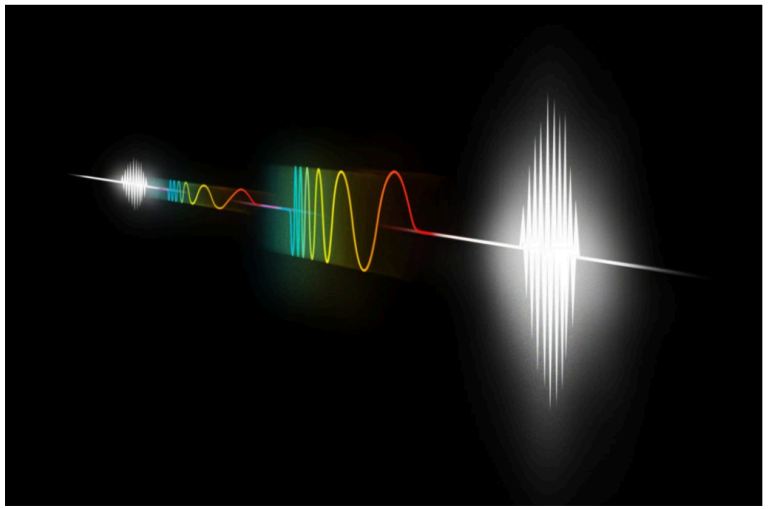
Nobel Prize in Physics 2018

Ill. Niklas Elmehed. © Nobel Media
Arthur Ashkin



© Nobel Media AB. Photo: A. Mahmoud
Gérard Mourou

© Nobel Media AB. Photo: A. Mahmoud
Donna Strickland

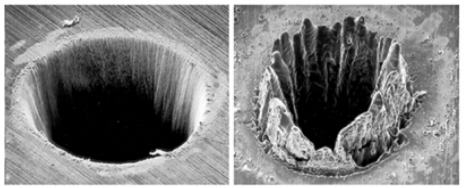


© Johan Jarnestad / The Royal Swedish Academy of Sciences

Frontiers of optical science using extreme pulsed light

High field

Laser processing, surgery



Electric field

Nonlinear

time

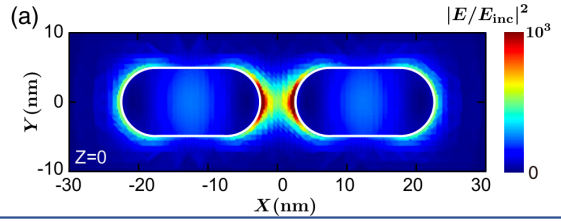
Ultrafast

length

Nonlocal

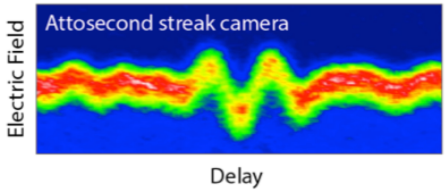
Nano-structure

Near field, plasmonics, meta-materials



Ultrashort pulse

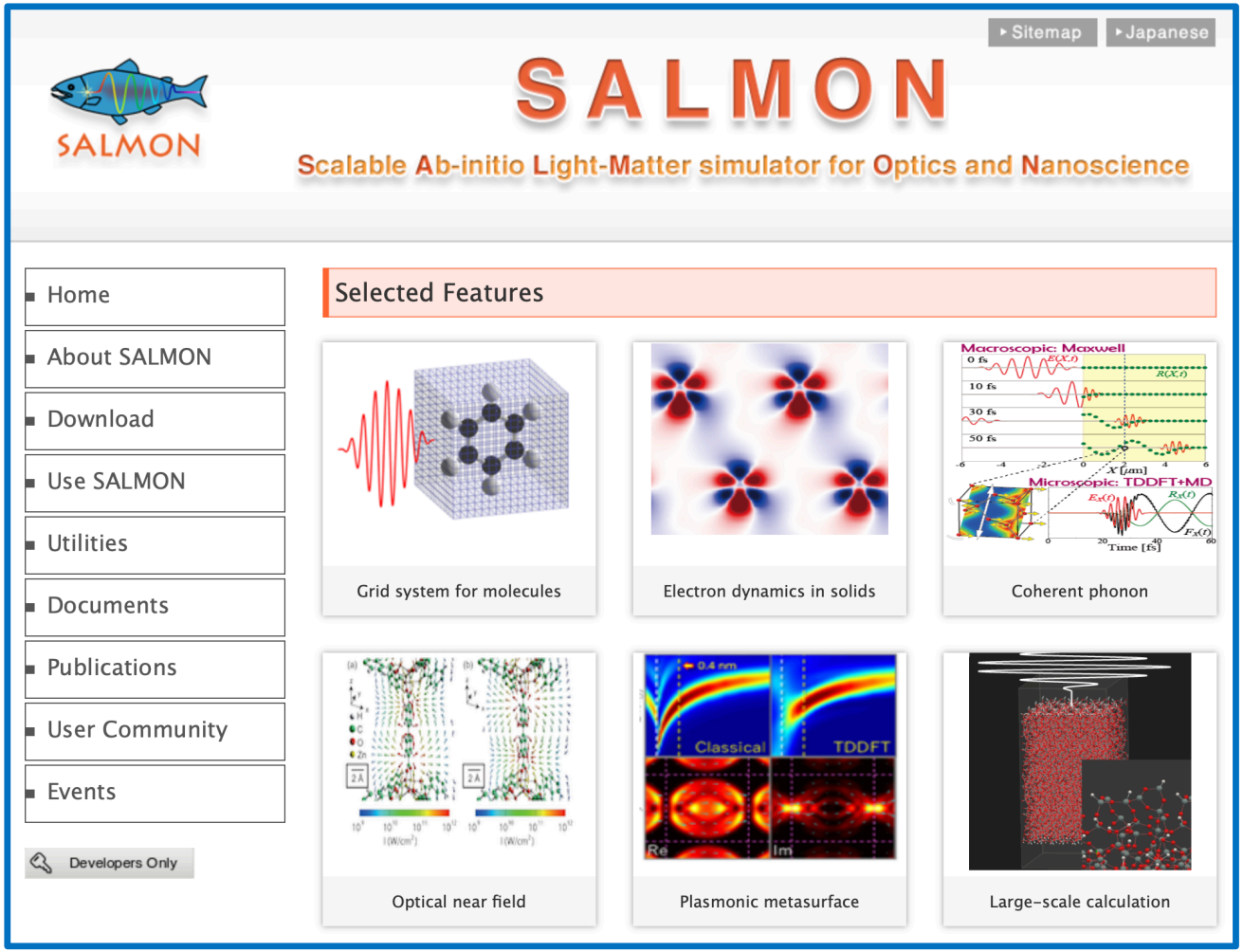
Attosecond Science



Here, QM + EM coupled approach is required.

Contents

Open-Source Software project: http://salmon-tddft.jp



Supported by CREST 戦略的創造研究推進事業 Core Research for Evolutionary Science and Technology, CdMSi, Q-LEAP

Light-Matter interaction: Multiphysics + Multscale phenomena

第一原理電子ダイナミクス計算プログラム

Scalable Ab-initio Light-Matter simulator for Optics and Nanoscience

https://salmon-tddft.jp/

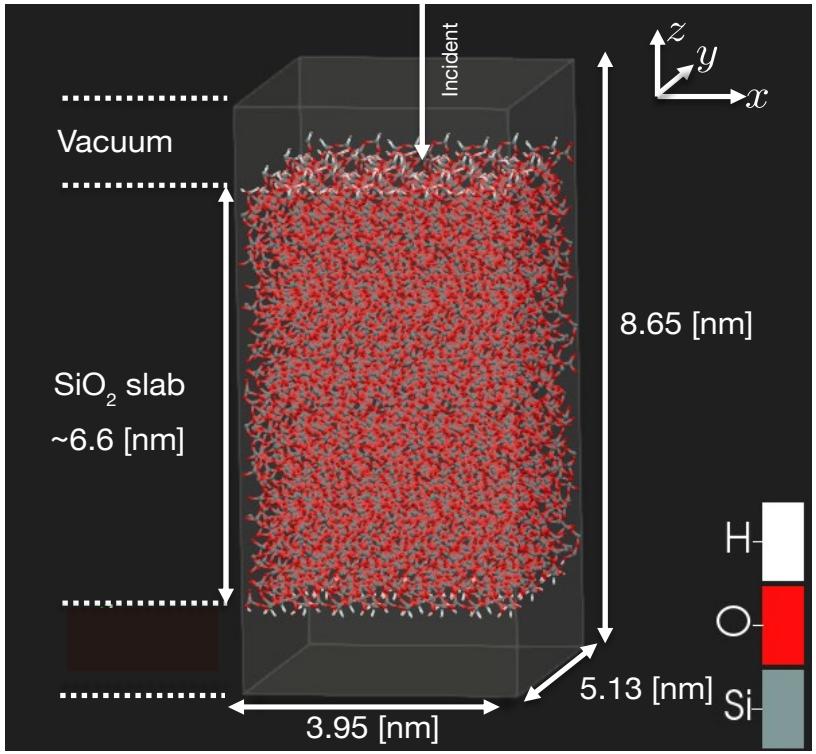
Dr. Mitsuharu Uemoto (Kobe Univ.)

Simulation of light-matter interaction using Fugaku

27,648 nodes, about 1/6 of full system

Dynamics of light EM fields, electrons, and ions of more than 10,000 atoms

Intense and ultrashort laser pulse Irradiates normally on SiO_2 thin film.



Basic Equations

Maxwell's Eq. for vector and scalar potentials (light fields)

$$\nabla^2 V_H = 4\pi\rho[\psi_b]$$

$$\frac{1}{c}\frac{\partial \mathbf{A}}{\partial t^2} - \nabla^2 \mathbf{A} = \frac{4\pi}{c^2}\mathbf{j}[\psi_b, \mathbf{R}_a]$$

Time-dep. Kohn-Sham eq. for electron orbitals

$$i\hbar\frac{\partial}{\partial t}\psi_b = \left\{\frac{1}{2m}\left(-i\hbar\nabla + \frac{e}{c}\mathbf{A}\right)^2 + V_{ion}[\mathbf{R}_a] + V_H + V_{xc}\right\}\psi_b$$

Newton eq. for ion coordinates

$$M_a\frac{d^2\mathbf{R}_a}{dt^2} = \mathbf{F}_a[\psi_b, \mathbf{A}]$$

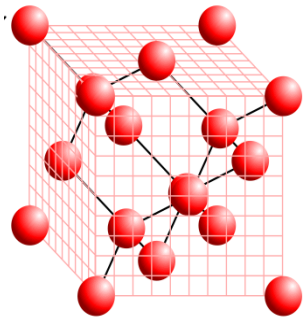
3D spatial grid

$V_H(\mathbf{r}, t)$

$\mathbf{A}(\mathbf{r}, t)$

$\psi_b(\mathbf{r}, t)$

$O(N^2)$



Number of atoms: 10k
Number of orbitals: 40k
Maximum grid size $< 256^3$

To prepare initial condition, need to solve static Kohn-Sham equation

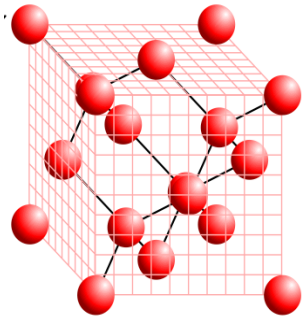
Initial condition: ground state

$$\phi_b(\mathbf{r}) = \left\{ \frac{1}{2m}\nabla^2 + V_{ion} + V_H + V_{xc} \right\} \phi_b(\mathbf{r})$$

Bottleneck:
Gram-Schmidt orthogonalization
Subspace diagonalization

$$O(N^3)$$

100,000 atoms by RSDFT
SC11 Gordon Bell Prize
using K computer,
H. Hasegawa, J.-I. Iwata et.al.



Optical response: time evolution

$$i\hbar\frac{\partial}{\partial t}\psi_b = \left\{ \frac{1}{2m}\left(-i\hbar\nabla + \frac{e}{c}\mathbf{A}\right)^2 + V_{ion}[\mathbf{R}_a] + V_H + V_{xc} \right\} \psi_b$$

Bottleneck:
Operation of Kohn-Sham Hamiltonian
on electron orbitals
- Stencil operation
- Nonlocal pseudopotential operation

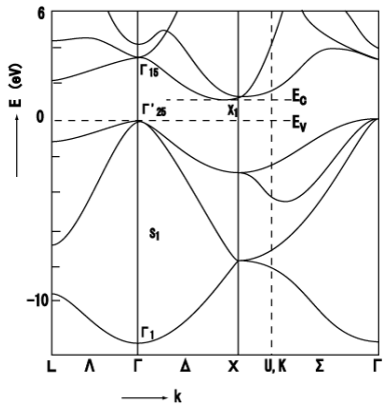
$$O(N^2)$$

More than 10,000 atoms
for the first time

Electron dynamics in a unit cell of solid (bulk Si)

16 atoms, 32 electrons
in the unit cell

Energy band of silicon



Time-dependent Kohn-Sham equation

$$i\hbar\frac{\partial}{\partial t}u_{nk}(r,t)=\left[\frac{1}{2m}\left(p+\hbar k+\frac{e}{c}A(t)\right)^2+V_H(r,t)+V_{xc}(r,t)\right]u_{nk}(r,t)$$

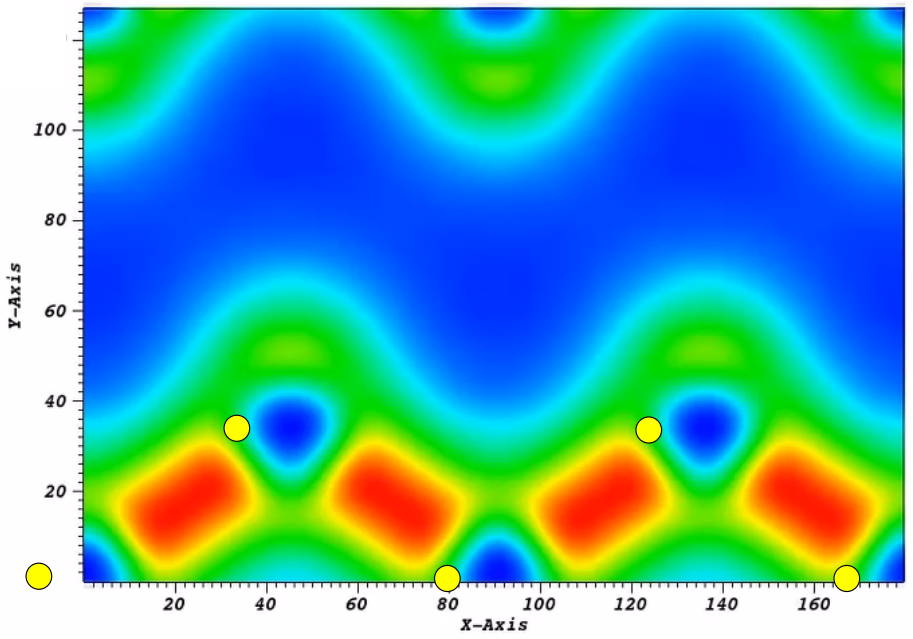
$$A(t)=-c\int^t E(t')\mathrm{dt'}$$

Electron density in crystalline silicon

Atomic positions fixed

$\hbar\omega=1.55\mathrm{eV}$

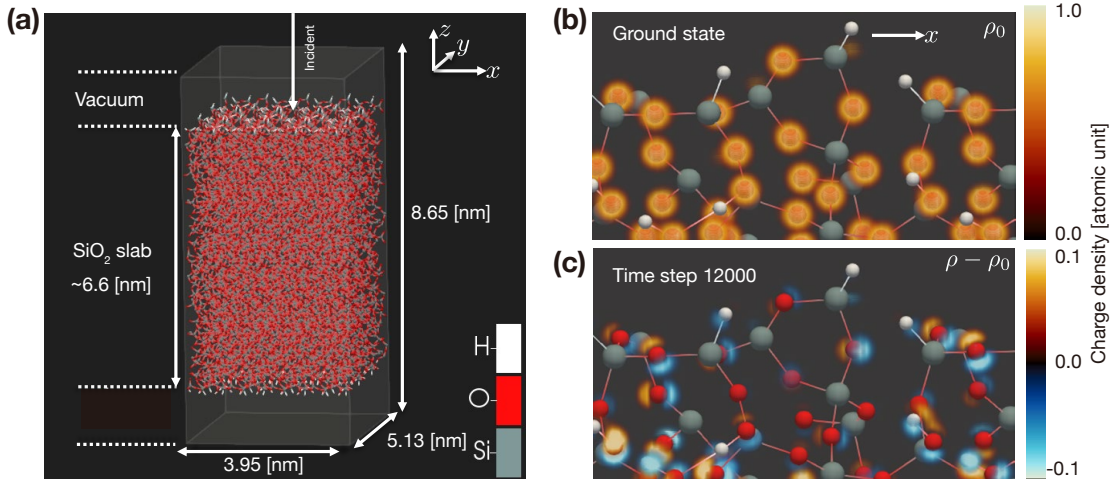
$I=1.0\times10^{14}\ \mathrm{W/cm^2}$

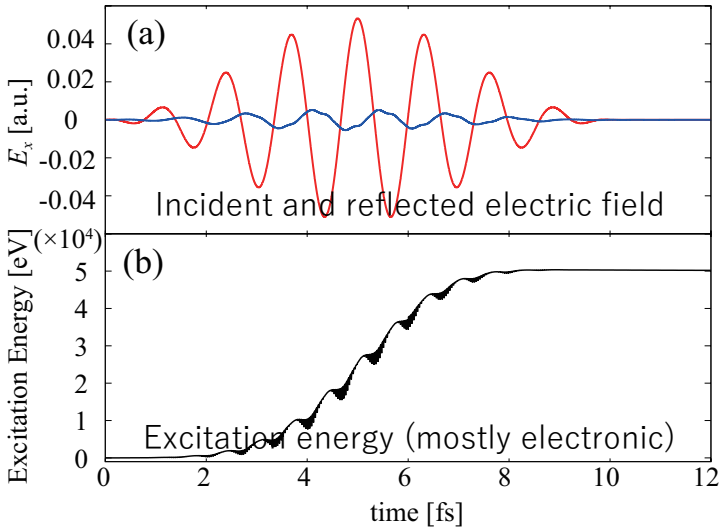


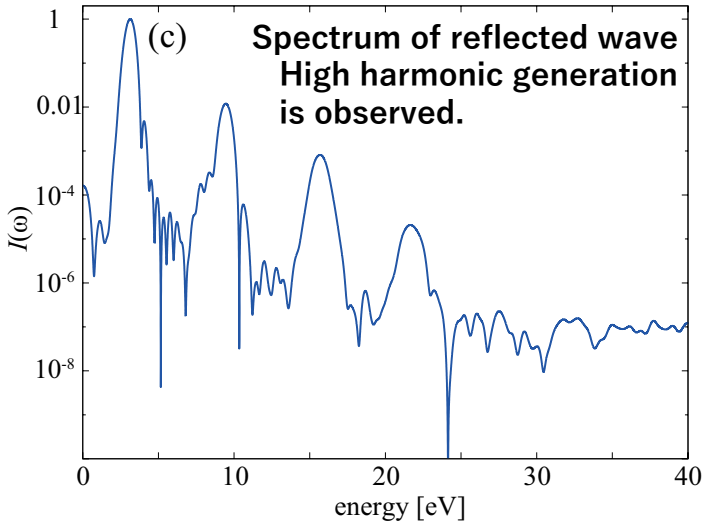
Production Run: 27,648 nodes, 6 hours in Fugaku

Material
10,224 atoms
(SiO_2 thin film)

Laser
3.1eV, 10fs,
10^{14}W/cm^2







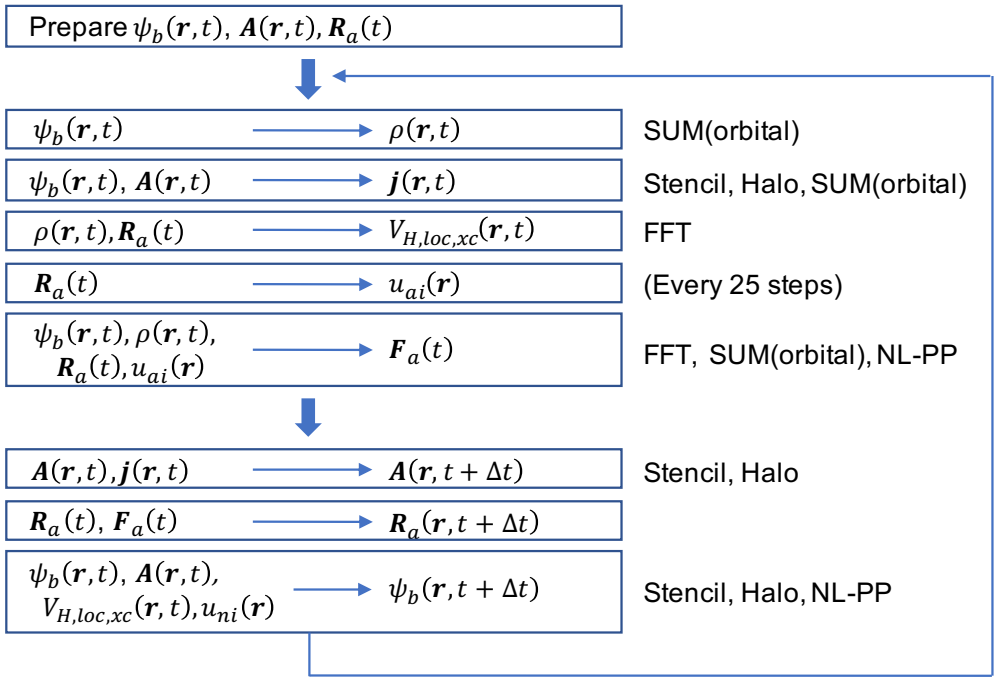
Computational parameters

Atoms: 10,224
Orbitals: 26,640
Grids: 216*288*480
Time steps: 16,000

MPI process allocation:
MPI process #: 27,648*4
Orbital: 288
Grid: 4*8*12

Contents

Characteristic of computation and communication

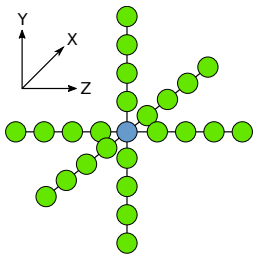


$O(N^2)$

$\psi_b(r,t)$	Wave function (Electron orbital)
$\rho(r,t)$	Charge density of electrons
$V_H(r,t)$	Hartree potential (Poisson equation)
$A(r,t)$	Vector potential (FDTD)
$R_a(t)$	Ion coordinate

SUM (orbital)	MPI_Allreduce(DP vector)
Stencil	25-points stencil (DP complex)
Halo	Halo communication
FFT	3-D FFT by FFTE
NL-PP	SpMV-like calculation

Optimization of computation and communication



- 25 points stencil computation is dominant
  - Target loop size: 128^3 - 256^3
  - Already carried out 512-bit SIMD optimization in Xeon Phi (KNC & KNL) processors
  - **Optimization in Fujitsu compiler**

- 6 to 8 MPI sub-communicators are used
  - Orbitals are 4D array (3D spatial + orbital index), whereas Fugaku provides 3D network (Tofu-D)
  - **Need to find optimal mapping of MPI processes to the Tofu-D network**

Optimization of stencil computation

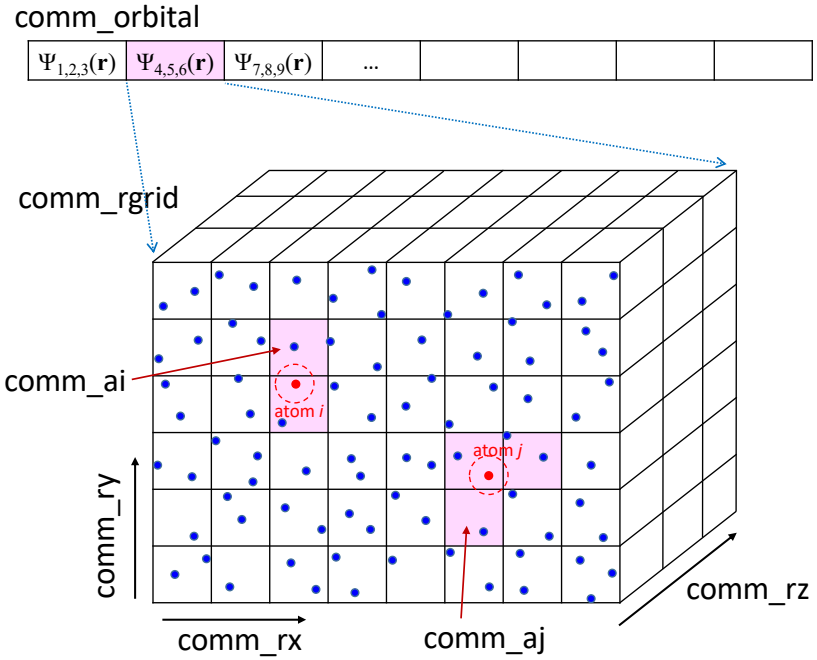
```
do ib = 1,Norbital ; do iz = 1,Lz; do iy = 1,Ly


! initialize
  do ix = 1,Lx
    htpsi(ix,iy,iz,ib,ik) = &
      (V_local(ix,iy,iz,ib,ik) + lap0(ik)) &
      * tpsi(ix,iy,iz,ib,ik)
  end do
! compute X-difference
  do ix = 1,Lx ; ... ; end do
! compute Y-difference
  do ix = 1,Lx ; ... ; end do
! compute Z-difference
  do ix = 1,Lx ; ... ; end do


end do; end do; end do
```

- SIMD + SWP (SoftWare Piplining) optimization
  - Loop unroll
    + Loop scheduling
    + Improve instruction-level parallelism
- Instruction latency hiding is important because of longer latency
  - FADD, FMAD: 9 cycle
  - ref. https://github.com/fujitsu/A64FX
- Reduce register usage by reducing vectorized-loop body size

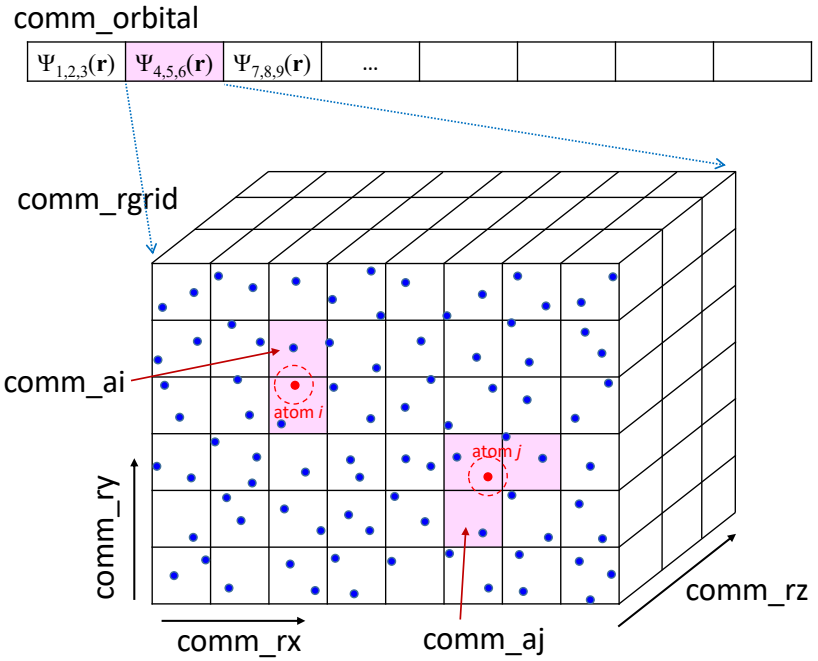
MPI process distribution strategy

- Localize communications by multiple MPI subcommunicators

Communicator	
comm_orbital	
comm_rgrid	
comm_rx, ry, rz	Comm. along x, y, z axes in comm_rgrid
comm_ai, aj	Grids associated with atoms to treat nonlocal pseudopotential



Communication optimization: MPI process mapping for Tofu-D(1)



- Electron orbitals: 4D array (x, y, z, b)
  - Spatial grid: comm_rgrid (x, y, z)
  - Orbital: comm_orbital (b)
- Time evolution calculation
  - comm_rgrid dominates
  - Mapping to network, minimizing hop counts (halo + collective) in comm_rgrid communications.

(Ground state calculation)

- comm_orbital dominates

Communication optimization: MPI process mapping for Tofu-D(2)

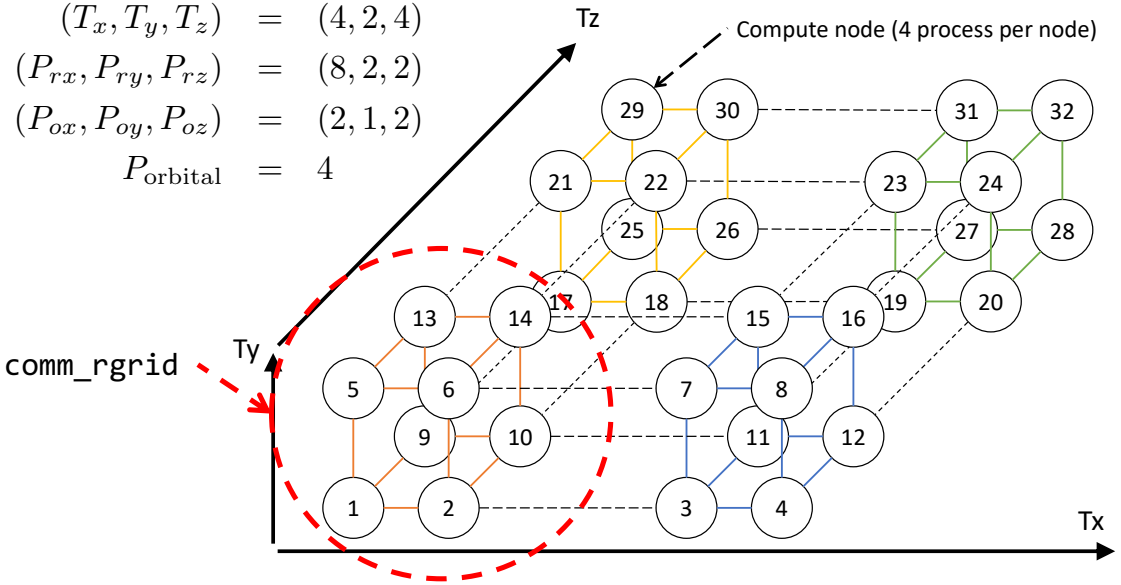
- Tofu-D appears as 3D network for users.
- 4 MPI process/node is recommended.

1. How to map 4D(grid+orbital) process space into 3D network?
   - ➢ Prepare process shape as 3D+1D system

2. How to treat 4 MPI process within a node?
   - ➢ Allocate one axis of 3D(grid) MPI process space.

Communication optimization: MPI process mapping for Tofu-D(3)

$$
\begin{aligned}
(T_x, T_y, T_z) &= (4, 2, 4) \\
(P_{rx}, P_{ry}, P_{rz}) &= (8, 2, 2) \\
(P_{ox}, P_{oy}, P_{oz}) &= (2, 1, 2) \\
P_{\mathrm{orbital}} &= 4
\end{aligned}
$$



1. Determine P_{rx}, P_{ry}, P_{rz}, P_{orbital}
2. Express P_{orbital} as a product of (P_{ox}, P_{oy}, P_{oz})
3. Introduce (T_x, T_y, T_z) = ($P_{rx} \times P_{\mathrm{node}}$, P_{ry}, P_{rz})

(T_x, T_y, T_z)	3-D node shape
P_{all}	Total # of process
(P_{rx}, P_{ry}, P_{rz})	# of proc. / spatial grid
P_{orbital}	# of proc. / orbital
$P_{\mathrm{node}} = 4$	# of proc. / node

$$
\begin{aligned}
P_{rx} \times P_{ry} \times P_{rz} \times P_{\mathrm{orbital}} &= P_{\mathrm{all}} \\
(T_x \times P_{\mathrm{node}}) \times T_y \times T_z &= P_{\mathrm{all}} \\
(T_x \times P_{\mathrm{node}}) \div P_{rx} &= P_{ox} \\
T_y \div P_{ry} &= P_{oy} \\
T_z \div P_{rz} &= P_{oz} \\
P_{ox} \times P_{oy} \times P_{oz} &= P_{\mathrm{orbital}}
\end{aligned}
$$

Contents

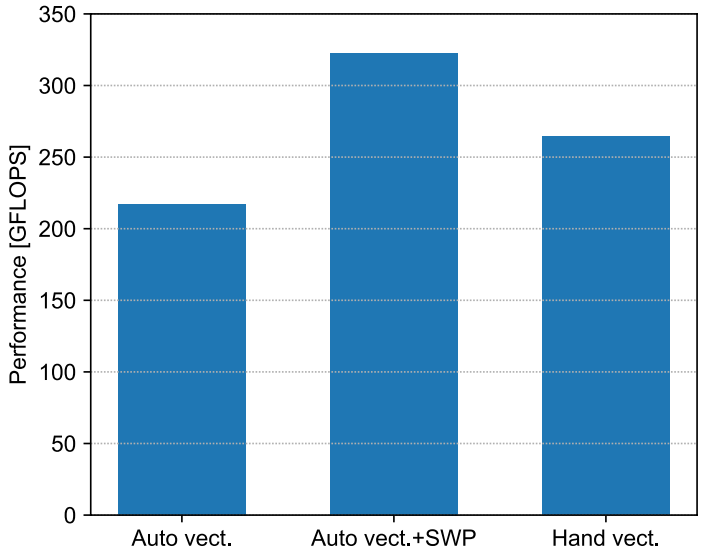
Evaluation environment: Supercomputer Fugaku

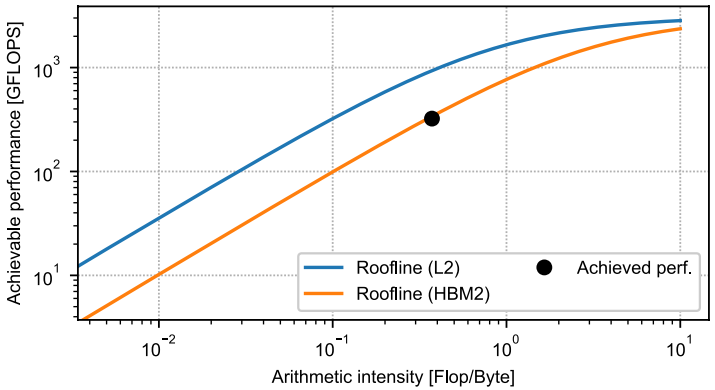
	Fugaku
Processor	Fujitsu A64FX (based on Arm v8.2-a)
Compute core	12 physical cores (core memory group, CMG) with 4 CMGs = 48 physical cores with 2 GHz
Cache	L1D$: 64 KiB, L2$: 8 MiB (CMG shared)
Memory	32 GiB: HBM2 (High bandwidth memory)
SIMD instruction	Arm SVE with 512-bit
Theoretical peak performance	3072 GFLOPS
Evaluated system	27,648 node, 84.935 PFLOPS
Network	Tofu-D (6-dimensional mesh/torus)

Disclaimer
- The results obtained on the evaluation environment in the trial phase do not guarantee the performance, power and other attributes of the supercomputer Fugaku at the start of its public use operation.

1. Single-node performance
   - Stencil calculation
   - Entire application
2. Weak scaling evaluation
   - Fixed size of orbitals
     - (54, 48, 40, 123.34)
     - 5 GB/process (20 GB/node)
   - $O(N^2)$ scaling, two times for both orbitals and grids

Single-node performance

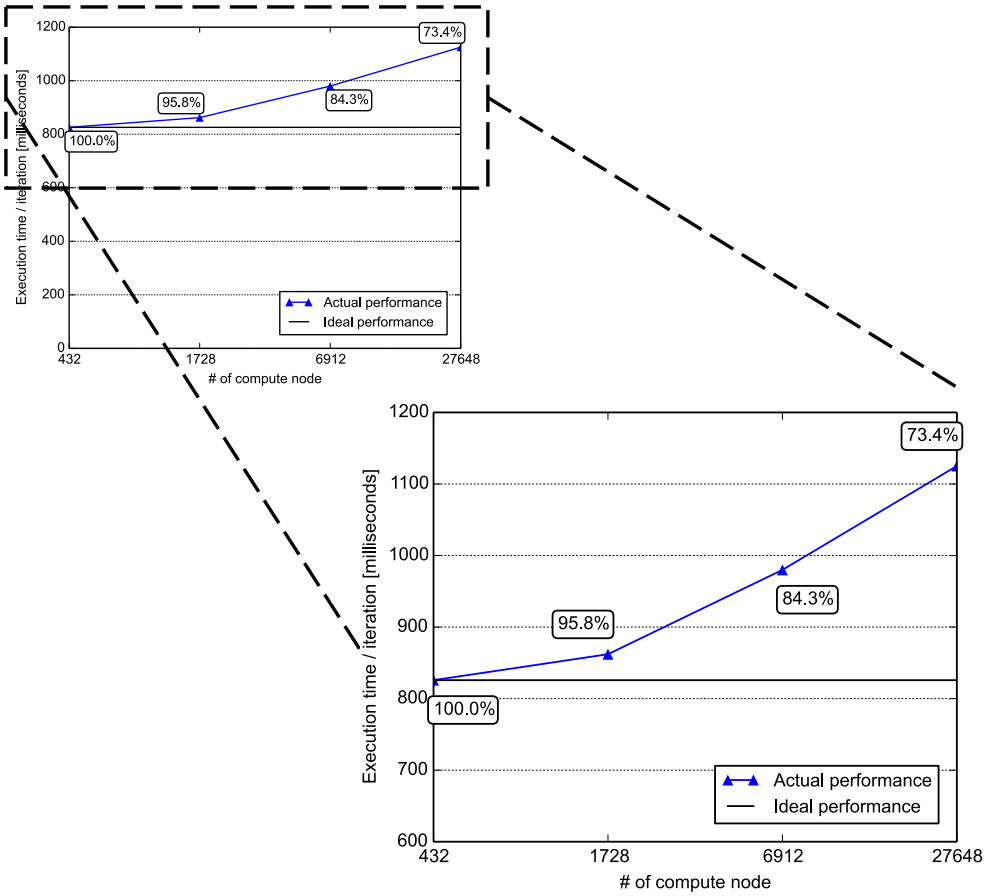




Auto vect.	Vector optimization by compiler
Auto vect. + SWP	+ Software pipelining optimization
Hand vect.	512-bit SVE hand-coding optimization

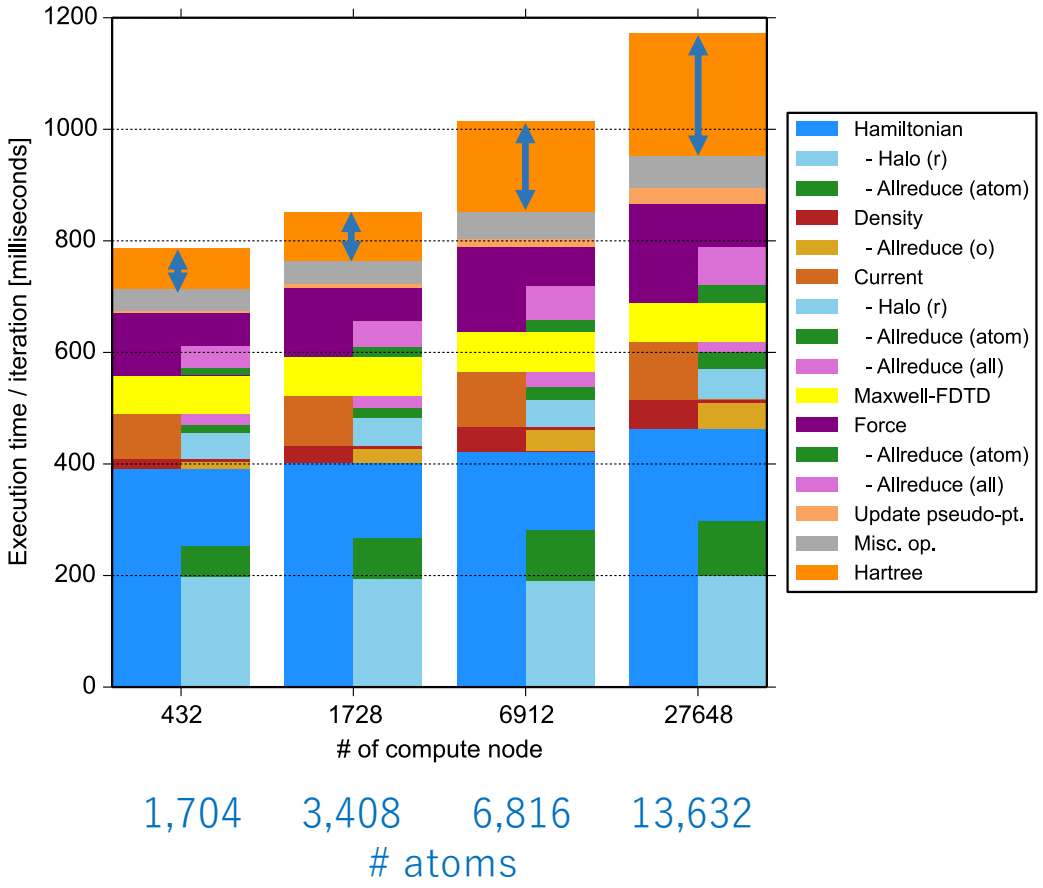
- Auto vect. + SWP : 322.7 GFLOPS
  - 1.4－5 times difference in performance w/wo SWP
  - 10.5% effective performance
  - Close to performance by Roofline model using HBM2 memory bandwidth
- Actual memory bandwidth is 866 GB/s
  - Bandwidth efficiency is approx. 84.5%
- Total efficiency of TDDFT calculation is 321.6 GFLOPS
  - Mostly bounded by Stencil performance

Weak scaling (1)



- Execution time of single time step
  - 1 s/iter: expected performance for daily simulations
- Effective performance gets down
  73.4% from 432 to 27648=432*4^3 nodes
- Linear reduction (about 11%)
  - **Expect 56% at 150k full nodes (1.5 sec/iter.)**
- What makes reduction?
  - Look into breakdown

Weak scaling (2)



- Breakdown of weak scaling results
  - Communication time is also plotted
  - No hiding of calculation/communication
- Hartree (FFT) is a bottleneck
  - **FFT is achieved within each grid process (comm_rgrid) redundantly.**
    - 96 nodes maximum (384 process)
  - We use FFTE of 3D-FFT, dividing (Y, Z) direction
  - Since the whole system is divided into 3D process, MPI_Alltoall message increases 2 times in 1,728 -> 6,912
    → increase communication time

Validity of achieved performance (1)

- Performance expected from Byte/FLOP (w/o communication costs)
  - Stencil calculation : 2.68 Byte/FLOP
  - System value : 0.33 Byte/FLOP (1024 GB/s ÷ 3072 GFLOPS)
  - Ratio to peak performance : 0.33 / 2.68 ≒ **12.4%**

[PFLOPS]	Theoretical peak	Estimated	Actual (vs. Estimated)
432	1.327	0.165	0.057 (34.5%)
1,728	5.308	0.661	0.219 (33.1%)
6,912	21.234	2.646	0.773 (29.2%)
27,648	84.935	10.583	2.692 (25.4%)

- Due to memory-bandwidth bound, we could not expect high performance in view of peak performance.

Validity of achieved performance (2)

Comparison with previous calculations

- TDDFT calculation using full Sequoia BG/Q (Draeger, et. al.: IEEE Cluster 2016)

	Peak perf.	Achieved perf.	Time/iter.	Problem size	Dominant kernel
Sequoia BG/Q	20 PFLOPS	8.6 PFLOPS	53.2 sec	5400 atom	(D/Z)GEMM
Fugaku	85 PFLOPS	2.692 PFLOPS	1.13 sec	13632 atom	Stencil

- BG/Q shows high FLOPS of 43%
  - Draeger et.al use plane-wave basis in which bottleneck is nonlocal pseudopotential operation and is treated by (D/Z)GEMM
  - In our grid representation, stencil is the first bottleneck
- Time-to-Solution
  - Assuming strong scaling, BG/Q is expected 12.5 sec/iter for the same problem.
  - In SALMON, 1 sec/iter for 2 times larger system.
    → 20 times better time-to-solution

Summary

- Optimization and performance measurements of SALMON in Fugaku
  - In stencil calculation, achieved effective memory-bandwidth of 866 GB/s (84%)
  - For weak scaling, achieved 75 – 95% effective performance
  - 0.8 – 1.13 second/iteration that satisfies conditions for simulations of real system. It is more than 20 times better time-to-solution than previous study.

- Toward extreme simulations using full system of Fugaku:
  - We may expect calculations of systems up to 20,000-30,000 atoms.
  - Performance degradation of FFT:
    - Still endurable since FFT is carried out within 128 – 192 nodes.
  - Possible shortage of memory
    - At present, use 60 – 75% (20 GB/node). Dominant requirement scales.