Low Precision Computing in Sparse Linear Solvers

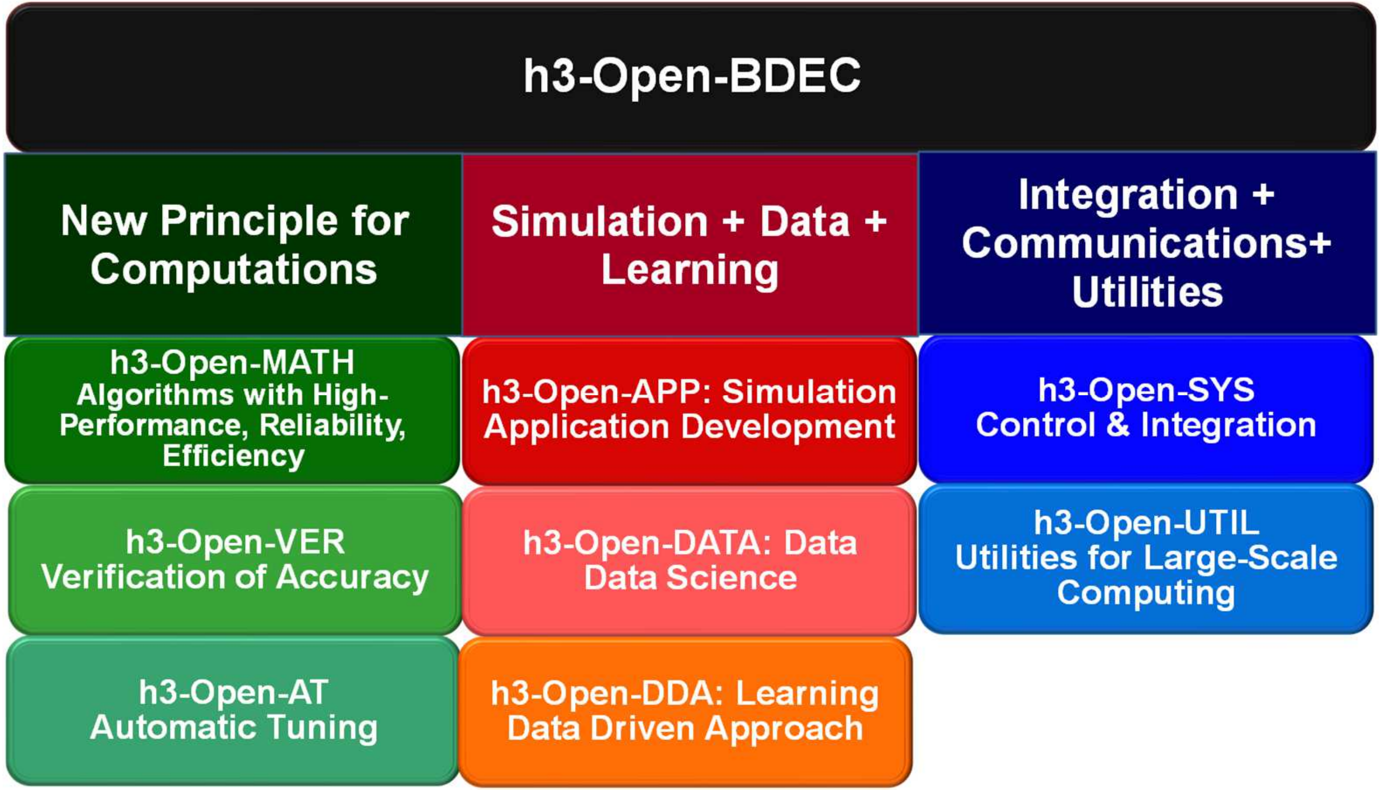
Kengo Nakajima
Information Technology Center
The University of Tokyo

h3-Open-BDEC: Innovative Software Platform for Integration of (Simulation+Data+Learning) (S+D+L) on the BDEC System

- **5-year project supported by Japanese Government through JSPS Grant-in-Aid for Scientific Research (S) since 2019**
  - **科研費基盤S**
- **Leading-PI: Kengo Nakajima (The University of Tokyo)**
- **Total Budget: 152.7M JPY= 1.41M USD**

h3-Open-BDEC

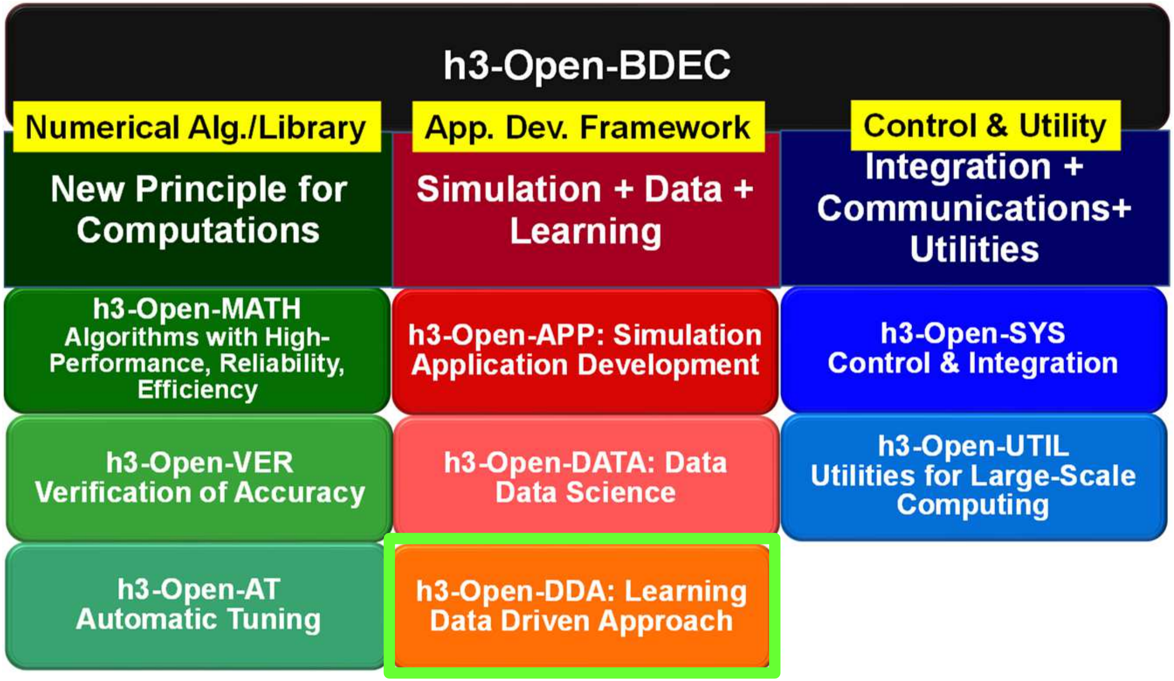
Innovative Software Platform for Integration of (S+D+L) on BDEC



h3-Open-BDEC: Two Significant Innovations

① Methods for Numerical Analysis with High-Performance/High-Reliability/Power-Saving based on the New Principle of Computing by

- ✓ Adaptive Precision
- ✓ Accuracy Verification
- ✓ Automatic Tuning



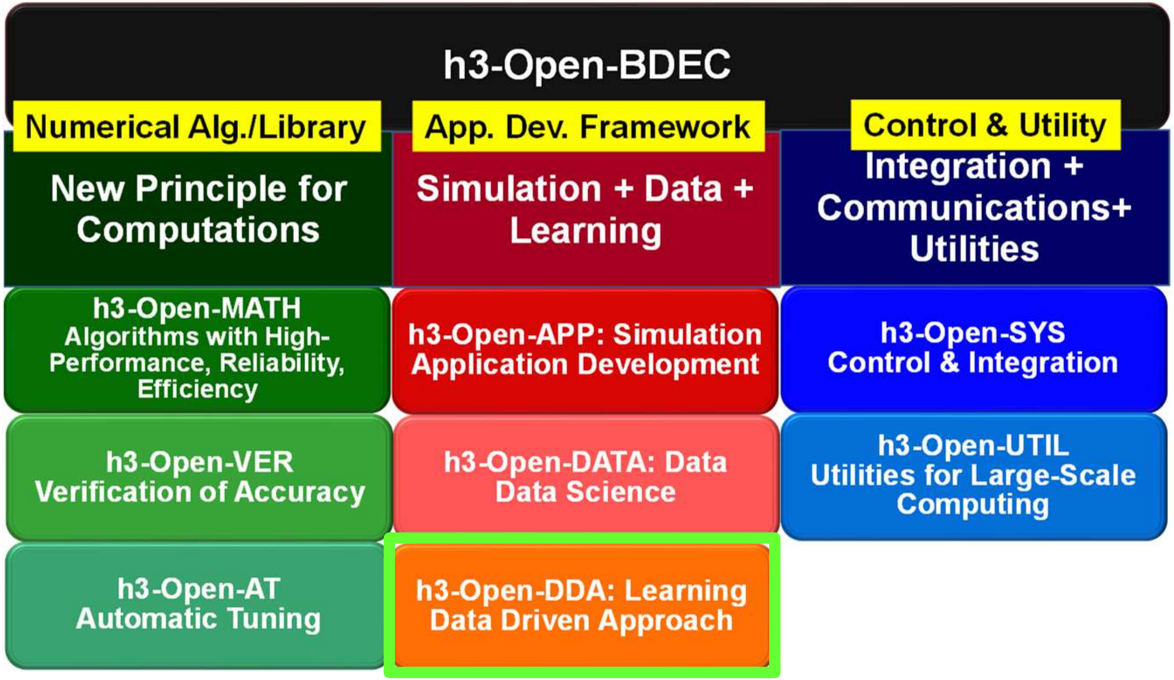
h3-Open-BDEC: Two Significant Innovations

① Methods for Numerical Analysis with High-Performance/High-Reliability/Power-Saving based on the New Principle of Computing by

- ✓ Adaptive Precision
- ✓ Accuracy Verification
- ✓ Automatic Tuning

② Hierarchical Data Driven Approach (*h*DDA) based on machine learning

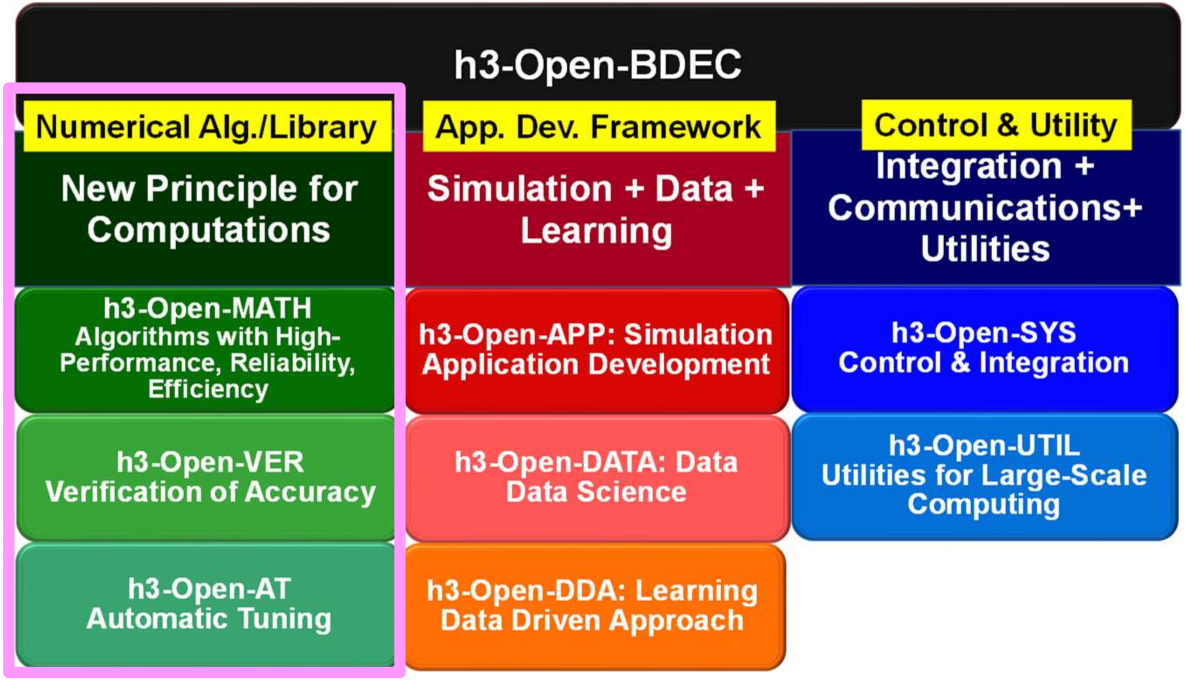
- ✓ Integration of (S+D+L) <u>AI for HPC</u>



h3-Open-BDEC: Two Significant Innovations

① Methods for Numerical Analysis with High-Performance/High-Reliability/Power-Saving based on the New Principle of Computing by

- ✓ Adaptive Precision
- ✓ Accuracy Verification
- ✓ Automatic Tuning

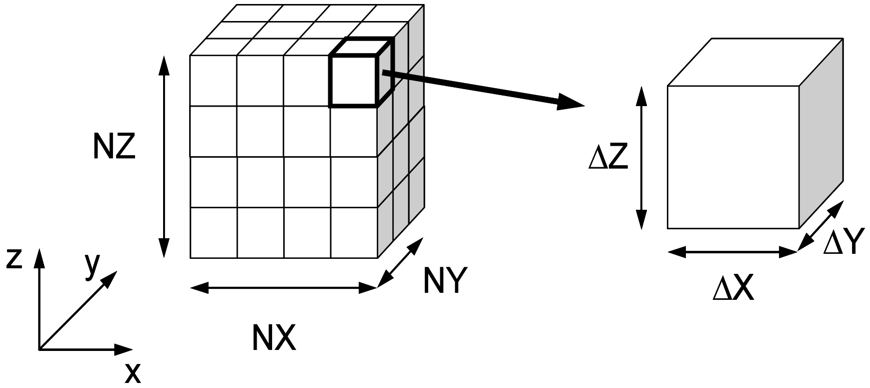


Approximate Computing with Low/Adaptive/Trans Precision

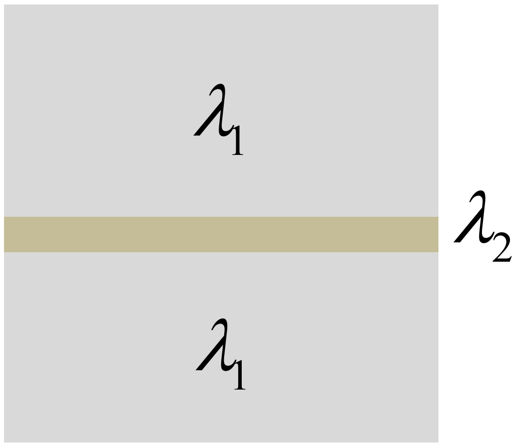
- Mostly, scientific computing has been conducted using FP64 (double precision, DP)
  – Sometimes, problems can be solved by FP32 (single precision, SP) or lower precision
- **Lower precision may save time, energy and memory**
- Approximate Computing
  – Originally for image recognition etc. where accuracy is not necessarily required
  – Also applied to numerical computations
- Computations by lower precision and by mixed precision may provide results with less accuracy

P3D: Steady State 3D Heat Conduction by FVM (1/2)

$$\nabla \cdot \left(\lambda \nabla \phi \right) + f = 0$$

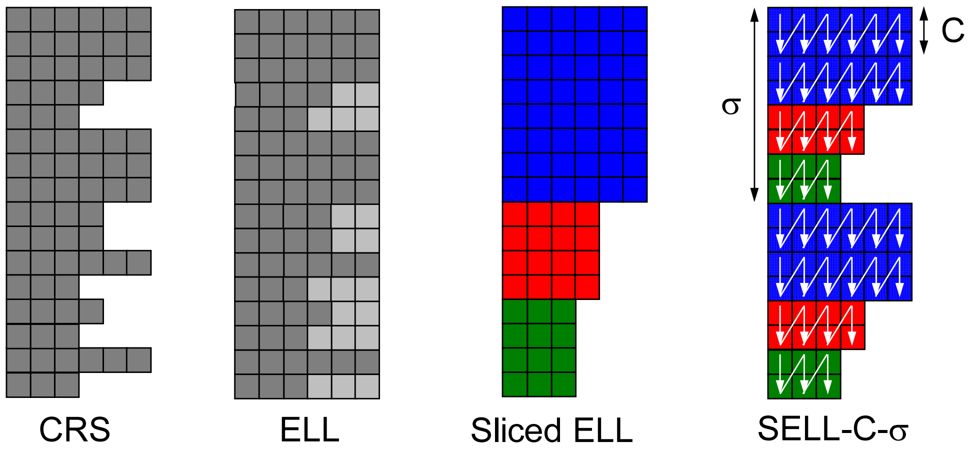


- 7-point Stencil
- Heterogenous Material Property
  - λ_1/λ_2 is proportional to the condition number of coefficient matrices
- Coefficient Matrix
  - Sparse, SPD
- ICCG Solver
- Fortran 90 + OpenMP
- CM-RCM Reordering



P3D: Steady State 3D Heat Conduction by FVM (2/2)

- Various Configurations
  - FP64 (Double), FP32 (Single), FP16 (Half) (just for preconditioning)
  - Matrix Storage Format (CRS, ELL, SELL-C-σ etc.)
    - CRS is applied in the present work

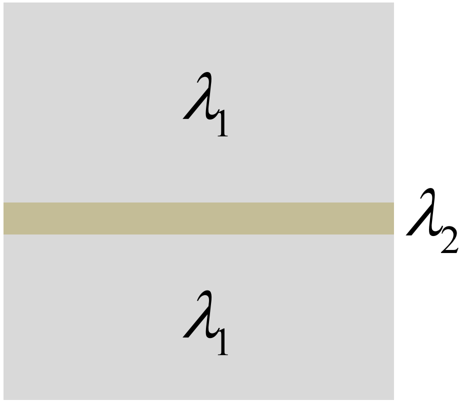


Ratio of FP32(SP)/FP64(DP)

Iterations● & Time△ for ICCG

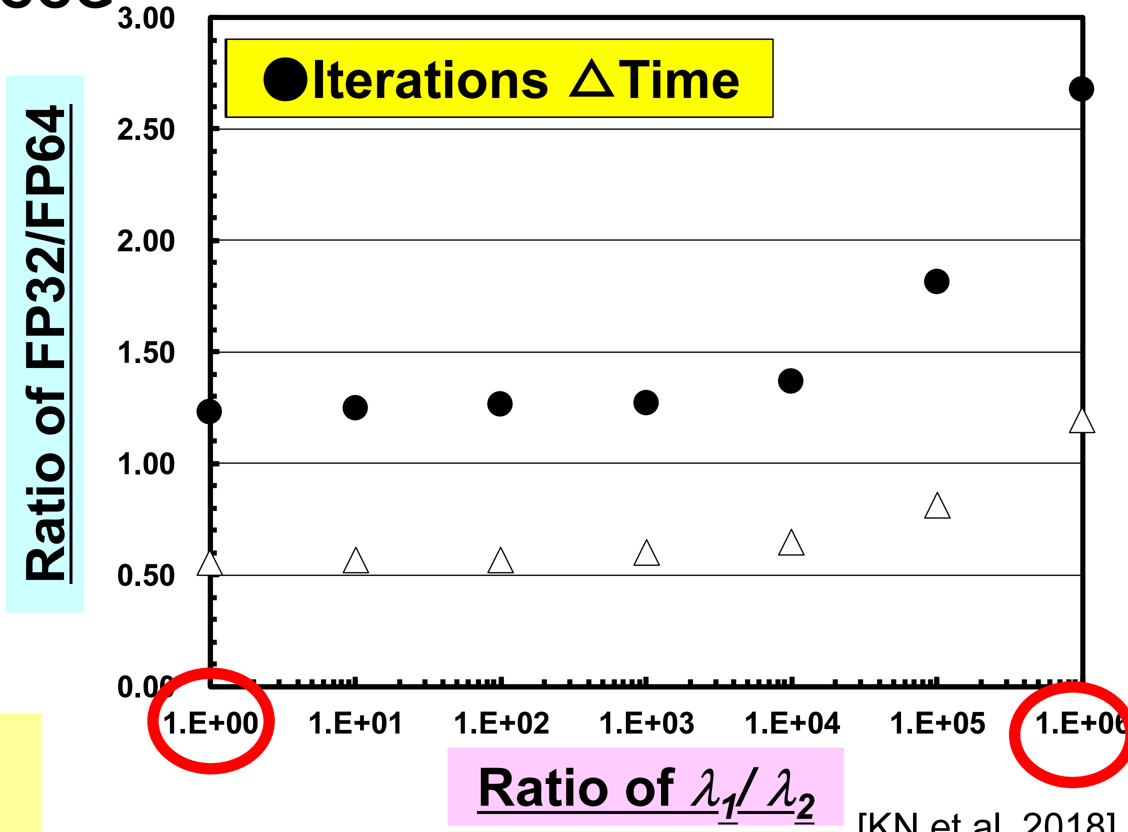
λ_1/λ_2, 128^3 DOF, CRS

Ratio<1 ⇒ FP32 is faster



$$\nabla \cdot (\lambda \nabla \phi) + f = 0$$

Intel Xeon BDW
1 Node: 18 cores x 2 soc's



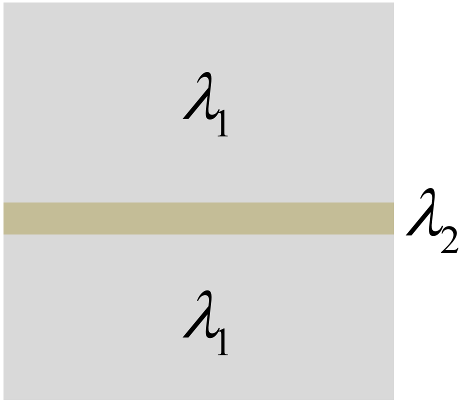
[KN et al. 2018]

Ratio of FP32(SP)/FP64(DP)

Iterations● & Time△ for ICCG

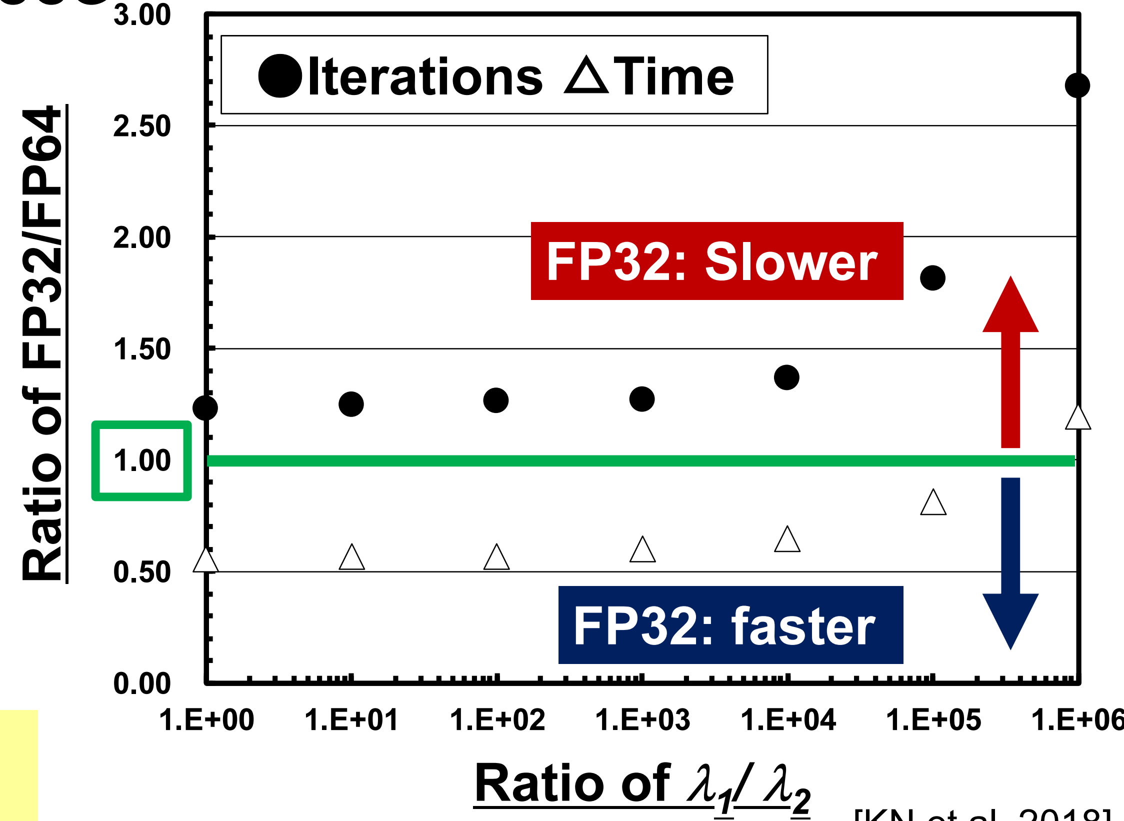
λ_1/λ_2, 128^3 DOF, CRS

Ratio<1 ⇒ FP32 is faster



$$\nabla \cdot (\lambda \nabla \phi) + f = 0$$

Intel Xeon BDW
1 Node: 18 cores x 2 soc's



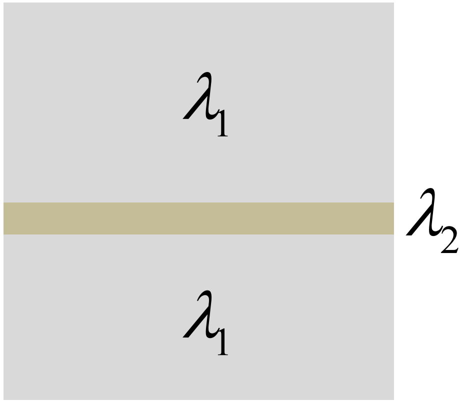
[KN et al. 2018]

Ratio of FP32(SP)/FP64(DP)

Iterations● & Time△ for ICCG

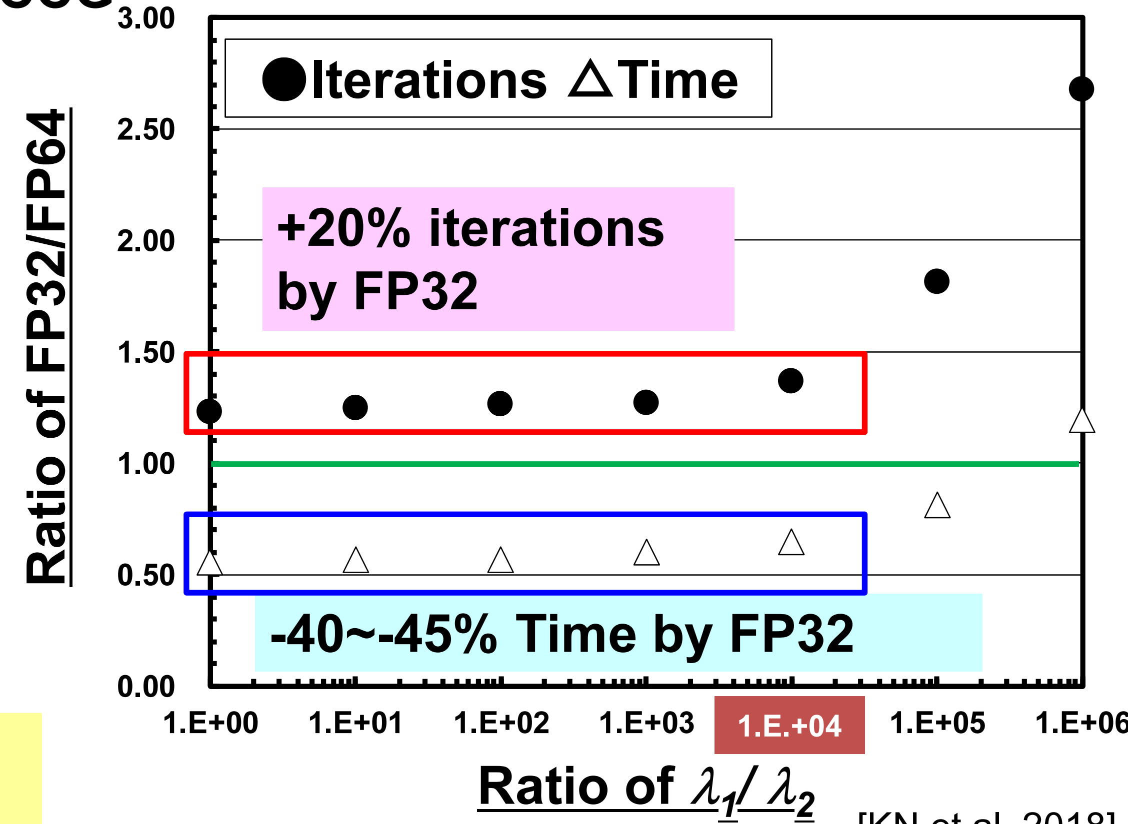
λ_1/λ_2, 128^3 DOF, CRS

Ratio<1 ⇒ FP32 is faster



$$\nabla \cdot (\lambda \nabla \phi) + f = 0$$

Intel Xeon BDW
1 Node: 18 cores x 2 soc's



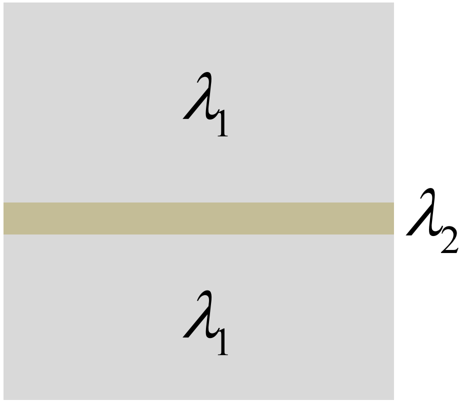
[KN et al. 2018]

Ratio of FP32(SP)/FP64(DP)

Iterations● & Time△ for ICCG

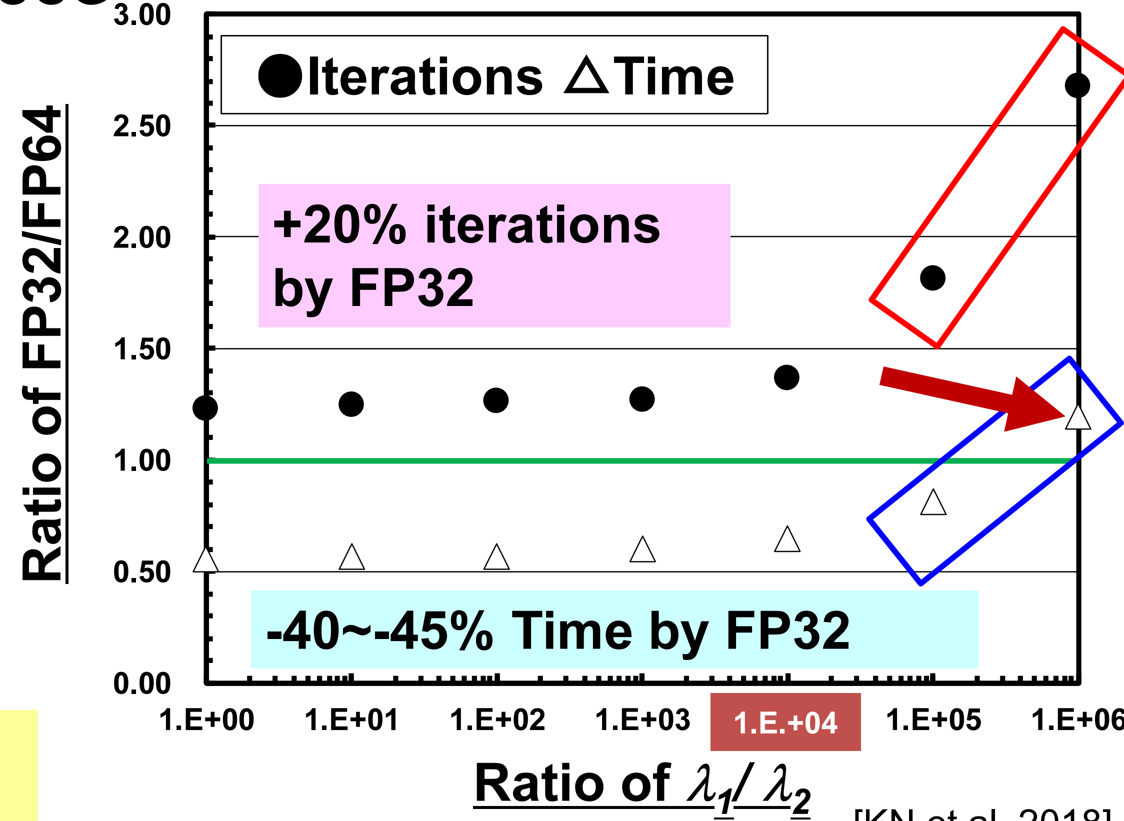
λ_1/λ_2, 128^3 DOF, CRS

Ratio<1 ⇒ FP32 is faster



$$\nabla\cdot(\lambda\nabla\phi)+f=0$$

Intel Xeon BDW
1 Node: 18 cores x 2 soc's

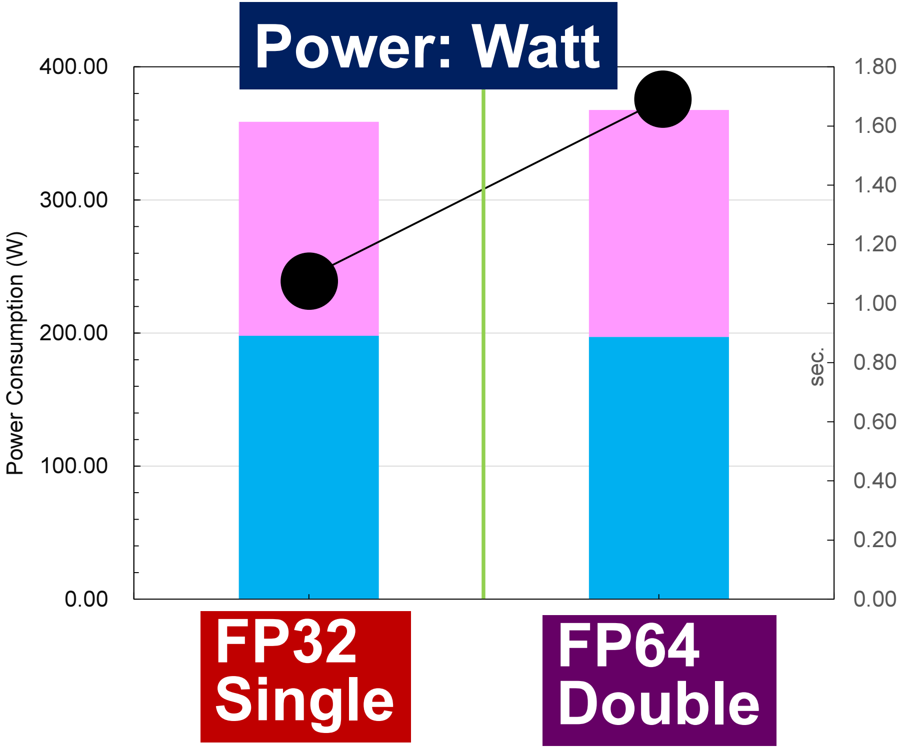


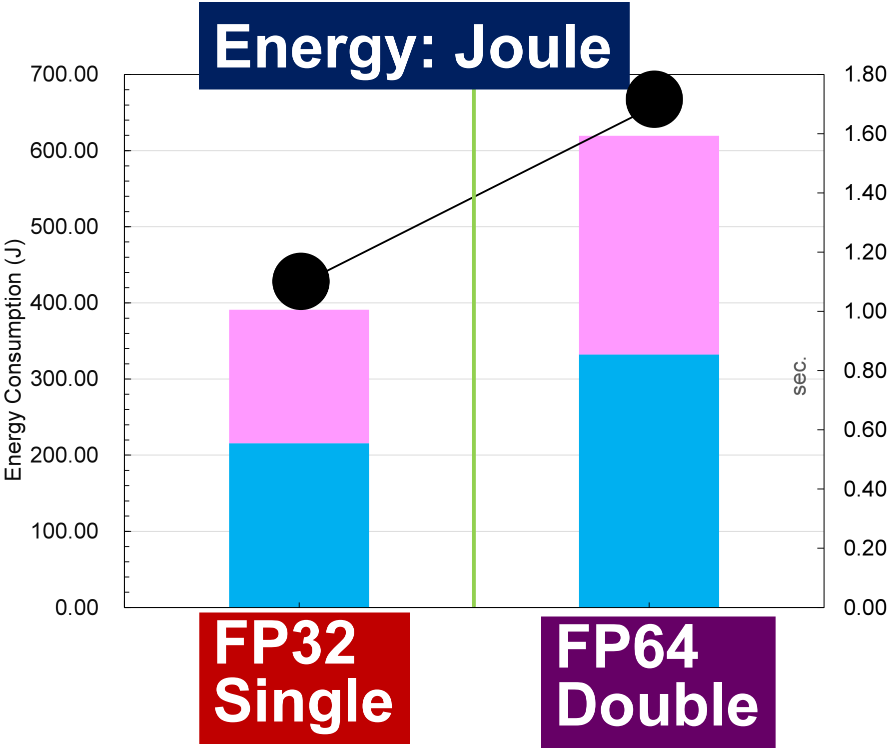
[KN et al. 2018]

Results on Intel Xeon BDW $\lambda_1 = \lambda_2$

[Sakamoto et al. 2020]

$N=128^3$, ■: CPU, ■: Memory, ●: Time





Summary

- ICCG Solver for Heat Conduction Problems by FVM with FP64 (Double Precision) and with FP32 (Single Precision)
  - Time for ICCG, Power Consumption (W), Energy Consumption (J)
- FP64(DP)⇒FP32(SP)
  - Number of iterations increases by 20% and time for ICCG decreases by 40-45% on a single node of Intel Xeon Broadwell with 36-cores if the ratio of λ_1/λ_2 is not larger than 10^4
  - Power consumption (W) does not change
  - Energy consumption (J) is proportional to computation time
- **Results of Accuracy Verification can be found in the separate slides**