An Innovative Method for Integration of Simulation/Data/Learning in the Exascale/Post-Moore Era

Kengo Nakajima
Information Technology Center
The University of Tokyo
RIKEN R-CCS

MS137 Toward Software Ecosystems for CSE
SIAM Conference on Computational Science & Engineering (CSE19)
February 26, 2019, Spokane, WA, USA

Innovative Computing Methods in the Exascale/Post-Moore Era

Kengo Nakajima
Information Technology Center
The University of Tokyo
RIKEN R-CCS

MS137 Toward Software Ecosystems for CSE
SIAM Conference on Computational Science & Engineering (CSE19)
February 26, 2019, Spokane, WA, USA

Acknowledgements

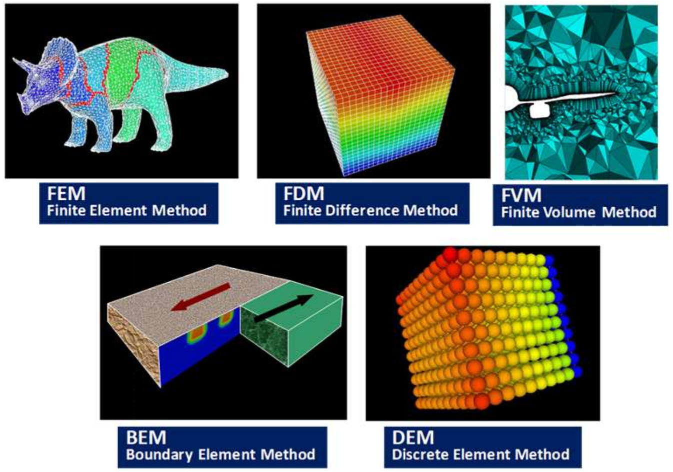
- Sponsors
  - ✓ CREST-JST, Japan
  - ✓ SPPEA-DFG, Germany
- Collaborators, Colleagues
  - ✓ Takeshi Iwashita (Hokkaido U.)
  - ✓ Takahiro Katagiri (Nagoya U.)
  - ✓ Takashi Shimokawabe (ITC/U.Tokyo)
  - ✓ Hisashi Yashiro (RIKEN R-CCS)
  - ✓ Hiroya Matsuba (RIKEN R-CCS)
  - ✓ Hiromichi Nagao (ERI/U.Tokyo)
  - ✓ **Takeshi Ogita (TWCU)**
  - ✓ **Ryuichi Sakamoto (ITC/U.Tokyo)**
  - ✓ Toshihiro Hanawa (ITC/U.Tokyo)
  - ✓ Akihiro Ida (ITC/U.Tokyo)
  - ✓ Tetsuya Hoshino (ITC/U.Tokyo)
  - ✓ Masatoshi Kawai (RIKEN R-CCS)
  - ✓ Takashi Furumura (ERI/U.Tokyo)
  - ✓ Hajime Yamamoto (Taisei)
  - ✓ Gerhard Wellein (Erlangen)
  - ✓ Achim Basermann (DLR)
  - ✓ Osni Marques (LBNL)
  - ✓ Weichung Wang (NTU, Taiwan)

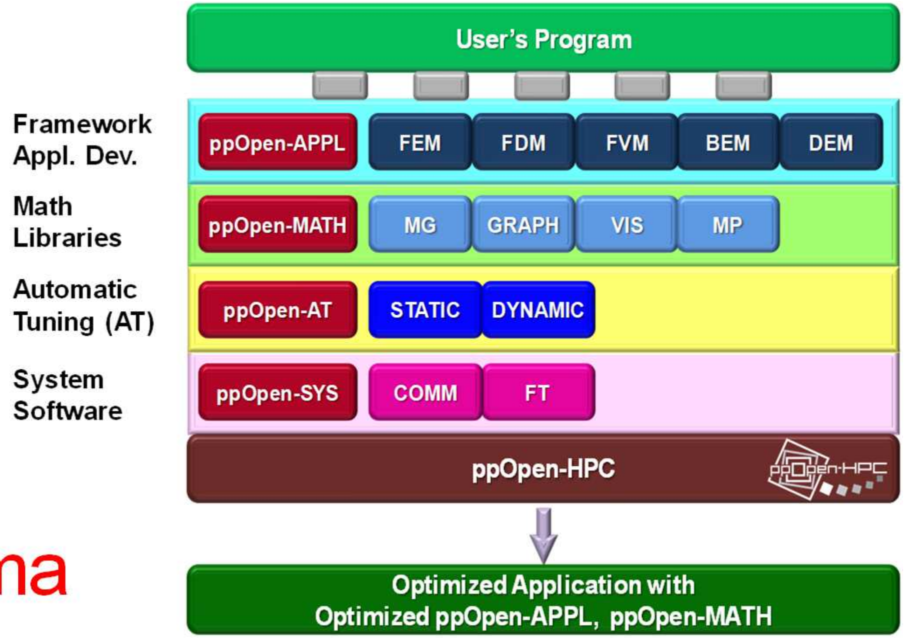
- Background
  - ppOpen-HPC
  - Society 5.0
- BDEC System in ITC/U.Tokyo
- Computing in the Exascale/Post Moore Era
  - Approximate Computing
  - Verification of Accuracy
  - Data Drive Approach
  - h3-Open-BDEC
- Summary

ppOpen-HPC

Application Framework with Automatic Tuning (AT)

- (5+2+α)-year project (FY.2011-2018) (since April 2011) supported by JST/CREST and DFG/SPPEXA
- Team with 7 institutes, >50 people (5 PDs) from various fields: Co-Design

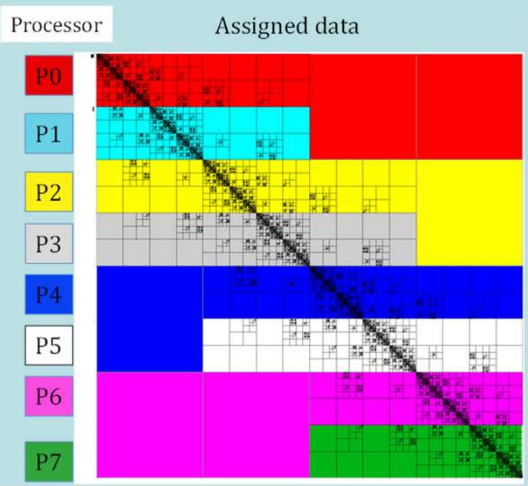




- Leading PI: Kengo Nakajima
- Open Source Software
  - ✓ https://github.com/Post-Peta-Crest/ppOpenHPC
  - ✓ English Documents, MIT License

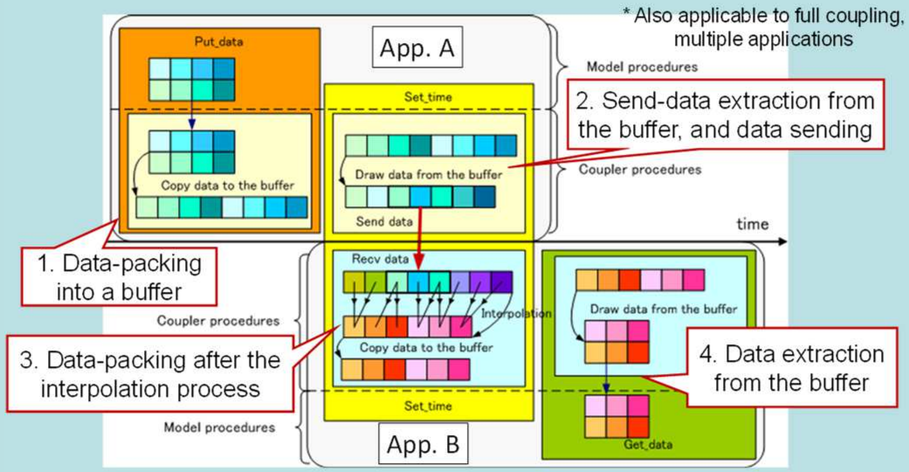
Featured Developments

- ppOpen-AT: AT Language for Loop Optimization
  - Focusing on Optimum Memory Access
- HACApK library for H-matrix comp. in ppOpen-APPL/BEM (OpenMP/MPI Hybrid Version)
  - First Open Source Library by OpenMP/MPI Hybrid
- **ppOpen-MATH/MP (Coupler for Multiphysics Simulations, Loose Coupling of FEM & FDM)**
- Sparse Linear Solvers



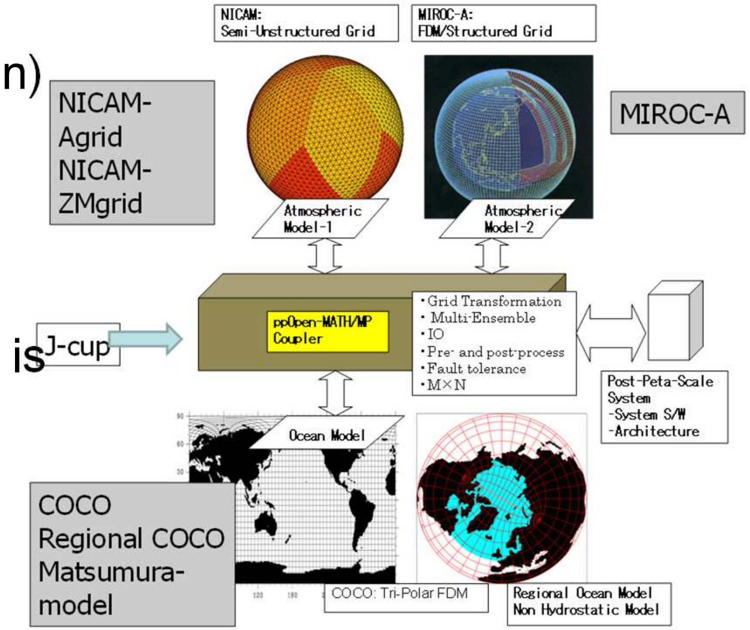
Assign a^c to P_k if $S(a^c{}_{ij})$ in $R(P_k)$

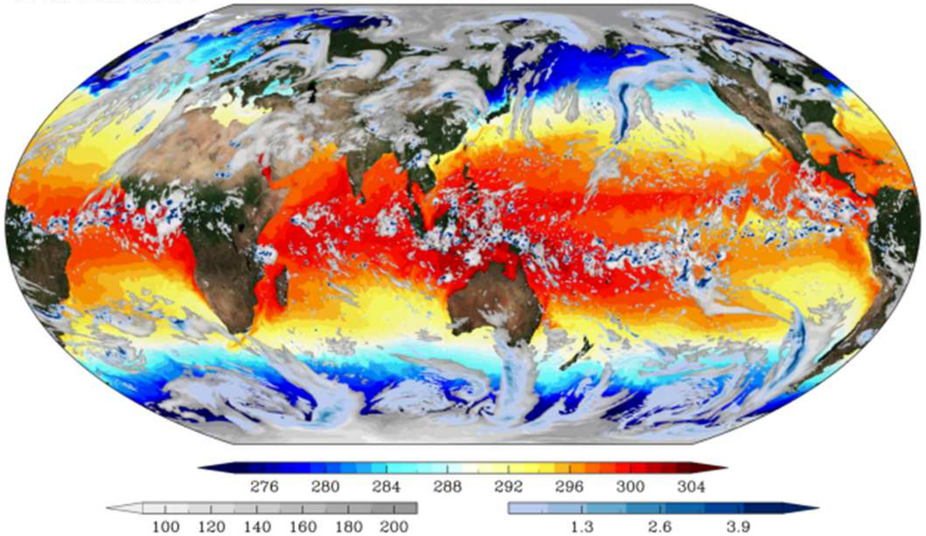
A : Whole matrix
A_{IJ} : The entry in the I-th row and the J-th column of A
a^c: A small submatrix of A
$a^c{}_{ij}$: The entry in the i-th row and the j-th column of a^c
N : Number of rows of A
n : Number of processors
P_k : The k-th processor
$l_k : 1 = l_0 < \cdots < l_k < \cdots < l_{n+1} = N + 1$
$R : R(P_k) = \{i \mid l_k \le i < l_{k+1}\}$
$S : S(a^c{}_{ij}) = I$ when $a^c{}_{ij} = A_{IJ}$

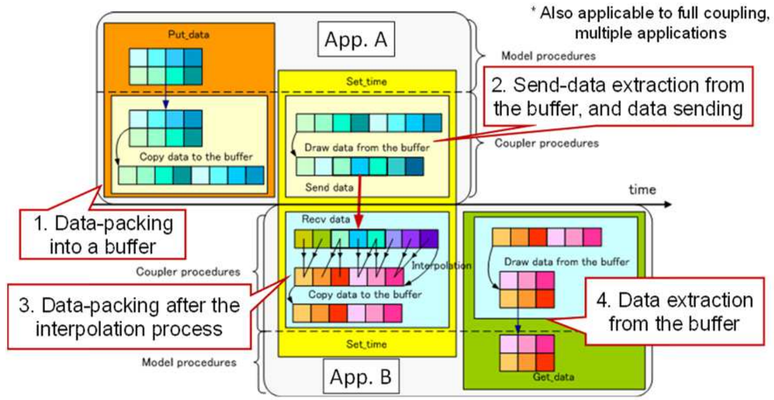


Atmosphere-Ocean Coupling
on OFP by NICAM/COCO/ppOpen-MATH/MP

- High-resolution global atmosphere-ocean coupled simulation by NICAM and COCO (Ocean Simulation) through ppOpen-MATH/MP on the K computer is achieved.
  - ppOpen-MATH/MP is a coupling software for the models employing various discretization method.
- An O(km)-mesh NICAM-COCO coupled simulation is planned on the Oakforest-PACS system.
  - A big challenge for optimization of the codes on new Intel Xeon Phi processor
  - New insights for understanding of global climate dynamics

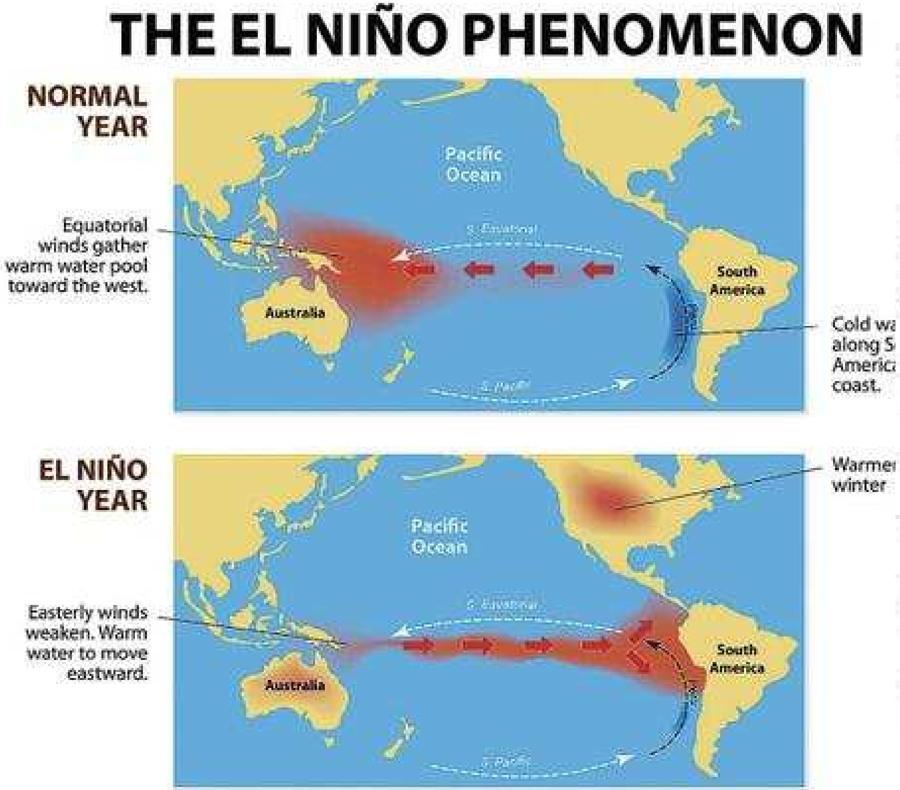


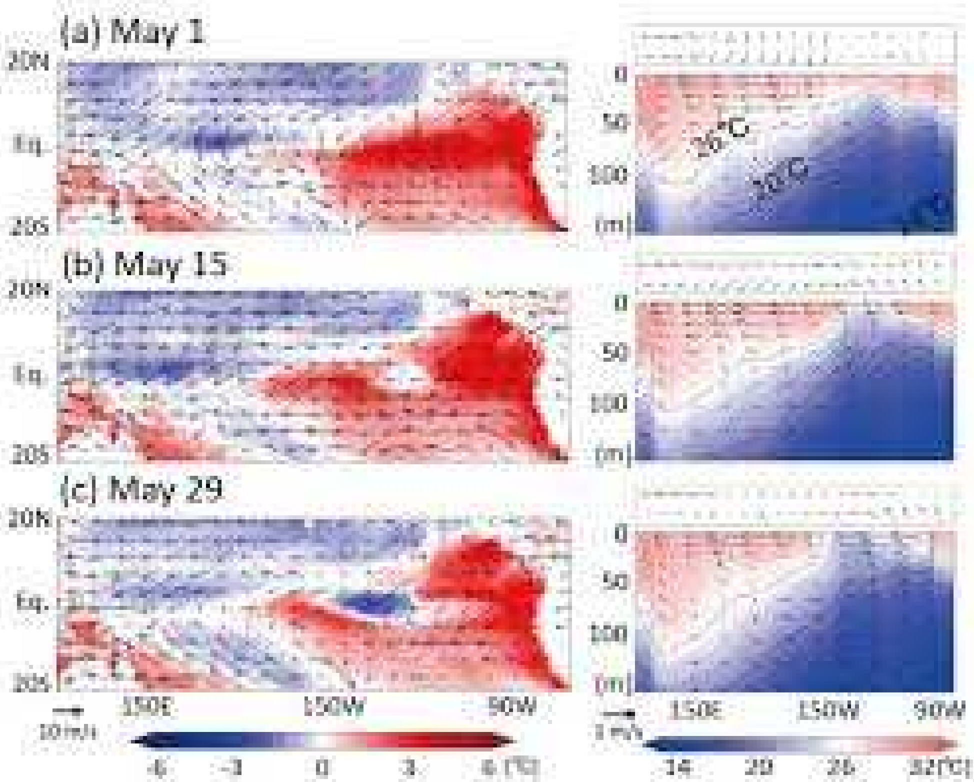




[C/O M. Satoh (AORI/UTokyo)@SC16]

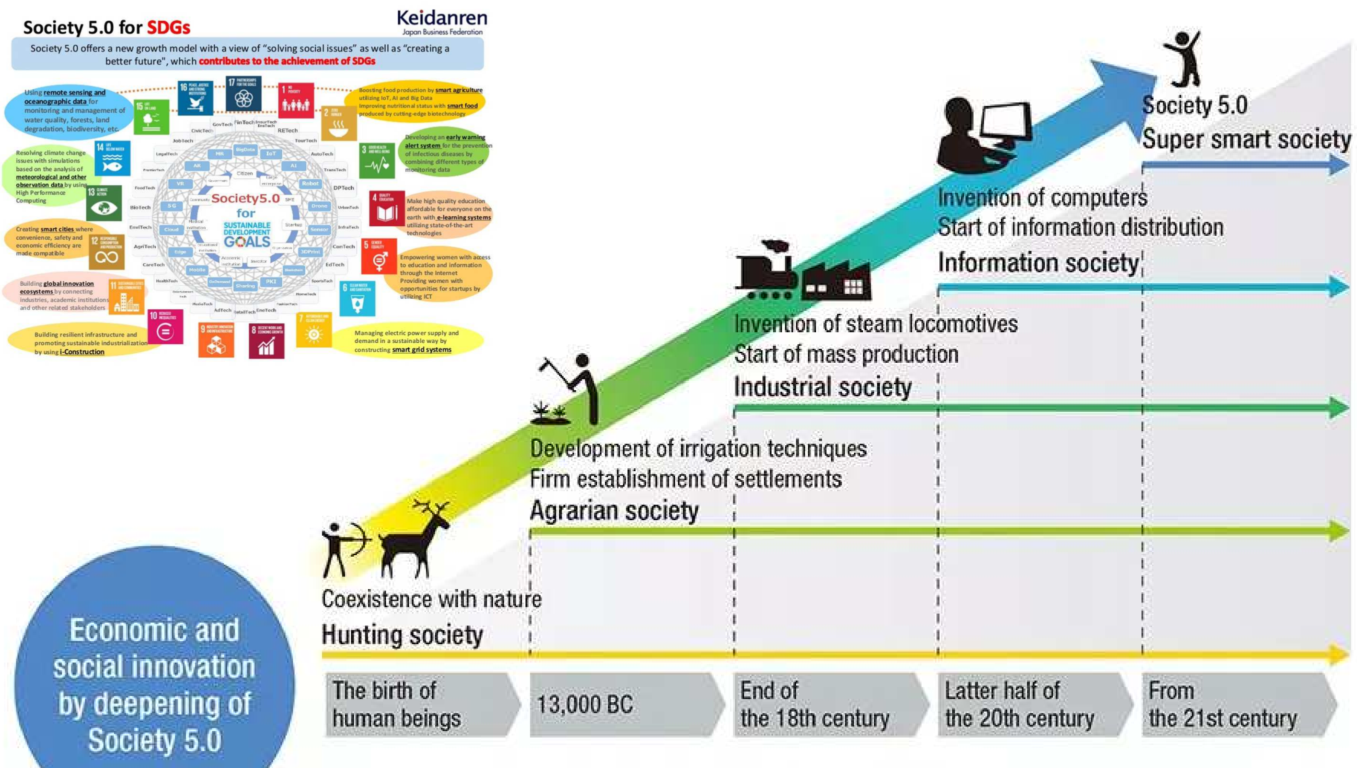
El Niño Simulations
[U.Tokyo, RIKEN September 2017]
Mechanism of the Abrupt Terminate of Super El Niño in 1997/1998 has been revealed by Atmosphere-Ocean Coupling Simulations for the Entire Earth using ppOpen-HPC on the K computer





Society 5.0 (= Super Smart Society) by the Cabinet Office of Japan

- Paradigm Shift towards Knowledge-Intensive & Super Smart Society by Digital Innovation (IoT, AI, Big Data etc.)



Source: Prepared based on materials from the Japan Business Federation (Keidanren)

5.0: Super Smart

4.0: Information

3.0: Industry

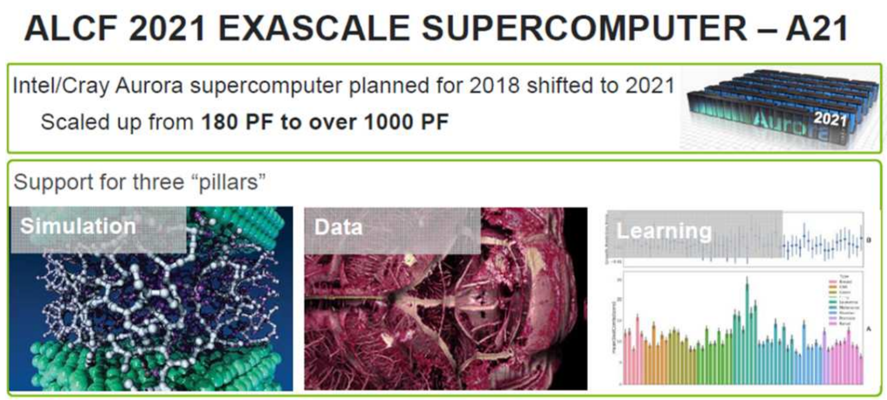
2.0: Agrarian

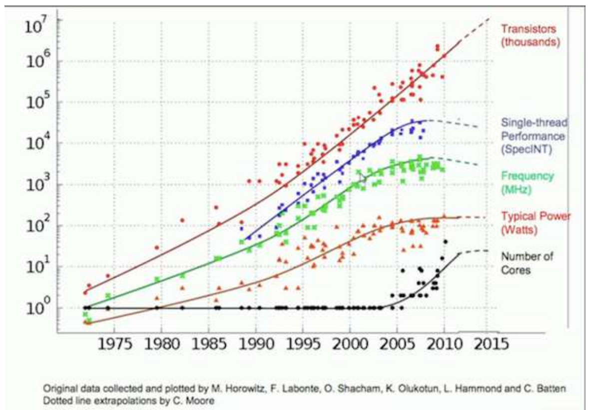
1.0: Hunting

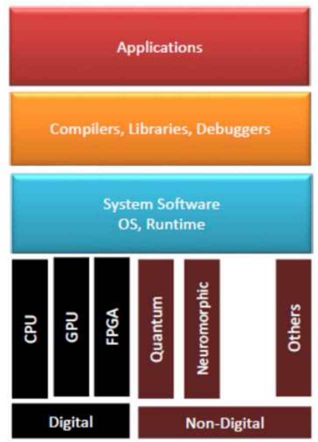
CSE towards Society 5.0 ?

- Integration of CSE, Data and Learning: AI for HPC
  - Simulation + Data + Learning (S+D+L) in A21 of US-DOE
    - The First Exascale System in 2021
  - AI + Big Data + Computing (A+B+C) ?

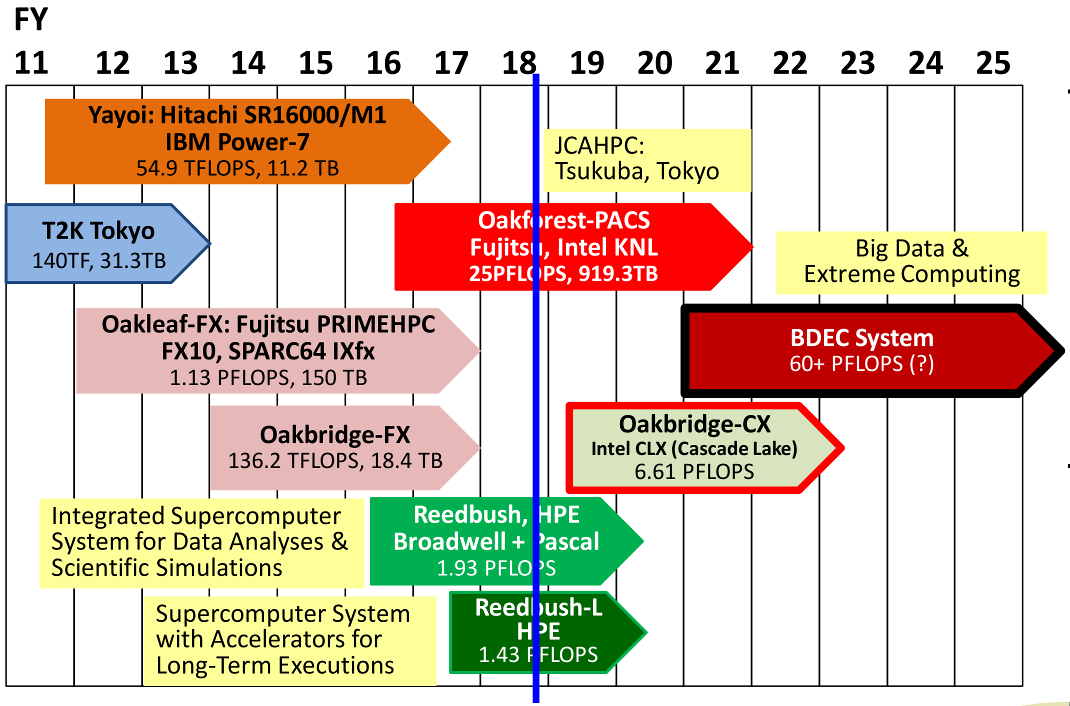
- Power Consumption
  - Important Issue in the Exascale/Post Moore Era
  - Heterogeneous Architecture
  - Various types of HW for Various types of Workload
    - CPU, GPU
    - FPGA
    - Quantum/Neuromorphic
    - Custom Chips







- Background
  - ppOpen-HPC
  - Society 5.0
- **BDEC System in ITC/U.Tokyo**
- Computing in the Exascale/Post Moore Era
  - Approximate Computing
  - Verification of Accuracy
  - Data Drive Approach
  - h3-Open-BDEC
- Summary

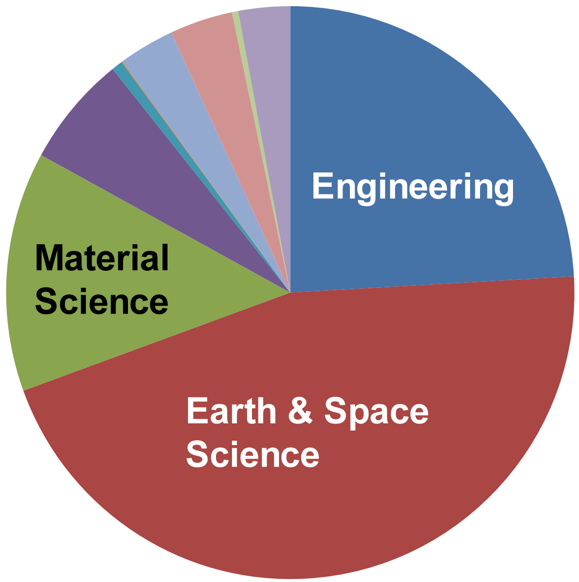


New Types of Users

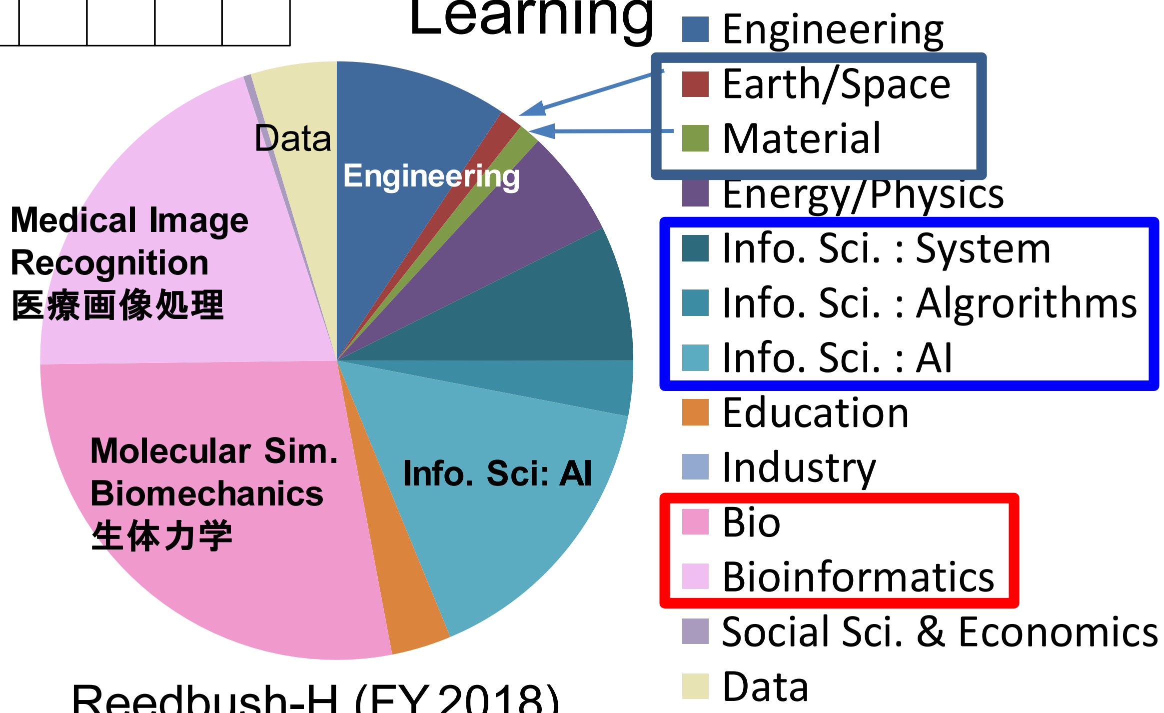
- Mostly CSE, so far
- Data, ML, AI etc.
  - Genome Analysis
  - Medical Image Recognition

New Methods

- Integration of CSE (Simulations) + Data + Learning



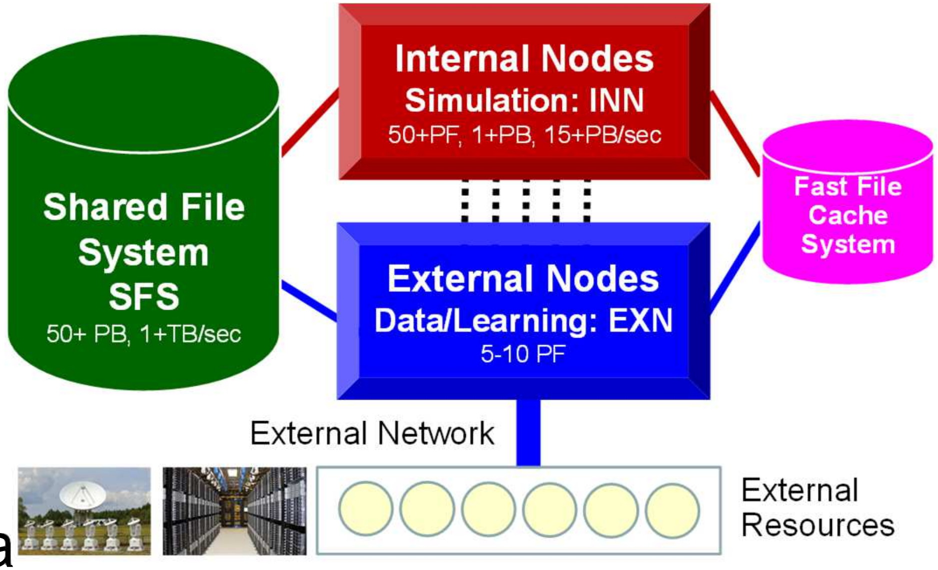
Oakleaf/Oakbridge-FX (FY.2017)
Commercial Version of K



Reedbush-H (FY.2018)
Intel BDW + NVIDIA P100

BDEC System at ITC/U.Tokyo

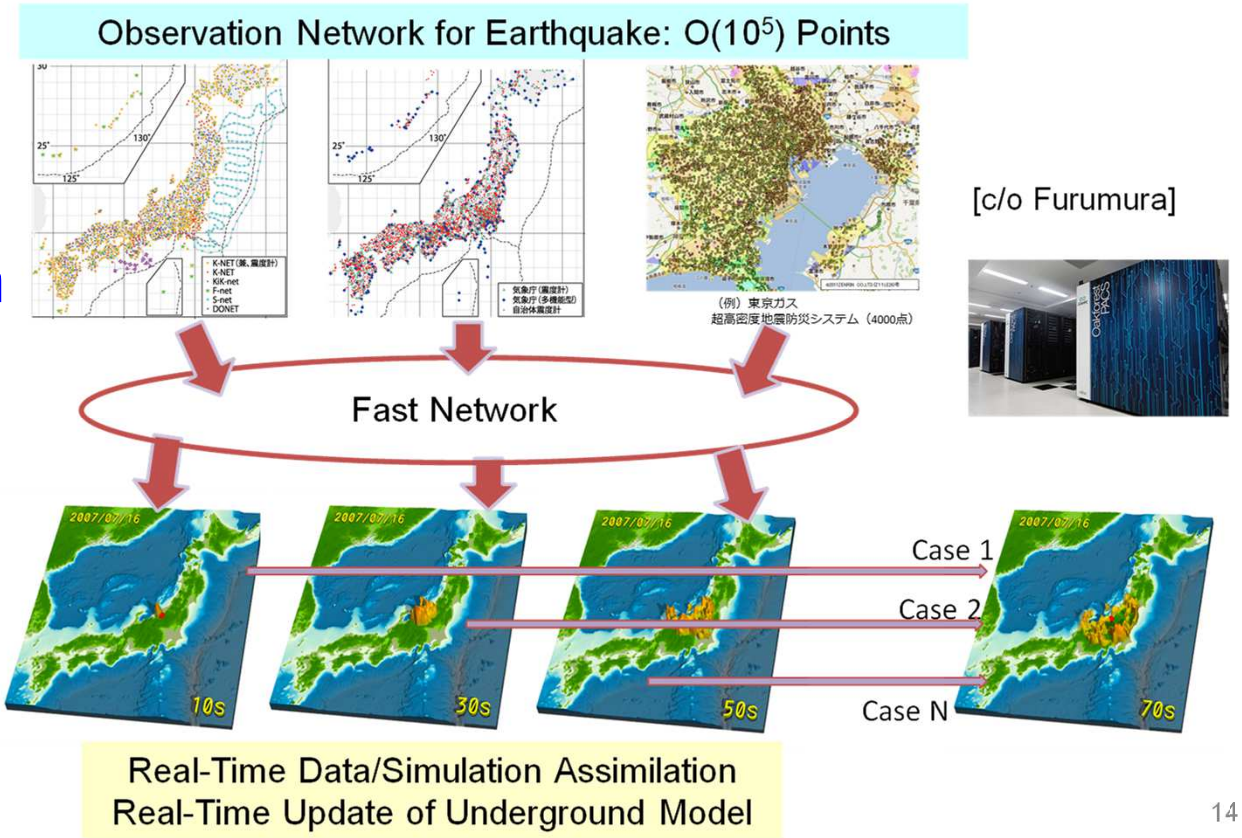
- Platform for (S+D+L)
  - Big Data & Extreme Comp.
- April 2021
- 60+ PF, 3.5-4.5 MW
  - External Nodes for Data Acquisition/Analysis (EXN)
    - 5-10 PF, 200+ TB
  - Internal Nodes for CSE/Data Analysis (INN)
    - 50+ PF, 1+ PB, 15+ PB/sec.
  - Shared File System (50+PB, 1+TB/sec) + File Cache
- Architectures of EXN and INN could be different
  - EXN could include GPU, FPGA, Quantum Device, and <u>more flexible</u>

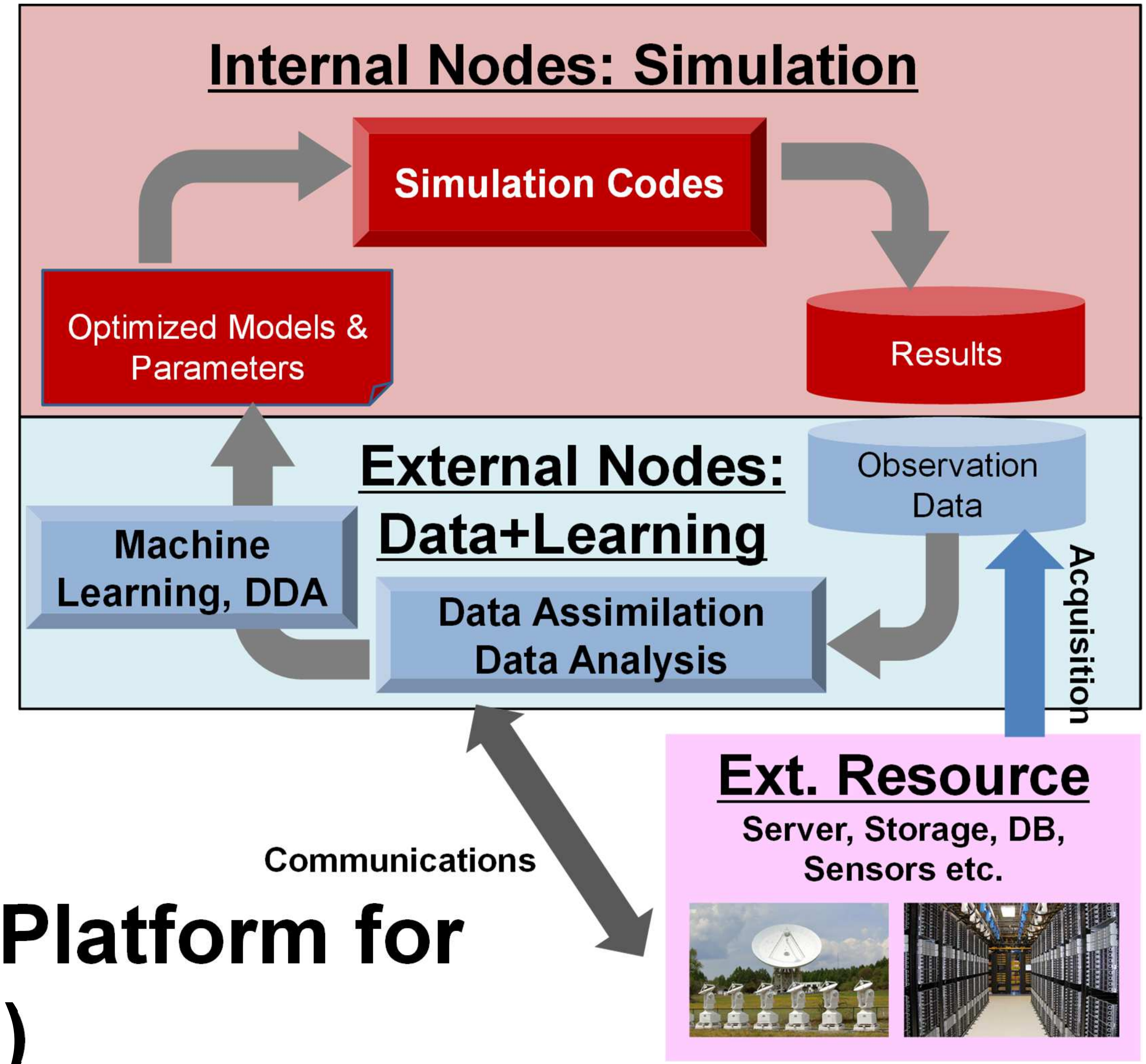


- Possible Applications
  - Atmosphere-Ocean Simulations with Data Assimilation
  - Real-Time Disaster Sim. (Flood, Earthquakes, Tsunami)
    - Earthquake Simulations with Data Assimilation
  - Data Driven Approach

Real-Time Earthquake Simulation with Data Assimilation

- Seismic Observation Data (100Hz/3-dir's/O(10^3) pts) by JDXnet is available through SINET <u>in Real Time</u>
  - Peta Server in ERI/U.Tokyo: O(10^2) GB/day⇒EXN of BDEC
  - O(10^5) pts in future including stations operated by industry
- External Nodes
  - Real-Time Data Acquisition
  - Data Assimilation
  - Update of Underground Model
- Internal Nodes
  - Large-Scale Multiple Simulations





BDEC: Platform for (S+D+L)

- Background
  - ppOpen-HPC
  - Society 5.0
- BDEC System in ITC/U.Tokyo
- **Computing in the Exascale/Post Moore Era**
  - **Approximate Computing**
  - **Verification of Accuracy**
  - Data Drive Approach
  - h3-Open-BDEC
- Summary

Computing in the Exascale/Post Moore Era

- Power Consumption is the Most Important Issue in the Post Moore Era
  - It is already important now.
  - Memory performance in the Post Moore Era is relatively better than now, but data movement should be reduced from the view point of energy consumption.
- Quantum Computing, FPGA ?: "Partial" Solution
  - Could be a solution in certain applications (e.g. searching, graph, data clustering etc.)
  - Contributions to (S+**D+L**)/(A+**B+C**)
- **How to save Energy for Sustainability ?**
  - **(1) Approximate Computing by Low/Adaptive Precision**
  - **(2) Reduction of Computations: Data Driven Approach**

Approximate Computing with Low/Adaptive/Trans Precision

- **Lower Precision: Save Time & Energy & Memory**
- Approximate Computing: originally for image recognition etc.
  - **Approach for Numerical Computations**
    - **SIAM PP18 Sessions, ICS-HPC 2018 Workshop**
  - OPRECOMP: Open transPREcision COMPuting (Horizon 2020)
- **Computations with Low Precision**
- **Mixed Precision Approach (FP16-32-64-128)**
- **Iterative Refinement**
  - **such computations may provide results with less accuracy**

Numerical Library with High-Performance/Adaptive-Precision/High-Reliability

Extension of ppOpen-HPC towards the Post Moore Era

JHPCN

- Lower/Adaptive Precision + Accuracy Verification
  - Collaboration with "Pure" Applied Mathematicians
  - Iterative Refinement
- Automatic Tuning (AT): Selection of the optimum precision, which minimizes computation time and power consumption under certain target accuracy
  - implemented to "ppOpen-HPC".
- Preconditioned Iterative Solvers for Practical Problems with Ill-Conditioned Matrices with Adaptive Precision
  - FP16-32-64-128
- Staring from April 2018, as a part of JHPCN Project in Japan (Preliminary Works in FY.2018)

Numerical Library with High-Performance/Adaptive-Precision/High-Reliability

20+ Members from 13 Institutions (Japan, Germany)

JHPCN

北海道大学 HOKKAIDO UNIVERSITY

東京大学 THE UNIVERSITY OF TOKYO

東京工業大学 Tokyo Institute of Technology

Tokyo Woman's Christian University 東京女子大学

芝浦工業大學 SHIBAURA INSTITUTE OF TECHNOLOGY Established 1927

名古屋大学

九州大学 KYUSHU UNIVERSITY

RIST

R-CCS

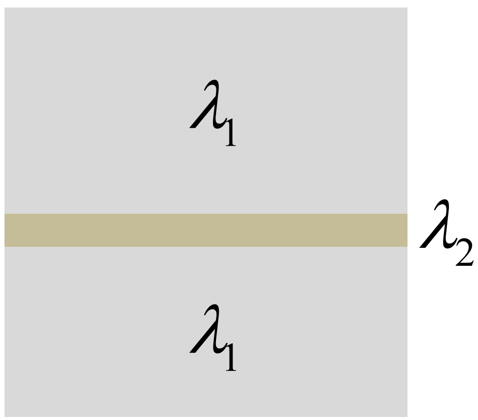
FAU FRIEDRICH-ALEXANDER UNIVERSITÄT ERLANGEN-NÜRNBERG

intel

NVIDIA

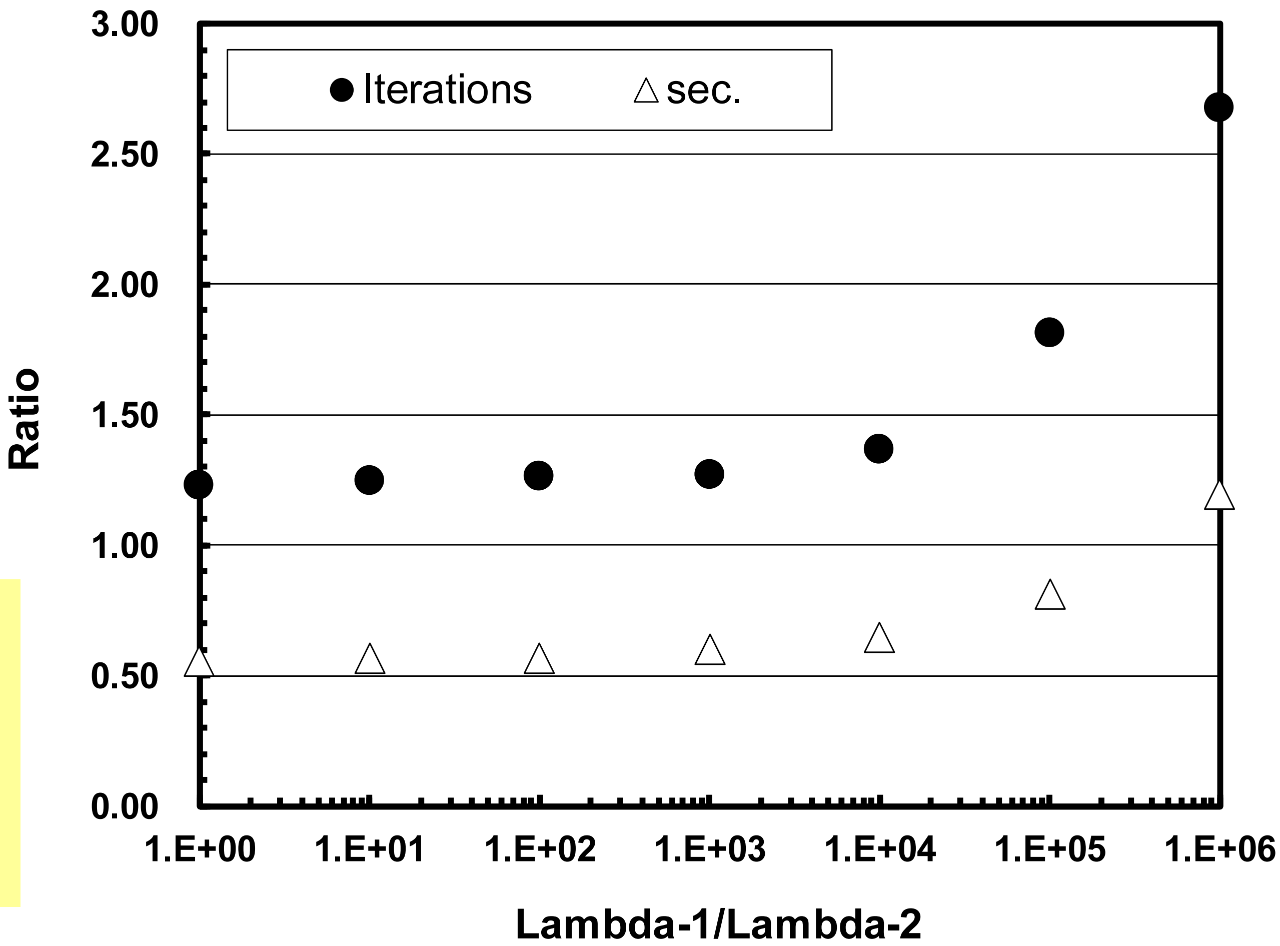
DLR Deutsches Zentrum für Luft- und Raumfahrt German Aerospace Center

Results: λ_1/λ_2~ Condition Number Ratio of Iterations & Computation Time Single/Double: Down is Good

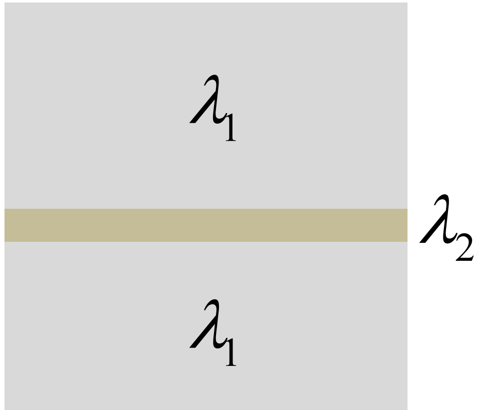


$$\nabla \cdot (\lambda \nabla \phi) + f = 0$$

Intel Xeon BDW
Single Node:
18cores x 2soc's
(Reedbush-U)

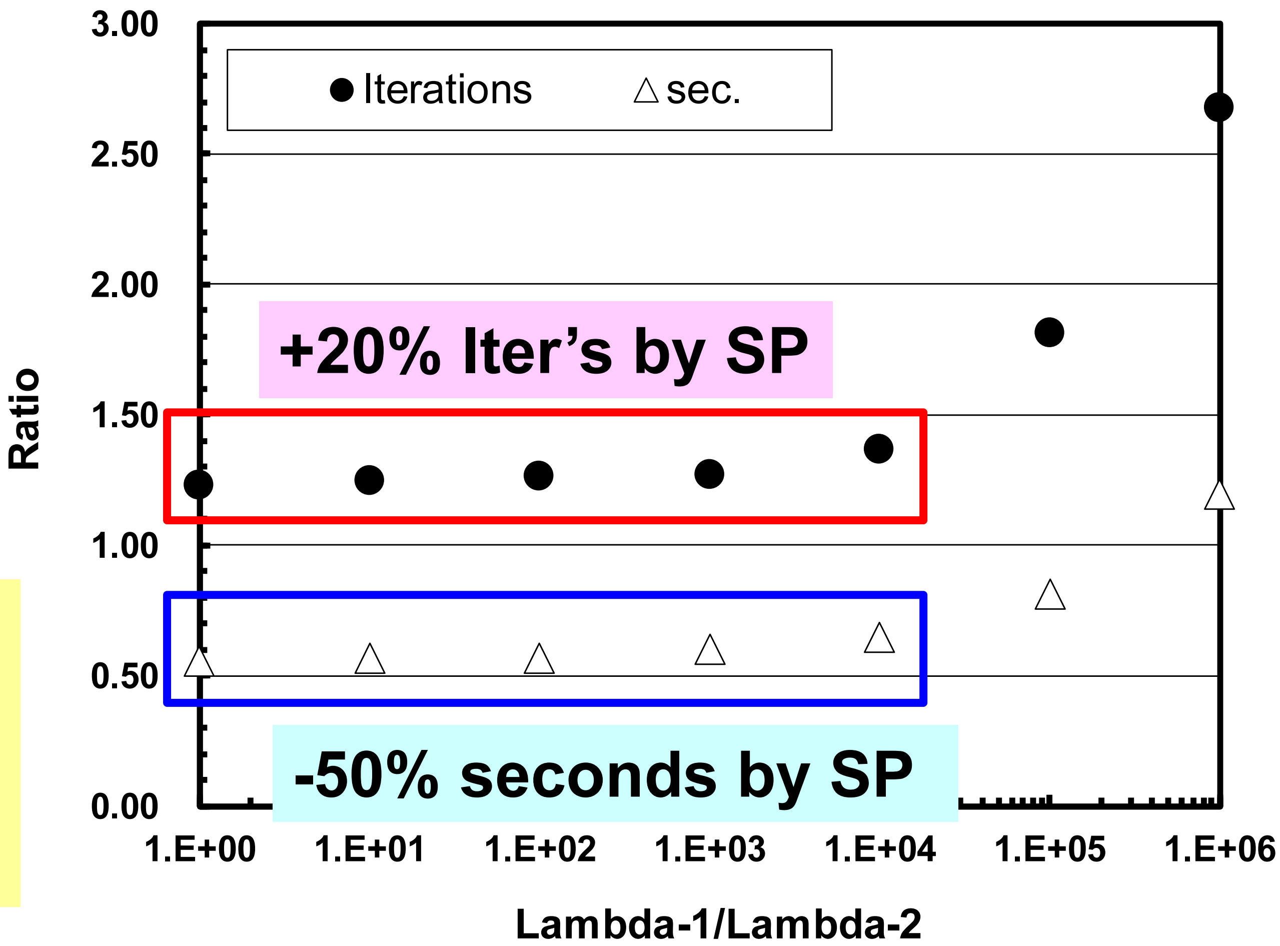


Results: λ_1/λ_2~ Condition Number Ratio of Iterations & Computation Time Single/Double: Down is Good



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

Intel Xeon BDW
Single Node:
18cores x 2soc's
(Reedbush-U)

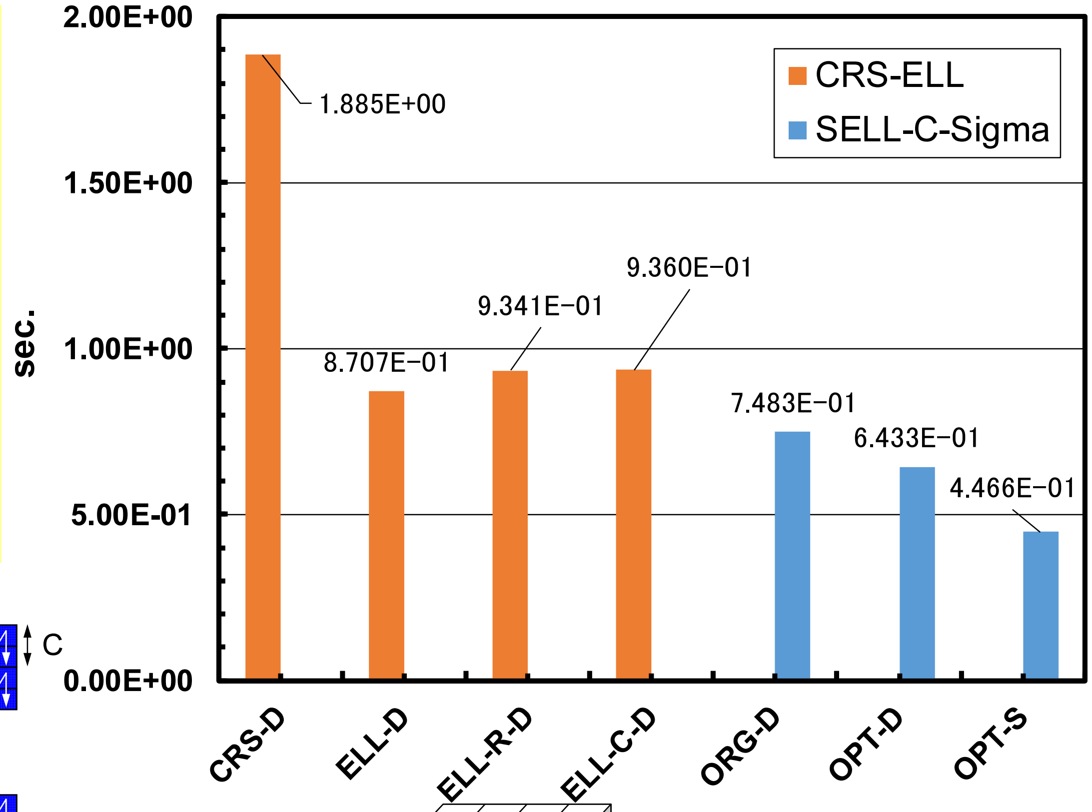


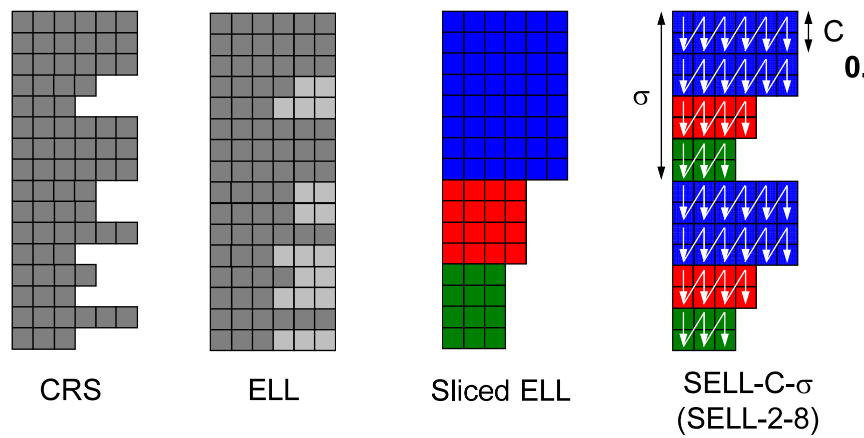
ICCG: ELL/Sliced ELL/SELL-C-σ

$\lambda_1 = \lambda_2$

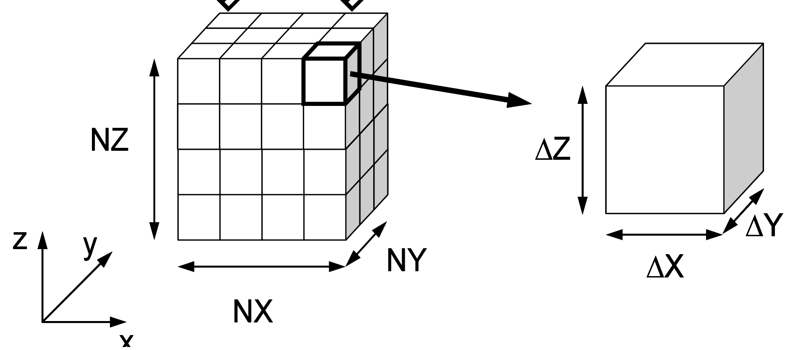
ICCG Solvers on
Intel Xeon/Phi (KNL)
(Oakforest-PACS)
Single Node:
64/68 cores

SELL-C-σ for ICCG

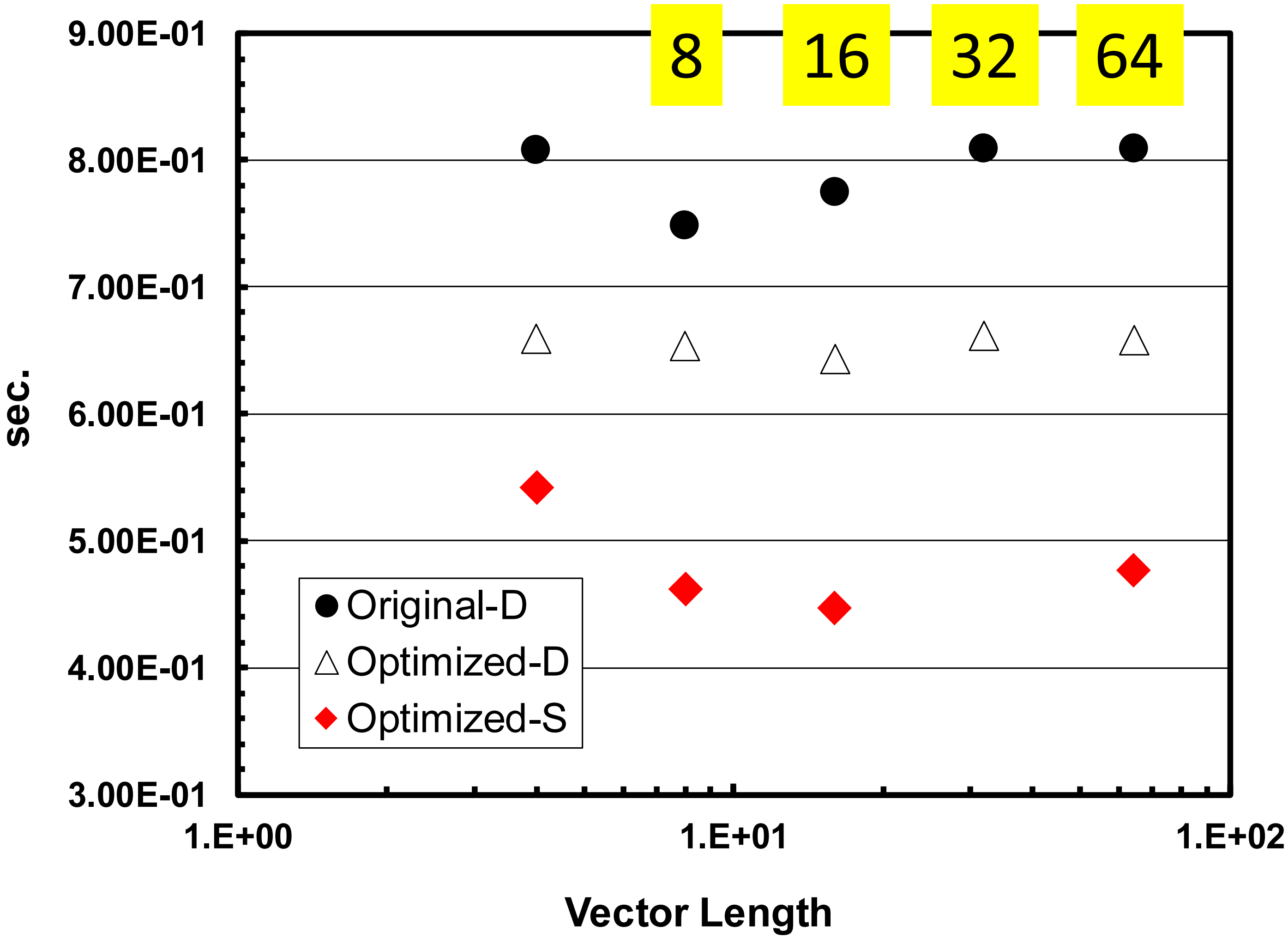




[Kreutzer, Wellein et al. 2014]

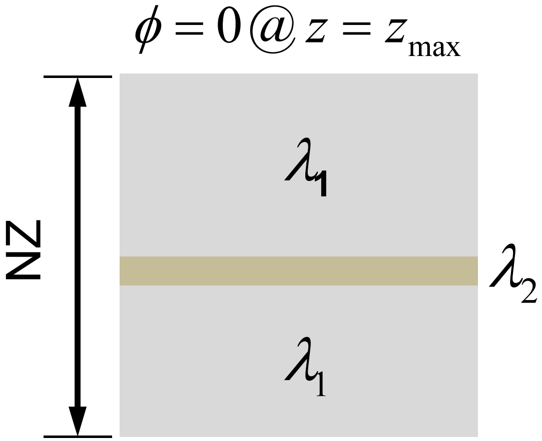


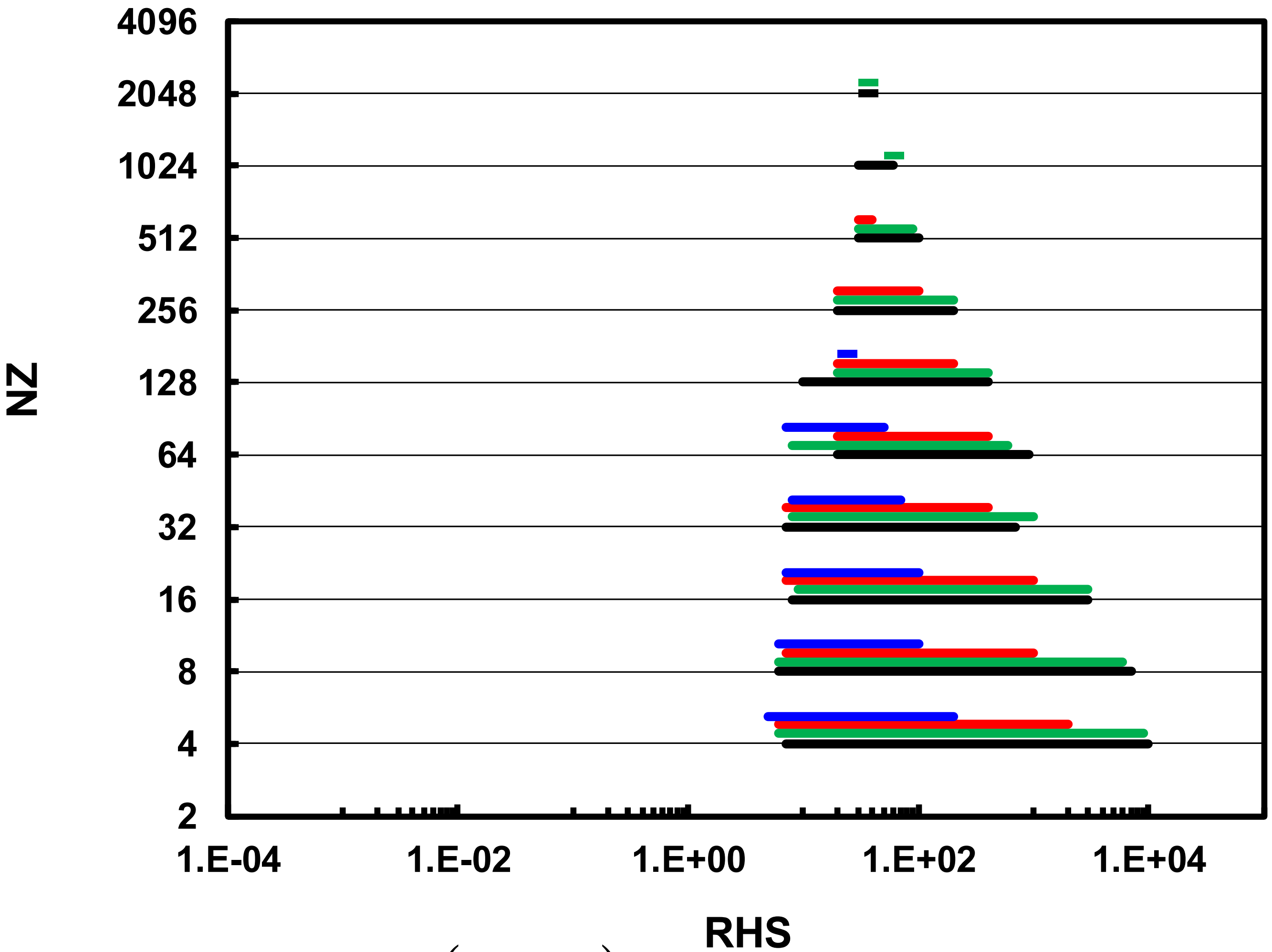
Results on OFP, Poisson-3D-OMP Effect of SIMD Vector Length in SELL-C-σ 10 colors, 128^3



FP32 (Single) with FP16 Precond.

V100, All Problems converge in FP32/64





λ_1/λ_2=1.0e+3

λ_1/λ_2=1.0e+2

λ_1/λ_2=1.0e+1

λ_1/λ_2=1.0e+0

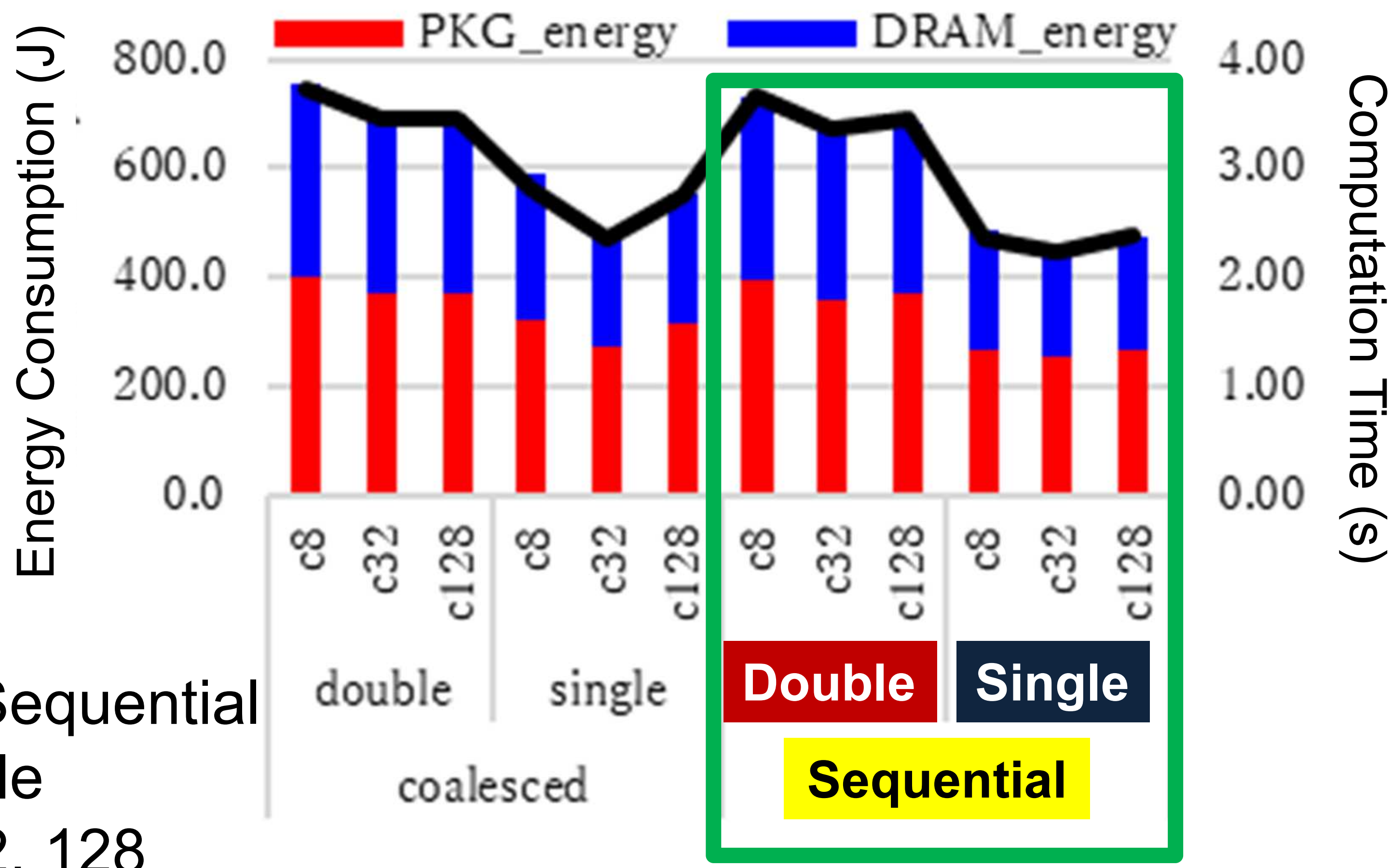
[Hoshino 2018]

$$\nabla \cdot \left(\lambda \nabla \phi \right) = -RHS$$

3D Poisson Solvers on Reedbush-H

$\lambda_1 = \lambda_2$

CPU only: Intel BDW: sec. & Joule

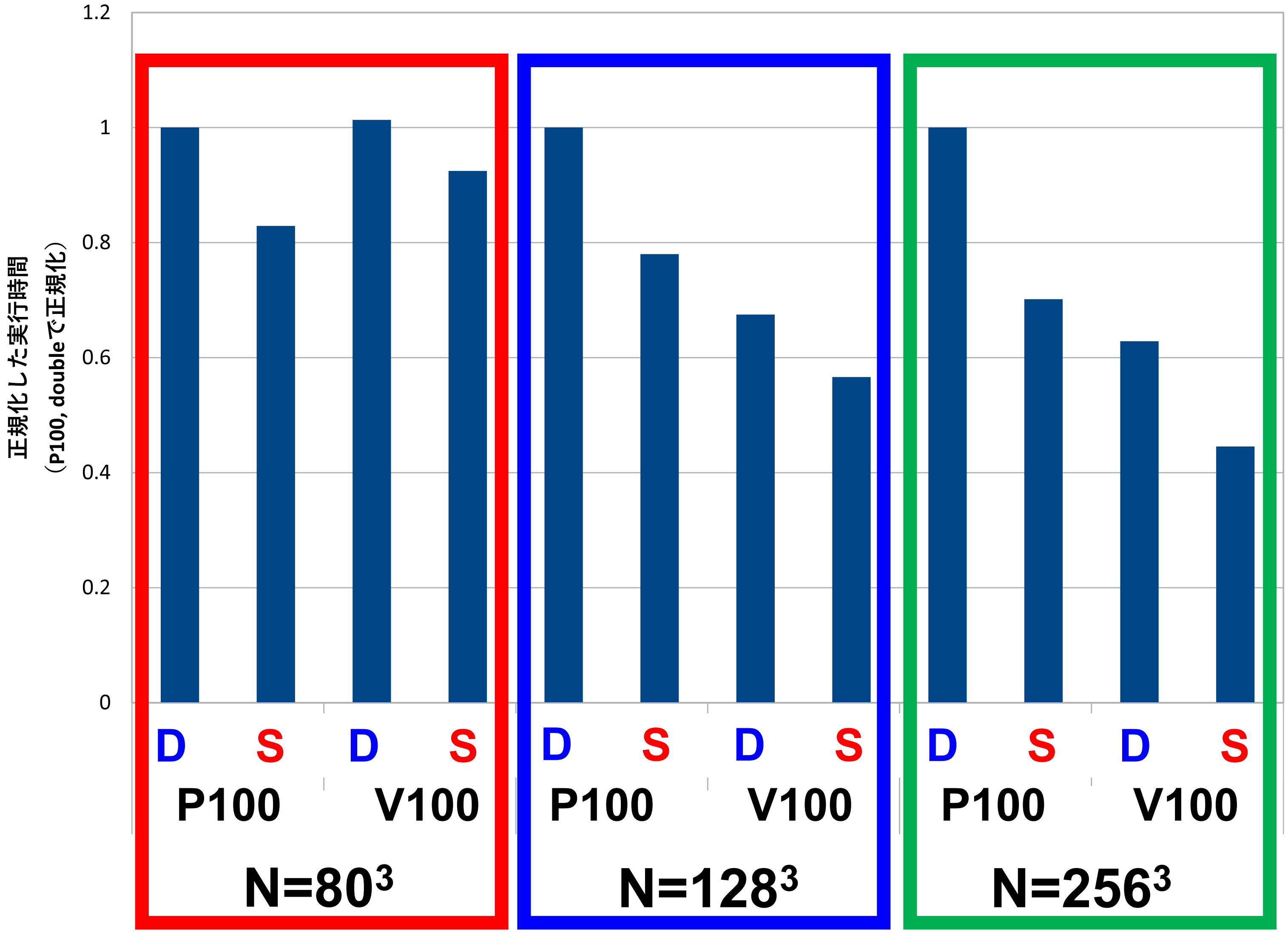


- 128^3 DOF
- Coalesced/Sequential
- Double/Single
- Colors: 8, 32, 128
- **Watt-value of SP may increase due to larger density of comp.**

[Sakamoto et al. 2018]

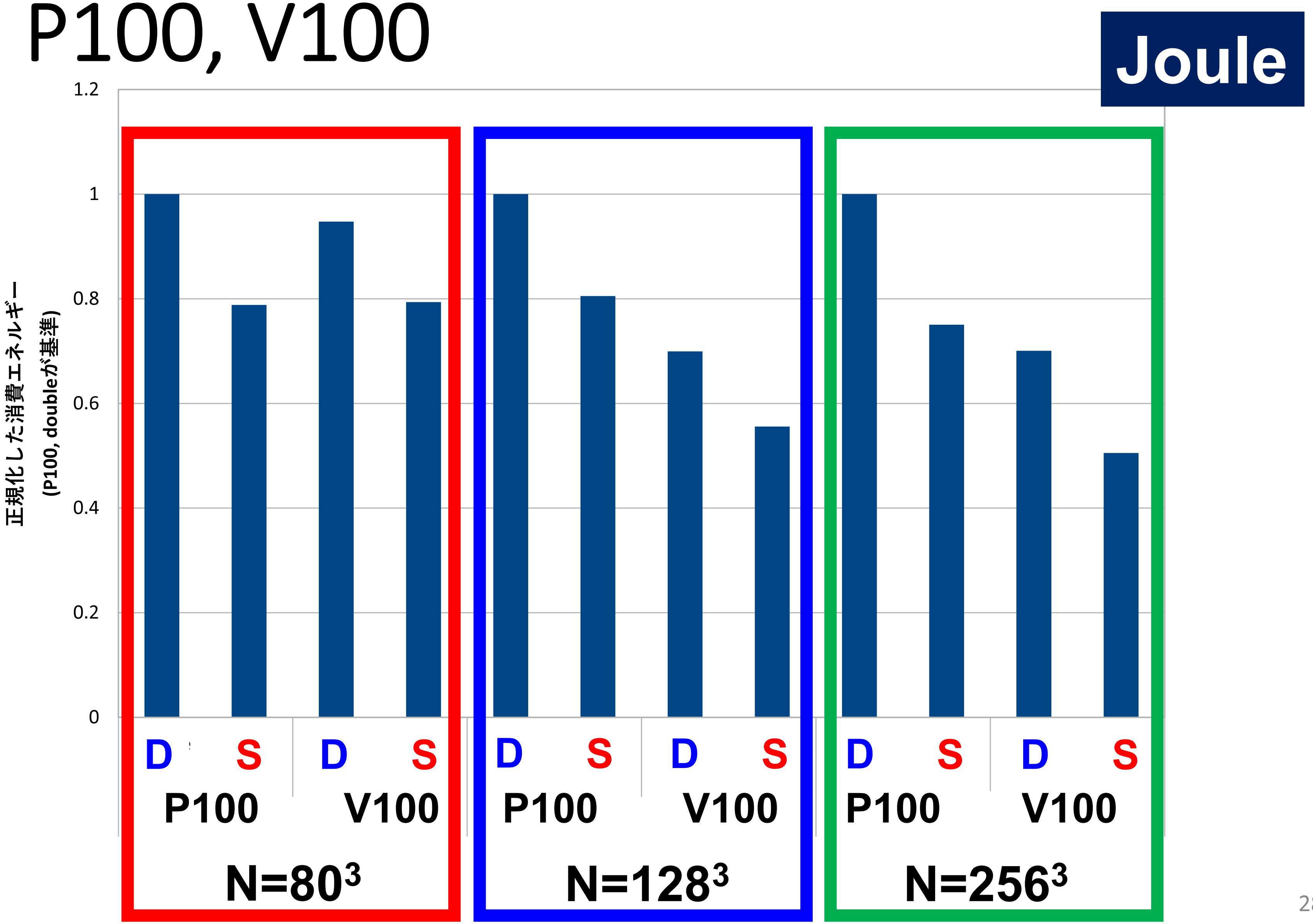
Computation Time (Normalized): P100, V100

[Sakamoto et al. 2018]



[Sakamoto et al. 2018]

Energy Consumption (Normalized): P100, V100



Approximate Computing with Low/Trans Precision

- Accuracy verification is important
  - Iterative Refinement
- A lot of methods for accuracy verification have been developed for problems with dense matrices
  - But very few examples for sparse matrices & H-matrices
- Generally speaking, processes for accuracy verification is very expensive
  - Sophisticated Method needed
  - Automatic Selection of Optimum Precision by Technology of AT (Auto Tuning)

Accuracy Verification of Sparse Linear Solver (1/2)

[Ogita, Ushiro, Oishi 2001]

1. Solve $Ax = b$ where $\tilde{x}$ is the numerical solution
2. Calculate upper bound of $\left\| A^{-1} \right\|_\infty$
3. Calculate lower/upper bound of $r = A\tilde{x} - b \Rightarrow r_{low}$ and r_{upp} (in higher preecision)
4. Solve $Az = r_{low}$ and/or $Az = r_{upp}$
5. Calculate upper bound of absolute error: $\varepsilon_{abs} \geq \left\| \tilde{x} - x^* \right\|_\infty$ (x^* : exact solution of $Ax = b$)
5. Calculate upper bound of relative error: $\varepsilon_{rel} \geq \dfrac{\left\| \tilde{x} - x^* \right\|_\infty}{\left\| x^* \right\|_\infty}$

Special Method for Rather Well-Conditioned Matrices (M-Matrix)

If "monotone" matrix A satisfies $\left\| A\tilde{y} - e \right\|_\infty < 1$ where $e = (1,1,\ldots,1)^T$ and $\tilde{y}$: solution of $Ay = e$

$$\left\| A^{-1} \right\|_\infty \leq \frac{\left\| \tilde{y} \right\|_\infty}{1 - \left\| A\tilde{y} - e \right\|_\infty}$$

Verification Algorithm

1. Solve a discretized linear system $Ax = b$.
   - $\hat{x}$: a computed solution
2. Solve a linear system $Ay = e$ where all elements of e are 1's.
   - $\hat{y}$: a computed solution
3. Verify M-property of A using $\hat{y}$. ($\hat{y} > 0 \Rightarrow A\hat{y} > 0$)
4. Compute an error bound using

$$\|x - \hat{x}\|_\infty \leq \frac{\|\hat{y}\|_\infty \|b - A\hat{x}\|_\infty}{1 - \|e - A\hat{y}\|_\infty}$$

if $\|e - A\hat{y}\|_\infty < 1$.

Numerical Results

- Computer: Reedbush-U (1 node)
  - Intel Xeon E5-2695v4 (Broadwell-EP, 2.1GHz 18 cores) x 2 sockets
  - 1.21 TFLOP/s per socket, 256 GiB (153.6GB/s)
- Solver: ICCG with CM-RCM, MC(20)
- Stopping criteria:
  - For $Ax = b$, $\frac{\|b - A\hat{x}\|_2}{\|b\|_2} < 10^{-12}$
  - For $Ay = e$, $\|e - A\hat{y}\|_\infty < 10^{-2}$
- FP64 (double precision), OpenMP (36 threads)

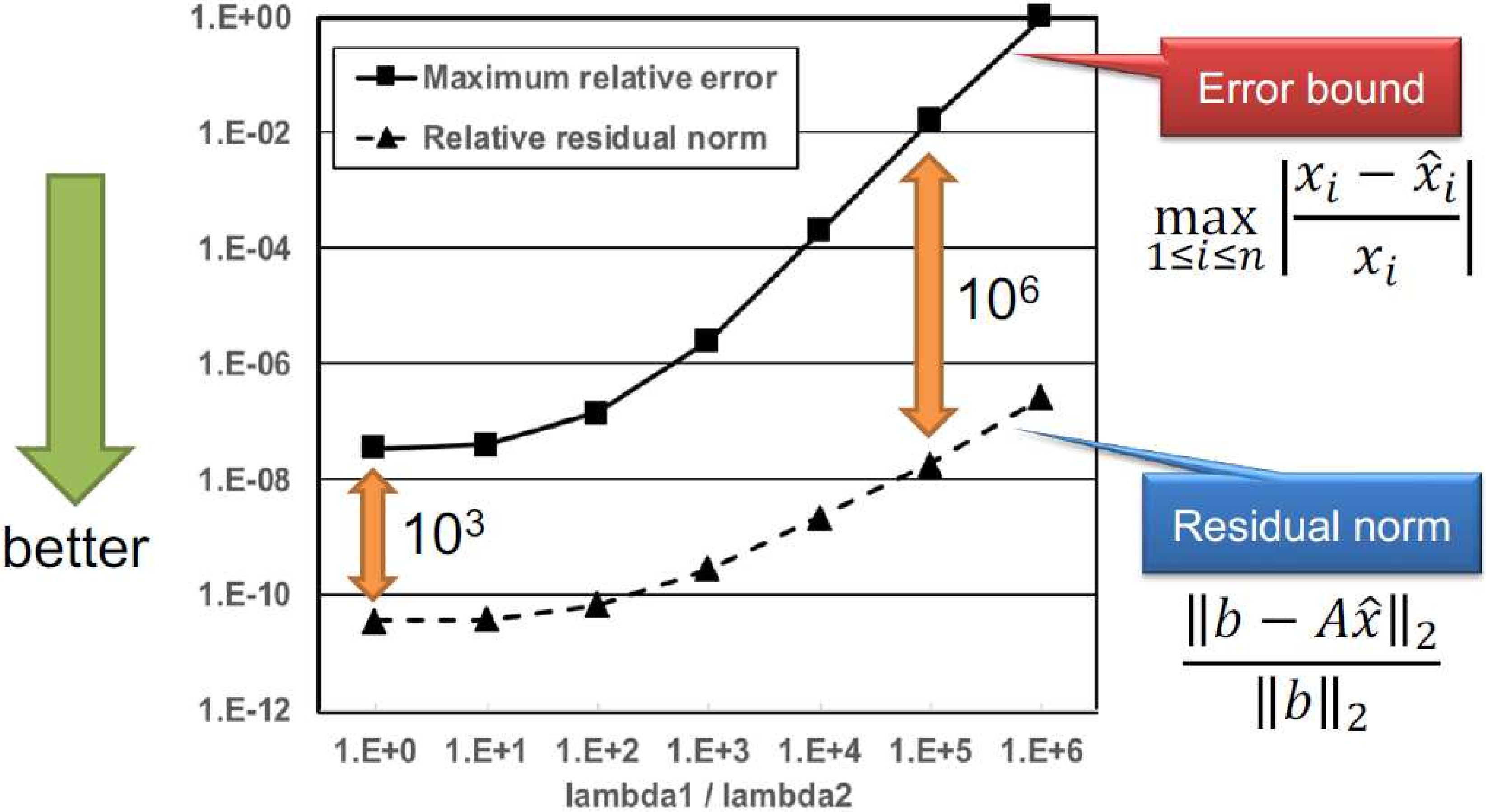
Result (1): $\lambda_1 = \lambda_2 = 1.0$
NX=NY=NZ=128 (n = 2,097,152)

- Upper bounds of maximum relative error and relative residual norm:
  - $\max_{1 \le i \le n} \left| \frac{x_i - \hat{x}_i}{x_i} \right| \le 3.38 \times 10^{-8}$
  - $\frac{\|b - A\hat{x}\|_2}{\|b\|_2} < 3.66 \times 10^{-11}$
- Computing time

	Approximation Solve Ax=b (415 iter's)	Verification-1 Solve Ay=e (211 iter's)	Verification-2	Total
Method-1	2.38	1.18		3.56
Method-2 (2 RHS's)	2.99		1.17e-02	3.00

Result (2):

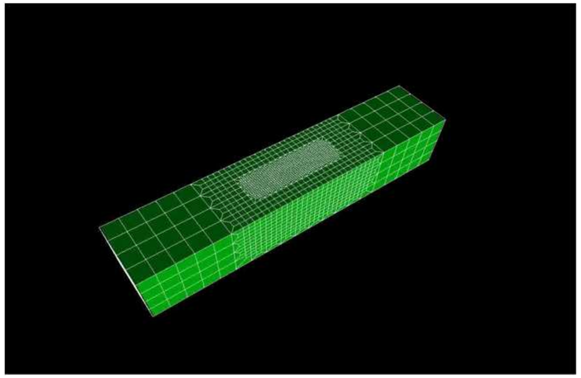
Vary $\lambda_1/\lambda_2 \sim \mathrm{cond}$ between 1 and 10^6

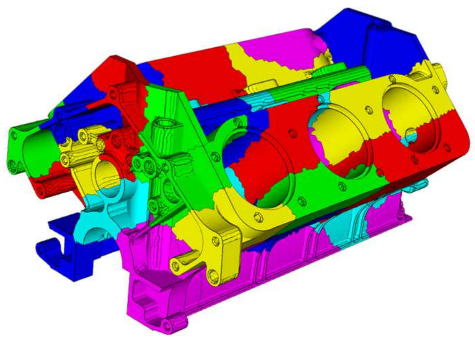


It is difficult to estimate the error of a computed solution only from residual norm!

(Near) Future Works in FY.2019

- Accuracy Verification + AT
  - More Reasonable Method for Accuracy Verification
    - Ill-Conditioned Sparse/H Mat.: Combined with Iterative Refinement
  - Strategy for Selection of Optimum Precision by AT (and ML)
    - Accuracy, Computation Time, Power Consumption
  - Trans-Precision
    - Challenging Approach: e.g. AT + FPGA
- FEM with Local Adaptive Precision
  - Precision changes on each element
    - New Idea
  - Heterogeneity, Stress Concentration, Elastic-Plastic (Linear-NL), Separation
  - Load In-Balancing in Parallel Computing
  - Discussions in WCCM 2018 in NYC
- **Towards "Appropriate Computing"**
  - **Approximate Computing + Accuracy Verification + Automatic Tuning (AT)**





- Background
  - ppOpen-HPC
  - Society 5.0
- BDEC System in ITC/U.Tokyo
- **Computing in the Exascale/Post Moore Era**
  - Approximate Computing
  - Verification of Accuracy
  - **Data Drive Approach**
  - **h3-Open-BDEC**
- Summary

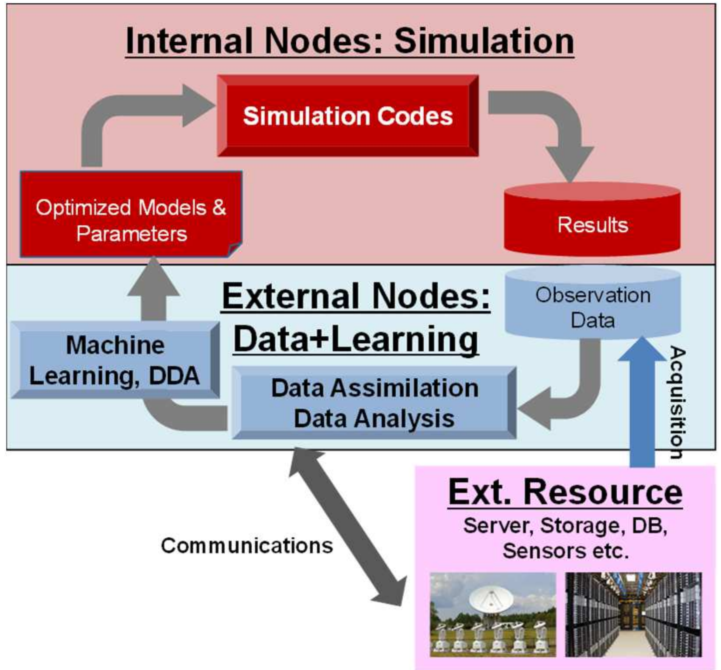
Computing in the Exascale/Post Moore Era

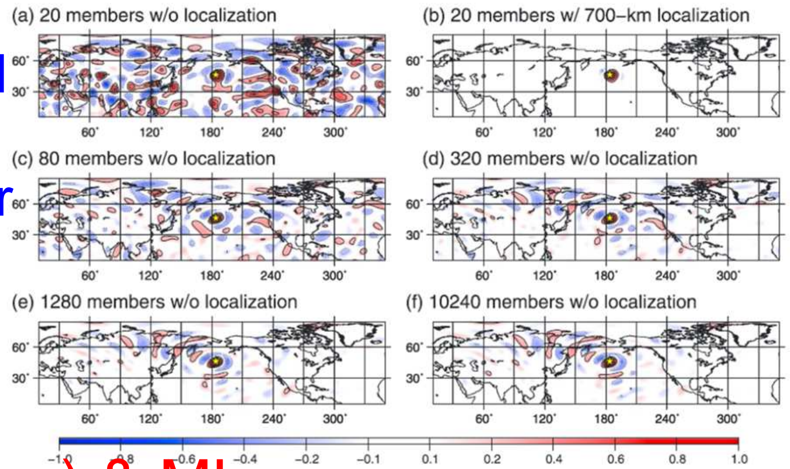
- Power Consumption is the Most Important Issue in the Post Moore Era
  - It is already important now.
  - Memory performance in the Post Moore Era is relatively better than now, but data movement should be reduced from the view point of energy consumption.
- Quantum Computing, FPGA ?: "Partial" Solution
  - Could be a solution in certain applications (e.g. searching, graph, data clustering etc.)
  - Contributions to (S+**D+L**)/(A+**B+C**)
- **How to save Energy for Sustainability ?**
  - **(1) Approximate Computing by Low/Adaptive Precision**
  - **(2) Reduction of Computations: Data Driven Approach**

Data Driven Approach

DDA, Integration of (S+D+L)

- Real-World Simulations
  - Non-Linear: Huge Number of Parameter Studies needed
    - Reduction of cases is a very crucial issue
  - Data Assimilation
    - Mid-Range Weather Prediction: 50-100 Ensemble Cases, 1,000 needed for accurate solution.
    - 50-100 (or fewer) may be enough for accurate solution, if opt. parameters are selected (e.g. by ML),



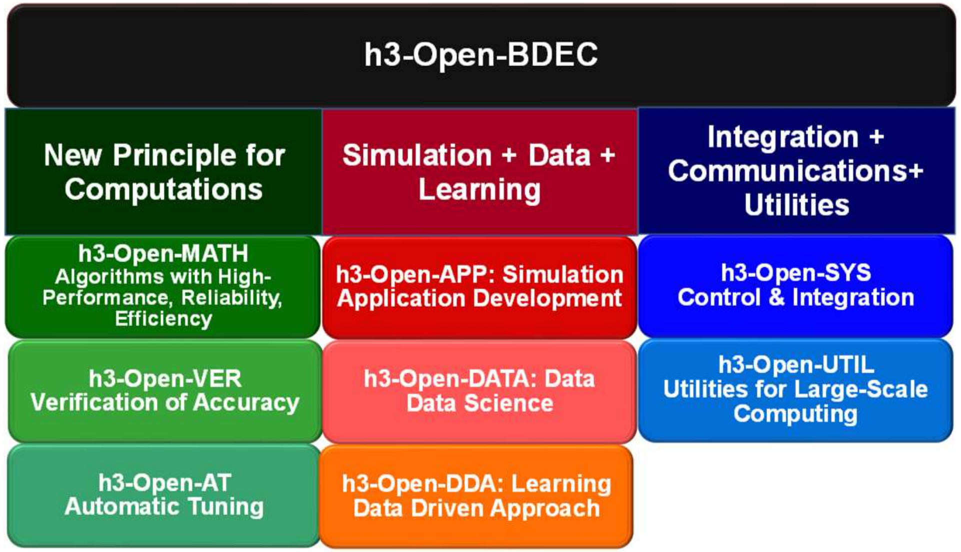


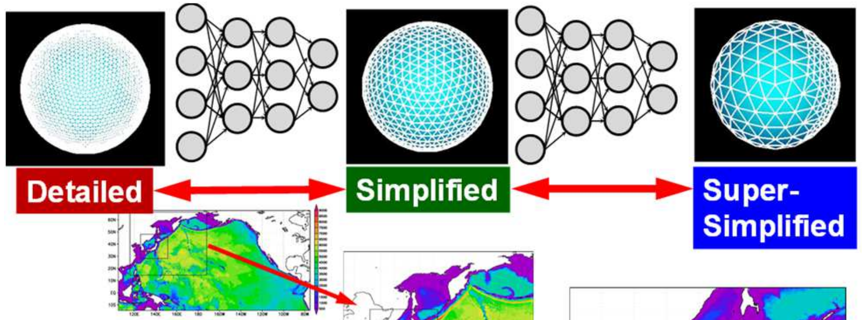
[Miyoshi et al. 2014]

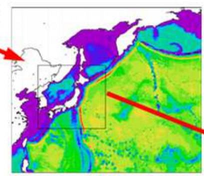
- Data Driven Approach (DDA)
  - Integration of CSE & (Observation) & ML
  - O(10^3-10^4) Training Data Sets: Difficult
    - Successful under Only Limited Conditions using Simplified Models
  - hDDA: Hierarchical DDA by More Efficient Training Approach

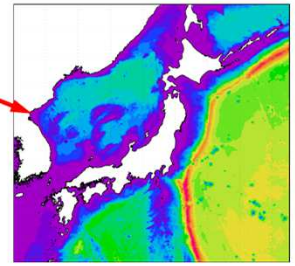
h3-Open-BDEC

- Innovative Software Infrastructure for (S+D+L)
  - h3: Hierarchical, Hybrid, Heterogeneous
  - Extension of ppOpen-HPC
  - Plug-in Existing Tools/Lib.
- Innovative/New Ideas
  - Adaptive Precision + Accuracy Verification + AT
    - Appropriate Computing
  - hDDA for General Problems by ML -> Reduction of Computations
    - Simplified/Local/Surrogate Model by ML
    - Multilevel/Multi-nested Approach using AMR
    - MOR (Model Order Reduction)
    - UQ (Uncertainty Quantification)









  - Control & Integration by Heterogeneous Containers
    - Various Functions on Heterogeneous Architectures
    - Including CPU, GPU, FPGA, Quantum Devices

- Background
  - ppOpen-HPC
  - Society 5.0
- BDEC System in ITC/U.Tokyo
- Computing in the Exascale/Post Moore Era
  - Approximate Computing
  - Verification of Accuracy
  - Data Drive Approach
  - h3-Open-BDEC
- **Summary**

Summary

- ppOpen-HPC
- Society 5.0 in Japan
- BDEC System, Next Stage
  - Platform for Simulation + Data + Learning (S+D+L)
  - Prototype for the Post-Moore Era Computing
    - Heterogeneous
    - Power Consumption
- Development of Software is also needed
  - h3-Open-BDEC
    - Extension of ppOpen-HPC towards Society 5.0
    - Low/Adaptive/Trans Precision
    - Reduction of Computations: Data Driven Approach
  - Proposals to Japanese Government

ICPP 2019 in Kyoto

48th International Conference on Parallel Processing

August 5-8, 2019

http://www.icpp-conf.org/

Submission Open:	February 01, 2019
Deadline for Submission (10-pages):	April 15, 2019
Author Notification:	May 17, 2019
Camera-Ready Due:	June 07, 2019




INTERNATIONAL CONFERENCE ON PARALLEL PROCESSING

Invited Speakers

Depei Qian (Sun Yat-Sen University & Beihang University, China)
Satoshi Sekiguchi (AIST, Japan)
Richard Vuduc (Georgia Tech, USA)

We are also calling for Exhibitors !!!