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](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

\[
\begin{align*}
A & : \text{Whole matrix} \\
A_{ij} & : \text{The entry in the } i\text{-th row} \\
& \text{and the } j\text{-th column of } A \\
n & : \text{Number of rows of } A \\
n & : \text{Number of processors} \\
R & : \text{The } k\text{-th processor} \\
R_k & : \{i \mid l_k \leq i < l_{k+1}\} \\
S & : S(a^c_{ij}) = 1 \text{ when } a^c_{ij} = A_{ij}
\end{align*}
\]
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]
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.)

1.0: Hunting
2.0: Agrarian
3.0: Industry
4.0: Information
5.0: Super Smart

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

- 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 in Real Time
  - 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 viewpoint of energy consumption.

• Quantum Computing, FPGA ?: “Partial” Solution
  – Could be a solution in certain applications (e.g. searching, graph, data clustering etc.)
  – Contributions to \( \frac{S+D+L}{A+B+C} \)

• 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

• 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)
Results: $\lambda_1/\lambda_2 \sim$ Condition Number
Ratio of Iterations & Computation Time
Single/Double: Down is Good

\[ \nabla \cdot (\lambda \nabla \phi) + f = 0 \]

Intel Xeon BDW
Single Node:
18 cores x 2 soc's (Reedbush-U)
Results: $\lambda_1/\lambda_2$ ~ Condition Number

Ratio of Iterations & Computation Time

Single/Double: Down is Good

Intel Xeon BDW
Single Node:
18cores x 2soc’s (Reedbush-U)
ICCG: ELL/Sliced ELL/SELL-C-ν

ICCG Solvers on Intel Xeon/Phi (KNL) (Oakforest-PACS) SINGLE NODE: 64/68 CORES

SELL-C-ν for ICCG

[Kreutzer, Wellein et al. 2014]
Results on OFP, Poisson-3D-OMP
Effect of SIMD Vector Length in SELL-C-σ
10 colors, 128³
FP32 (Single) with FP16 Precond.
V100, All Problems converge in FP32/64

$\phi = 0 \text{ at } z = z_{\text{max}}$

$\lambda_1 / \lambda_2 = 1.0e+3$
$\lambda_1 / \lambda_2 = 1.0e+2$
$\lambda_1 / \lambda_2 = 1.0e+1$
$\lambda_1 / \lambda_2 = 1.0e+0$

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

[Hoshino 2018]
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]
Energy Consumption (Normalized): P100, V100

[Sakamoto et al. 2018]
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 $\|A^{-1}\|$
3. Calculate lower/upper bound of $r = A\tilde{x} - b \Rightarrow r_{low}$ and $r_{upp}$ (in higher preecision)
4. Solve $A\tilde{z} = r_{low}$ and/or $A\tilde{z} = r_{upp}$
5. Calculate upper bound of absolute error: $\varepsilon_{abs} \geq \|\tilde{x} - x^*\|_\infty$
   
   \[ (x^* : \text{exact solution of } Ax = b) \]

5. Calculate upper bound of relative error: $\varepsilon_{rel} \geq \frac{\|\tilde{x} - x^*\|_\infty}{\|x^*\|_\infty}$

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

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

$$\|A^{-1}\|_\infty \leq \frac{\|\tilde{y}\|_\infty}{1 - \|A\tilde{y} - e\|_\infty}$$
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 $Ax = b$, $\frac{\|b - Ax\|_2}{\|b\|_2} < 10^{-12}$
  - For $Ay = e$, $\|e - Ay\|_\infty < 10^{-2}$
- 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:
  
  \[
  \max_{1 \leq i \leq n} \left| \frac{x_i - \hat{x}_i}{x_i} \right| \leq 3.38 \times 10^{-8}
  \]
  
  \[
  \frac{\|b - Ax\|_2}{\|b\|_2} < 3.66 \times 10^{-11}
  \]

- Computing time

<table>
<thead>
<tr>
<th>Method</th>
<th>Approximation Solve $Ax=b$ ((415\ iter’s))</th>
<th>Verification-1 Solve $Ay=e$ ((211\ iter’s))</th>
<th>Verification-2</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Method-1</td>
<td>2.38</td>
<td>1.18</td>
<td></td>
<td>3.56</td>
</tr>
<tr>
<td>Method-2</td>
<td>2.99</td>
<td>1.17e-02</td>
<td></td>
<td>3.00</td>
</tr>
<tr>
<td>(2 RHS’s)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Result (2):
Vary $\lambda_1/\lambda_2 \sim \text{cond}$ 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 \( \frac{S+D+L}{A+B+C} \)

• 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),

- 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

[Miyoshi et al. 2014]
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

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 !!!