Welcome and Announcements

Webinar Schedule
https://sc20.supercomputing.org/program/studentssc/studentssc-webinars/
Speakers

Stephen Lien Harrell - SC20 Reproducibility Challenge Chair
Michela Taufer - Founder of Reproducibility Effort at SC
Beth Plale - SC20 Transparency & Reproducibility Chair
Verónica Melesse Vergara - SC20 vSCC Co-chair
Scott Michael - SC20 vSCC Co-chair
Mert Hidayetoglu - Author of reproducibility challenge paper
Tekin Bicer - Co-author of reproducibility challenge paper
Reproducibility in HPC and Our Everyday Lives

“Do we trust the experimental results published in, e.g., the Nature journal? “
Yes, because …
• The experiment was reproduced multiple times converting to the same scientific conclusions

Illustration by Kallum Best

Image sources:
https://www.alfadispenser.com/2018/05/30/repeatability-vs-reproducibility-a-little-talk/
https://undsci.berkeley.edu/article/0_0_0/howscienceworks_17
“Do we trust the experimental results published in, e.g., the Nature journal?”
Yes, because …
• The experiment was reproduced multiple times converging to the same scientific conclusions
• The experimental process was documented step by step
Reproducibility in HPC and Our Everyday Lives

“Do we trust the experimental results published in, e.g., the Nature journal? “
Yes, because …
- The experiment was reproduced multiple times converting to the same scientific conclusions
- The process was documented step by step

“What is one of the advantages of a reproducible and well documented experiment?”
- Everyone can reproduce the experiment
- Anyone can build new science by leveraging the reproduced results

Illustration by Kallum Best

Image sources:
https://www.alfadispenser.com/2018/05/30/repeatability-vs-reproducibility-a-little-talk/
https://undsci.berkeley.edu/article/0_0_0/howscienceworks_17
Reproducibility in HPC and Our Everyday Lives

“Do we trust the experimental results published in, e.g., the Nature journal? “
Yes, because …
- The experiment was reproduced multiple times converting to the same scientific conclusions
- The process was documented step by step

“What is one of the advantages of a reproducible and well documented experiment?”
- Everyone can reproduce the experiment
- Anyone can build new science by leveraging the reproduced results

What if we replace “experimental results” with “computational results”? 

Illustration by Kallum Best

Image sources:
https://www.alfadispenser.com/2018/05/30/repeatability-vs-reproducibility-a-little-talk/
https://undsci.berkeley.edu/article/0_0_0/howscienceworks_17
Life Cycle of Reproducibility at SC

Technical Program @ SC X
Select BP/BSP candidates

Assign badge
Check AD or CRA
Review papers

Artifacts Available

SC X Papers
Computational Results (CRA)
Artifact Descriptor (AD)
Life Cycle of Reproducibility at SC

Technical Program @ SC X
Select BP/BSP candidates

Assign badge
Check AD or CRA

Review papers

SC X Papers

Select one (1) SC X paper for SC X+1 ECC
Life Cycle of Reproducibility at SC

Technical Program @ SC X
Select BP/BSP candidates

Artifacts Available
Assign badge
Check AD or CRA

SC X Papers
Review papers

Computational Results (CRA)
Artifact Descriptor (AD)

Select one (1) SC X paper for SC X+1 SCC

Generate replication benchmark for diverse set of HPC platforms
Life Cycle of Reproducibility at SC

Technical Program @ SC X
Select BP/BSP candidates

Assign badge
Check AD or CRA
Review papers

Select one (1) SC X paper for SC X+1 SCC

Generate replication benchmark for diverse set of HPC platforms

Student Cluster Competition @ SC X+1
Partner with vendors
Build a cluster
Test performance benchmarks
Replicate SC X Paper
Generate replication reports
Life Cycle of Reproducibility at SC

**Technical Program @ SC X**
- Select BP/BSP candidates
- Assign badge
- Check AD or CRA
- Review papers

**Technical Program @ SC X+1**
- Assign badge to SC X paper
- Give SIGHPC certificate to SC X paper authors
- Present Journal SI with SCC reports from SCC @ SC X-1

**Select one (1) SC X paper for SC X+1 SCC**

**Generate replication benchmark for diverse set of HPC platforms**

**Student Cluster Competition @ SC X+1**
- Partner with vendors
- Build a cluster
- Test performance benchmarks
- Replicate **SC X** Paper
- Generate replication reports
Life Cycle of Reproducibility at SC

**Technical Program @ SC X**
- Select BP/BSP candidates
- Assign badge
- Check AD or CRA
- Review papers

**Technical Program @ SC X+1**
- Assign badge to SC X paper
- Generate replication benchmark for diverse set of HPC platforms

**Technical Program @ SC X+2**
- Review IEEE TPDS paper with SCC reports from SCC @ SC X+1
- Give SIGHPC certificate to SC X paper authors
- Present Journal SI with SCC reports from SCC @ SC X-1
- Partner with vendors
- Build a cluster
- Test performance benchmarks
- Replicate SC X Paper
- Generate replication reports
Overview

- Four applications plus benchmarks -- GROMACS, CESM, MemXCT, Mystery application
- Reproducibility application scoring is unique; based on report

Recommendations

- Start your reports early; you already know a lot about your infrastructure and can plan your experiment in advance
- Plan for contingencies; have a target number of runs but also know what is the minimum you need
- Reproduction is not required for a good score, however, a good explanation of why the experiment was reproduced or not is required for a good score.
Beginning: Summer Internship at Argonne
**Advanced Photon Source**
- Hosts 5,300 researchers each year
- High-energy, high-brightness, highly penetrating, x-ray beams
  - 1,000,000,000 times brighter than dentist X-rays
- Helps answer the big questions about materials, superconductors, disease, battery electrodes, etc.

**Argonne Leadership Computing Facility**
- Theta supercomputer
  - Intel KNL, 11 petaflops
- 1,240+ users in 2019, 8.7B core-hours, 413 active projects, 275+ publications
- Aurora: Next exascale computer
  - Intel and Cray, 2021
Applications of MemXCT

X-Ray Ptychographic Imaging
- Reconstruction of a Integrated Circuit
- 25000 Hz pixel sampling
- 1cm$^2$ area - 41 nm pixels – 10GB/s
- 1 month of data collection

Shale Rock  IC Chip  Activated Charcoal  Mouse Brain
Motivation: Advancement Trends

Brilliance x3 in 18 months

Computer speed x2 in 18 months

X-Ray Ptychographic Imaging
- Reconstruction of a Integrated Circuit
- 25000 Hz pixel sampling
- 1cm² area - 41 nm pixels – 10GB/s
- 1 month of data collection
Data Acquisition

\[ Ax = b \]
\[ A^T A = A^T b \]

Projection: \( p_\theta(s) \)

Sample

Detector

\( f(x, y) \)

X-ray beam

Rotation stage

X-ray source

Sinogram

Tomogram
Iterative Image Reconstruction

Ax = b

1. Forward model
2. Compare
3. Inverse model

Stop? Yes → Output data
No → Reverse model
Memory-Centric Approach

Projection Matrix

Sparse Matrix Transpose

Backprojection Matrix

Access Order is Preserved!

Memory Wall Problem

Processor-DRAM Memory Gap

“Moore’s Law”

μProc 1.52/yr.
(2X/1.5yr)

μProc 1.20/yr.

DRAM 7%/yr.
(2X/10 yrs)

Processor-Memory Performance Gap:
(grows 50% / year)
How to Store 2-D Data?

- **Row-Major**: Bad Data Locality, Most Intuitive Storage
- **Blocked Row-Major**: Better Spatial Data Locality, Does Not Provide Connectivity
- **Morton**: Better Spatial Data Locality, Provides Connectivity
- **Hilbert**: Better Spatial Data Locality, Provides Connectivity
Access Pattern to Tomogram Data

Single Level - Single Line
Capacity: 64 bytes (16 floats)

Row-Major Ordering:
16 Cache Misses
1024 Bytes

Hilbert Ordering:
6 Cache Misses
384 Bytes
Sparsity Patterns

**Projection Matrix**

Row-Major Ordering

Hilbert Ordering

Sparse Matrix Transpose

**Backprojection Matrix**

Hilbert Ordering

Tomogram

Sinogram

Access Order is Preserved!

Multi-Stage Buffering

Tomogram Domain

Sinogram Domain

Sinogram Partition

256x256

256x256
Multi-Stage Buffering

Tomogram Domain

Memory Access Footprint

Sinogram Domain

Sinogram Partition
64x64

Data Reuse

256x256

x

y

θ

ρ
Leadership Class Supercomputers

Theta @ Argonne National Lab. (2016)

Summit @ Oak Ridge National Lab. (2019)

Blue Waters @ U. of. Illinois (2013)

Aurora @ Argonne National Lab. (2021)
Performance Optimizations on KNL

Higher is Better

<table>
<thead>
<tr>
<th>Sinogram Size</th>
<th>ADS1</th>
<th>ADS2</th>
<th>ADS3</th>
<th>ADS4</th>
</tr>
</thead>
<tbody>
<tr>
<td>(360x256)</td>
<td>1x (1.56x)</td>
<td>1x (1.59x)</td>
<td>1x (3.52x)</td>
<td>1x (2.29x)</td>
</tr>
<tr>
<td>430 MB</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(750x512)</td>
<td>1x (4.62x)</td>
<td></td>
<td>1x (3.78x)</td>
<td>1x (2.51x)</td>
</tr>
<tr>
<td>3.6 GB</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1500x1024)</td>
<td></td>
<td>1x (2.97x)</td>
<td></td>
<td>1x (3.38x)</td>
</tr>
<tr>
<td>28 GB</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(2400x2048)</td>
<td></td>
<td></td>
<td>1x (3.42x)</td>
<td></td>
</tr>
<tr>
<td>180 GB</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
**Strong Scaling**

**Strong Scaling on ALCF Theta**

- Grid: 11283x11283x11283
- Projections: 4500
- Memory Footprint: 10.2 TB

**Weak Scaling on NCSA Blue Waters**

- Grid: 2048x2048x2048
- Projections: 1500
- Memory Footprint: 112 GB

**Shale Sample**

- Grid: 2048x2048x2048
- Projections: 1500
- Memory Footprint: 112 GB
Weak Scaling

**Weak Scaling on ALCF Theta**

- Artificial Dataset
- 1500x1024 (28 GB)
- 24000x16384 (115 TB)

![Graph showing weak scaling on ALCF Theta](image)

**Weak Scaling on NCSA Blue Waters**

- Artificial Dataset
- 12000x8192 (14.7 TB)
- 750x512 (3.6 GB)

![Graph showing weak scaling on NCSA Blue Waters](image)
3-D Extension: SC20 Follow-up Paper

<table>
<thead>
<tr>
<th>Sample</th>
<th>Measurement Data Cube (K \times M \times N)</th>
<th>I/O Data Footprint</th>
<th>Memory Footprint</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shale Rock</td>
<td>1501×1792×2048</td>
<td>52.1 GB</td>
<td>120 GB</td>
</tr>
<tr>
<td>IC Chip</td>
<td>1210×1024×2448</td>
<td>36.7 GB</td>
<td>139 GB</td>
</tr>
<tr>
<td>Activated Charcoal</td>
<td>4500×4198×6613</td>
<td>1.23 TB</td>
<td>2.82 TB</td>
</tr>
<tr>
<td>Mouse Brain</td>
<td>4501×9209×11,283</td>
<td>6.56 TB</td>
<td>10.9 TB</td>
</tr>
</tbody>
</table>
3-D Extension: SC20 Follow-up Paper

Optimized SpMM Kernel Speedup
- Projection
- Backprojection
- Speedup vs Minibatch Size
  - Double
  - Single
  - Half
  - Mixed

Optimized SpMM Roofline Analysis
- Projection
- Backprojection
- Minibatch Size: 20
- Minibatch Size: 18
- Minibatch Size: 28
- Memory BW Bound
- Arithmetic Intensity (FLOPS/Byte)

Shale Recon. on 4 Nodes (24 GPUs)
- Wall Time (s)
- Kernel
- Comm.
- Idle
- CG
- I/O
- Synchronized
- Opt.
- +Kernel +Comm.
- Opt.*
- Double (1x4)
- Single (2x2)
- Mixed (4x1)

Charcoal Recon. on 128 Nodes (768 GPUs)
- Wall Time (m)
- Kernel
- Comm.
- Idle
- CG
- I/O
- Synchronized
- Opt.
- +Kernel +Comm.
- Opt.*
- Double (1x128)
- Single (2x64)
- Mixed (4x32)
Reproducibility Reports

What to Expect

- Describe your architecture, hardware and software
- Describe your experimental setup
- Describe the design of your experiment (amount of runs, implementation details, etc)
- Describe (and plot) the outcomes of your experiment
- Visualize the output data
- Describe the differences between the paper’s results and your experiment's results

Suggestions for Preparation

- Take a look at past year’s reports (will send these out to the SCC Google Group next week)
- Write any part that you can write before the competition (architecture, experiment setup, etc)
- Think and ask about (before the competition) what could cause any differences that you may see between your results and the results of the paper.
  - Some hints on where to look: architecture, scale, software stack, interconnect, etc
- Make sure you understand weak and strong scaling
Questions or Comments?