Note

- The following research was performed under the HPC Advisory Council activities
  - Compute resource - HPC Advisory Council Cluster Center

- The following was done to provide best practices
  - RELION performance overview over AMD based platforms
  - Understanding RELION communication patterns

- More info on RELION
  - https://github.com/3dem/relion
  - https://www3.mrc-lmb.cam.ac.uk/relion/index.php/Main_Page
RELION (REgularized LIkelihood OptimizatioN) is an open-source program for the refinement of macromolecular structures by single-particle analysis of electron cryo-microscopy (cryo-EM) data.

RELION (REgularized LIkelihood OptimizatioN) implements an empirical Bayesian approach for analysis of electron cryo-microscopy (Cryo-EM).

RELION provides refinement methods of singular or multiple 3D reconstructions as well as 2D class averages.

RELION is an important tool in the study of living cells.
Cluster Configuration

**Dallas cluster**
- Dual Socket  AMD EPYC 7742 CPU @ 2.25GHz, 2 sockets/node
- NVIDIA ConnectX-6 HDR InfiniBand
- NVIDIA Quantum Switch HDR InfiniBand
- Memory: 192GB DDR4 2677MHz RDIMMs per node
- Lustre Storage

**Software**
- OS: RHEL 8.3, MLNX_OFED 5.2.1
- MPI: HPC-X 2.8.1
- Relion 3.1.0
- Input: shiny_2sets.star (Plasmodium Ribosome, 3D)
RELION Performance and Scalability

![RELION Performance and Scalability](image)

**Relion**

*(Plasmodium ribosome 3D)*

- **Elpased time (s)**
- **Number of Nodes**
  - 4
  - 8
  - 16

- **AMD EPYC 7742 CPU @ 2.25GHz**

*Lower is better*
• SHARP reduced MPI time by 12%
RELION MPI Profiling

- 47% MPI time

16 Nodes, SHARP disabled

- 42% MPI time

16 Nodes, SHARP enabled
• 84% of MPI Communication spent on MPI_Barrier
• 13% MPI_Wait

<table>
<thead>
<tr>
<th>Comm Size</th>
<th>Buffer Size</th>
<th>Ncalls</th>
<th>Total Time</th>
<th>Avg Time</th>
<th>Min Time</th>
<th>Max Time</th>
<th>%MPI</th>
<th>%Wall</th>
</tr>
</thead>
<tbody>
<tr>
<td>MPI_Barrier</td>
<td>16</td>
<td>0</td>
<td>5152</td>
<td>9.454008e+03</td>
<td>1.835017e+00</td>
<td>1.907300e-06</td>
<td>5.159100e+01</td>
<td>83.79</td>
</tr>
<tr>
<td>MPI_Wait</td>
<td>0</td>
<td>0</td>
<td>22595</td>
<td>1.487701e+03</td>
<td>6.584205e-02</td>
<td>0.000000e+00</td>
<td>3.829000e+00</td>
<td>13.19</td>
</tr>
<tr>
<td>MPI_Bcast</td>
<td>16</td>
<td>0</td>
<td>335544320</td>
<td>2.736226e+02</td>
<td>1.140094e-01</td>
<td>8.918400e-02</td>
<td>1.310200e-01</td>
<td>2.43</td>
</tr>
<tr>
<td>MPI_Send</td>
<td>0</td>
<td>0</td>
<td>536870912</td>
<td>2.323529e+01</td>
<td>5.532212e-02</td>
<td>2.365400e-02</td>
<td>1.609400e-01</td>
<td>0.21</td>
</tr>
<tr>
<td>MPI_Bcast</td>
<td>15</td>
<td>450</td>
<td>25165824</td>
<td>9.822899e+00</td>
<td>2.182866e-02</td>
<td>1.845600e-02</td>
<td>3.063200e-02</td>
<td>0.09</td>
</tr>
<tr>
<td>MPI_Bcast</td>
<td>15</td>
<td>270</td>
<td>16777216</td>
<td>5.291484e+00</td>
<td>1.959809e-02</td>
<td>1.565900e-02</td>
<td>2.765700e-02</td>
<td>0.05</td>
</tr>
</tbody>
</table>

16 Nodes, 1 rank per node, SHARP enabled
RELION MPI Profiling

- Rank 0 – is not a compute rank, use to distribute the job

32 Nodes, 1 rank per node, SHARP enabled
RELION MPI Profiling

• Ring communication
RELION Profiling

- Memory footprint
Performance Analysis Summary

• RELION performance testing
  – Relion is scaling good up to 8 nodes and then slower up to 16 nodes
  – SHARP In-Network Computing reduces MPI time by 12%
  – Performance advantages increases with system size, up to 16 nodes were tested

• RELION Profile
  – Rank #0 does not perform computation
  – Mostly MPI_Barrier (83%)
  – Ring communication matrix
Thank You