**MVAPICH2-X: High Performance MPI+PGAS Unified Runtime**

Dhabaleswar K. (DK) Panda, Khaled Hamidouche – The Ohio State University, {panda, hamidouc}@cse.ohio-state.edu

- High performance open-source MPI for InfiniBand, 10-40Gig/iWARP, and RoCE
  - MVAPICH (MPI-1), MVAPICH2 (MPI-2.2 and MPI-3.0), Available since 2002
  - MVAPICH2-X (MPI + PGAS), Available since 2011
  - Support for GPGPUs (MVAPICH2-GDR), Available since 2014
  - Support for MIC (MVAPICH2-MIC), Available since 2014
  - Support for Virtualization (MVAPICH2-Virt), Available since 2015
  - Support for Energy-Awareness (MVAPICH2-EA), Available since 2015
  - Used by more than 2,475 organizations in 76 countries
  - More than 304,000 downloads from the OSU site directly
  - Empowering many TOP500 clusters (Nov’15 ranking)
  - [http://mvapich.cse.ohio-state.edu](http://mvapich.cse.ohio-state.edu)
- Empowering Top500 systems for over a decade

**Benefits of Hybrid MPI+OpenSHMEM on Graph500**

<table>
<thead>
<tr>
<th>No. of Processes</th>
<th>Time (s)</th>
<th>MPI-Simple</th>
<th>MPI-CSC</th>
<th>MPI-CSR</th>
<th>Hybrid (MPI+OpenSHMEM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>4K</td>
<td></td>
<td>35</td>
<td>30</td>
<td>25</td>
<td>15</td>
</tr>
<tr>
<td>8K</td>
<td></td>
<td>35</td>
<td>30</td>
<td>25</td>
<td>15</td>
</tr>
<tr>
<td>16K</td>
<td></td>
<td>35</td>
<td>30</td>
<td>25</td>
<td>15</td>
</tr>
</tbody>
</table>

- 16,384 processes
  - 1.5X improvement over MPI-CSRe
  - 13X improvement over MPI-Simple

**Next Step for PGAS models: Accelerator Support**

- Extend Memory model for heterogeneous Memory domains
- **heap_on_device/heap_on_host**
  - (a way to indicate location of heap)
  - $host_buf = shmalloc(sizeof(int), 0);$
  - $dev_buf = shmalloc(sizeof(int), 1);$
- Accelerator-Aware MVAPICH2-X support for PGAS models