Accelerating Performance and Scalability with NVIDIA GPUs on HPC Applications

Pak Lui
The HPC Advisory Council Update

- World-wide HPC non-profit organization
- ~425 member companies / universities / organizations
- Bridges the gap between HPC usage and its potential
- Provides best practices and a support/development center
- Explores future technologies and future developments
- Leading edge solutions and technology demonstrations
HPC Advisory Council Centers

HPC ADVISORY COUNCIL CENTERS

HPCAC HQ
AUSTIN

SWISS (CSCS)

CHINA
HPC Advisory Council HPC Center

- Dell™ PowerEdge™ R730 GPU 36-node cluster
- Dell PowerVault MD3420 Dell PowerVault MD3460
  InfiniBand Storage (Lustre)
- HPE Apollo 6000 10-node cluster
- HPE ProLiant SL230s Gen8 4-node cluster
- HPE Cluster Platform 3000SL 16-node cluster

- Dell™ PowerEdge™ C6145 6-node cluster
- Dell™ PowerEdge™ R815 11-node cluster
- Dell™ PowerEdge™ R720xd/R720 32-node GPU cluster
- Dell™ PowerEdge™ M610 38-node cluster

- Dell™ PowerEdge™ C6100 4-node cluster
- Dell™ PowerEdge™ M610 4-node GPU cluster
- InfiniBand-based Storage (Lustre)

Dell PowerVault MD3420
Dell PowerVault MD3460
InfiniBand Storage (Lustre)
Exploring All Platforms / Technologies

X86, Power, GPU, FPGA and ARM based Platforms
HPC Training

- **HPC Training Center**
  - CPUs
  - GPUs
  - Interconnects
  - Clustering
  - Storage
  - Cables
  - Programming
  - Applications

- **Network of Experts**
  - Ask the experts
University Award Program

- **University award program**
  - Universities / individuals are encouraged to submit proposals for advanced research
- **Selected proposal will be provided with:**
  - Exclusive computation time on the HPC Advisory Council’s Compute Center
  - Invitation to present in one of the HPC Advisory Council’s worldwide workshops
  - Publication of the research results on the HPC Advisory Council website
- **2010 award winner is Dr. Xiangqian Hu, Duke University**
  - Topic: “Massively Parallel Quantum Mechanical Simulations for Liquid Water”
- **2011 award winner is Dr. Marco Aldinucci, University of Torino**
- **2012 award winner is Jacob Nelson, University of Washington**
  - “Runtime Support for Sparse Graph Applications”
- **2013 award winner is Antonis Karalis**
  - Topic: “Music Production using HPC”
- **2014 award winner is Antonis Karalis**
  - Topic: “Music Production using HPC”
- **2015 award winner is Christian Kniep**
  - Topic: Dockers
- To submit a proposal – please check the HPC Advisory Council web site
ISC'15 – Student Cluster Competition Teams
ISC'15 – Student Cluster Competition Award Ceremony
Getting Ready to 2016 Student Cluster Competition
2015 HPC Conferences
### The Third Student RDMA Programming Competition

Organized by HPC Advisory Council

#### Results

<table>
<thead>
<tr>
<th>Prize</th>
<th>University</th>
<th>Advisor</th>
<th>Students</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st</td>
<td>National University of Defense Technology</td>
<td>Prof. Nong XIAO</td>
<td>Min YANG, RuJie YU, SongPing YU, Ya FENG</td>
</tr>
<tr>
<td>2nd</td>
<td>University of Science and Technology of China</td>
<td>Prof. Hong AN</td>
<td>Wenhao YAO, QiZe XU, SiYuan, ZHUANG, Xing GUO, Xiao LUO</td>
</tr>
<tr>
<td>3rd</td>
<td>Xidian University</td>
<td>YiNing QUAN</td>
<td>Yang YU, ZhenXin MA, Chao WANG</td>
</tr>
<tr>
<td></td>
<td>Northwestern Polytechnical University</td>
<td>Tao Wang, Prof.</td>
<td>Jian WU, Hao WANG, ChangPeng XU, HaoNan HE, Dong LIANG</td>
</tr>
<tr>
<td></td>
<td>Chongqing University</td>
<td>JianHua GU</td>
<td>Jianhui DING, Chao QI, Tianyuan XU, Liang Liang YIN</td>
</tr>
<tr>
<td></td>
<td>Huazhong University of Science and Technology</td>
<td>Prof. YueMing LI</td>
<td>XuanHua SHI, YingZhong GUO, HongFan LU, XiaoQi JIAN</td>
</tr>
<tr>
<td></td>
<td>Ocean University of China</td>
<td>LinJie ZHANG</td>
<td>QinYu WANG, ChenTao WANG, YingQi ZHAO</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Zhi WANG, XiaoLiang WANG, Ting CHEN, Nai CHEN</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Liang SHI, HaoQi XU, ZuoChang SHU</td>
</tr>
<tr>
<td>Honorable</td>
<td>Nanjing University</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mention</td>
<td>Nanjing University of Aeronautics and Astronautics</td>
<td>Prof. ShengLin YU</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Huazhong University of Science and Technology Team B</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Sponsors:
- Sugon
- Argonne National Laboratory
- SDSC
- HPCTC
2015 HPC Advisory Council Conferences

- **HPC Advisory Council (HPCAC)**
  - Application best practices, case studies
  - Benchmarking center with remote access for users
  - World-wide workshops

- **2015 Workshops / Activities**
  - USA (Stanford University) – February
  - Switzerland (CSCS) – March
  - Student Cluster Competition (ISC) - July
  - Brazil (LNCC) – August
  - Spain (BSC) – Sep
  - China (HPC China) – November

- **For more information**
  - [www.hpcadvisorycouncil.com](http://www.hpcadvisorycouncil.com)
  - [info@hpcadvisorycouncil.com](mailto:info@hpcadvisorycouncil.com)
If you are interested to bring HPC Advisory Council conference to your area, please contact us.
Over 156 Applications Best Practices Published

<table>
<thead>
<tr>
<th>Application</th>
<th>Application</th>
<th>Application</th>
<th>Application</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abaqus</td>
<td>CPMD</td>
<td>LS-DYNA</td>
<td>MILC</td>
</tr>
<tr>
<td>AcuSolve</td>
<td>Dacapo</td>
<td>miniFE</td>
<td>OpenMX</td>
</tr>
<tr>
<td>Amber</td>
<td>Desmond</td>
<td>MILC</td>
<td>PARATEC</td>
</tr>
<tr>
<td>AMG</td>
<td>DL-POLY</td>
<td>MSC Nastran</td>
<td>PFA</td>
</tr>
<tr>
<td>AMR</td>
<td>Eclipse</td>
<td>MR Bayes</td>
<td>PFLOTRAN</td>
</tr>
<tr>
<td>ABySS</td>
<td>FLOW-3D</td>
<td>MM5</td>
<td>Quantum ESPRESSO</td>
</tr>
<tr>
<td>ANSYS CFX</td>
<td>GADGET-2</td>
<td>MPQC</td>
<td>RADIOSS</td>
</tr>
<tr>
<td>ANSYS FLUENT</td>
<td>GROMACS</td>
<td>NAMD</td>
<td>SPECFEM3D</td>
</tr>
<tr>
<td>ANSYS Mechanics</td>
<td>Himeno</td>
<td>Nekbone</td>
<td>WRF</td>
</tr>
<tr>
<td>BQCD</td>
<td>HOOMD-blue</td>
<td>NEMO</td>
<td></td>
</tr>
<tr>
<td>CCSM</td>
<td>HYCOM</td>
<td>NWChem</td>
<td></td>
</tr>
<tr>
<td>CESM</td>
<td>ICON</td>
<td>Octopus</td>
<td></td>
</tr>
<tr>
<td>COSMO</td>
<td>Lattice QCD</td>
<td>OpenAtom</td>
<td></td>
</tr>
<tr>
<td>CP2K</td>
<td>LAMMPS</td>
<td>OpenFOAM</td>
<td></td>
</tr>
</tbody>
</table>

The following research was performed under the HPC Advisory Council activities
- Participating vendors: Intel, Dell, Mellanox, NVIDIA
- Compute resource - HPC Advisory Council Cluster Center

The following was done to provide best practices
- LAMMPS performance overview
- Understanding LAMMPS communication patterns
- Ways to increase LAMMPS productivity

For more info please refer to
- http://www.dell.com
- http://www.intel.com
- http://www.mellanox.com
- http://www.nvidia.com
GROMACS

- **GROMACS (GROningen MAchine for Chemical Simulation)**
  - A molecular dynamics simulation package
  - Primarily designed for biochemical molecules like proteins, lipids and nucleic acids
    - A lot of algorithmic optimizations have been introduced in the code
    - Extremely fast at calculating the non-bonded interactions
  - Ongoing development to extend GROMACS with interfaces both to Quantum Chemistry and Bioinformatics/databases
  - An open source software released under the GPL
Objectives

• The presented research was done to provide best practices
  – GROMACS performance benchmarking
    • Interconnect performance comparison
    • CPUs/GPUs comparison
    • Optimization tuning

• The presented results will demonstrate
  – The scalability of the compute environment/application
  – Considerations for higher productivity and efficiency
Test Cluster Configuration

- **Dell PowerEdge R730 32-node (896-core) “Thor” cluster**
  - Dual-Socket 14-Core Intel E5-2697v3 @ 2.60 GHz CPUs (BIOS: Maximum Performance, Home Snoop)
  - Memory: 64GB memory, DDR4 2133 MHz, Memory Snoop Mode in BIOS sets to Home Snoop
  - OS: RHEL 6.5, MLNX_OFED_LINUX-3.2-1.0.1.1 InfiniBand SW stack
  - Hard Drives: 2x 1TB 7.2 RPM SATA 2.5” on RAID 1

- **Mellanox ConnectX-4 EDR 100Gb/s InfiniBand Adapters**

- **Mellanox Switch-IB SB7700 36-port EDR 100Gb/s InfiniBand Switch**

- **Mellanox ConnectX-3 FDR VPI InfiniBand and 40Gb/s Ethernet Adapters**

- **Mellanox SwitchX-2 SX6036 36-port 56Gb/s FDR InfiniBand / VPI Ethernet Switch**

- **Dell InfiniBand-Based Lustre Storage based on Dell PowerVault MD3460 and Dell PowerVault MD3420**

- **NVIDIA Tesla K40 and K80 GPUs; 1 GPU per node**

- **MPI: Mellanox HPC-X v1.4.356 (based on Open MPI 1.8.8) with CUDA 7.0 support**

- **Application: GROMACS 5.0.4 and 5.1.2 (Single Precision)**

- **Benchmark datasets: Alcohol dehydrogenase protein (ADH) solvated and set up in a rectangular box (134,000 atoms), simulated with 2fs step (http://www.gromacs.org/GPU_acceleration)**
PowerEdge R730
Massive flexibility for data intensive operations

- **Performance and efficiency**
  - Intelligent hardware-driven systems management with extensive power management features
  - Innovative tools including automation for parts replacement and lifecycle manageability
  - Broad choice of networking technologies from GigE to IB
  - Built in redundancy with hot plug and swappable PSU, HDDs and fans

- **Benefits**
  - Designed for performance workloads
    - from big data analytics, distributed storage or distributed computing where local storage is key to classic HPC and large scale hosting environments
    - High performance scale-out compute and low cost dense storage in one package

- **Hardware Capabilities**
  - Flexible compute platform with dense storage capacity
    - 2S/2U server, 6 PCIe slots
  - Large memory footprint (Up to 768GB / 24 DIMMs)
  - High I/O performance and optional storage configurations
    - HDD options: 12 x 3.5" - or - 24 x 2.5 + 2x 2.5 HDDs in rear of server
    - Up to 26 HDDs with 2 hot plug drives in rear of server for boot or scratch
GROMACS Installation

• **Build flags:**
  – $ CC=mpicc CXX=mpiCC cmake <GROMACS_SRC_DIR> -DGMX_OPENMP=ON -DGMX_GPU=ON -DGMX_MPI=ON -DGMX_BUILD OWN_FFTW=ON -DGPU_DEPLOYMENT KIT_ROOT_DIR=/path/to/gdk -DGMX_PREFER STATIC_LIBS=ON -DCMAKE_BUILD_TYPE=Release -DCMAKE_INSTALL_PREFIX=<GROMACS_INSTALL_DIR>

• **GROMACS 5.1: GPU clocks can be adjusted for optimal performance via NVML**
  – NVIDIA Management Library (NVML)

• **Setting Application Clock for GPUBOOST:**

• **References:**
  – GROMACS documentation
  – [Best bang for your buck: GPU nodes for GROMACS biomolecular simulations](http://hpcadvisorycouncil.com/pdf/GROMACS_GPU.pdf)

*Credits to NVIDIA*
GROMACS Run Time Options

- **GROMACS options to use when benchmarking**
  - `-resethway`
    - Reduces runtime needed to obtain stable results
  - `-noconfout`
    - Disables output of confout.gro which might take a long time to write
  - `-maxh <wall time>`
    - Controls the simulation should run
    - mdrun runs the time steps to run until the specified time in hours is reached
  - `-v`
    - Additional outputs to be printed to log file (md.log)

- **Considerations:**
  - `-nb`
    - Open MP threads, MPI, or hybrid:
    - Using -nb cpu, gpu, or gpu_cpu, -ntomp <NUM> or OMP_NUM_THREADS

*Credits to NVIDIA*
• **EDR InfiniBand provides higher scalability in performance for GROMACS**
  – InfiniBand delivers 465% higher performance than 10GbE on 8 nodes
  – Benefits of InfiniBand over Ethernet expect to increase as cluster scales
  – Ethernet would not scale; while InfiniBand scale continuously

**GROMACS Performance**
(adh_cubic, RF, K40)

**GROMACS Performance**
(adh_cubic, RF, K80)

**Higher is better**

GPU: 1 GPU / Node
GROMACS Profiling – % of MPI Calls

- The communication time for GPU stays roughly the same as cluster scales
  - While the compute time reduces as number of nodes increase
The most time consuming MPI calls for GROMACS (cuda):
- MPI_Sendrecv: 40% MPI / 19% Wall
- MPI_Bcast: 20% MPI / 9% Wall
- MPI_Comm_split: 16% MPI / 7% Wall

16 nodes, adh_cubic, RF

8 nodes, adh_cubic, PME
For the most time consuming MPI calls:
- MPI_Comm_split: 0B (14% MPI time)
- MPI_Sendrecv: 16KB (13% MPI time)
- MPI_Bcast: 4B (11% MPI time)

16 nodes, adh_cubic, RF

8 nodes, adh_cubic, PME
Some load imbalance is seen on the workload for MPI_SendRecv.

Memory consumption:
- About 400MB of memory is used on each compute node for this input data.
For adh_cubic, Tesla K80 generally outperforms the predecessor Tesla K40
- K80 can deliver up to 71% of higher performance on the adh_cubic data

GROMACS parameters used to control GPUs being used
- mdrun_mpi -gpu_id 01 -nb gpu_cpu (for K80, 2 MPI are being used for each GPU core)
For adh_dodec, the K80 performs 47% higher than Tesla K40.
For `adh_dodec_vsites`, the K80 performs 36% higher than Tesla K40.

**GROMACS Performance**
(adh_dodec_vsites, RF)

**GROMACS Performance**
(adh_dodec_vsites, PME)

*Higher is better*
• GPU has a performance advantage compared to just CPU cores on the same node
  – GPU outperforms the CPU only by 22%-55% for adh_cubic on a single node
• The scalability performance of CPUs as node count increases
  – The performance of CPU cluster delivers around 48% higher at 16 nodes (448 cores)
GROMACS Performance – CPU & GPU performance

- GPU has a performance advantage compared to just CPU cores on the same node
  - GPU outperforms the CPU only by 32%-44% for adh_dodec on a single node
- The scalability performance of CPUs as node count increases
  - The performance of CPU cluster delivers around 68% higher at 16 nodes (448 cores)

GROMACS Performance (adh_dodec, RF)

- GPU: 1 GPU / Node
- Higher is better
GROMACS – Summary

- **GROMACS demonstrates good scalability on cluster of CPU or GPU**
- The Tesla K80 outperforms the Tesla K40 by up to 71%
- **GPU outperforms CPU on a per node basis**
  - Up to 55% against the 28 core CPU per node
- **InfiniBand enables scalability performance for GROMACS**
  - InfiniBand delivers 465% higher performance than 10GbE on 8 nodes
  - Benefits of InfiniBand over Ethernet expect to increase as cluster scales
  - Ethernet would not scale; while InfiniBand scale continuously
  - Scalability performance on CPU cluster shown to be better than GPU cluster
- **The most time consuming MPI calls for GROMACS (cuda):**
  - MPI_Sendrecv: 40% MPI / 19% Wall
  - MPI_Bcast: 20% MPI / 9% Wall
  - MPI_Comm_split: 16% MPI / 7% Wall
HOOMD-blue

- Stands for Highly Optimized Object-oriented Many-particle Dynamics -- Blue Edition
- Performs general purpose particle dynamics simulations on workstation and cluster
- Takes advantage of NVIDIA GPUs to attain a level of performance equivalent to many processor cores on a fast cluster
- Is free, open source; anyone can change source for additional functionality
- Simulations are configured and run using simple python scripts, allowing complete control over the force field choice, integrator, all parameters, time steps, etc
- The scripting system is designed to be as simple as possible to the non-programmer
- The development effort is led by Glotzer group at University of Michigan
- Many groups from different universities have contributed code that is now part of the HOOMD-blue main package, see the credits page for the full list
Test Cluster Configuration

- Dell™ PowerEdge™ T620 128-node (1536-core) Wilkes cluster at Univ of Cambridge
  - Dual-Socket Hexa-Core Intel E5-2630 v2 @ 2.60 GHz CPUs
  - Memory: 64GB memory, DDR3 1600 MHz
  - OS: Scientific Linux release 6.4 (Carbon), MLNX_OFED 2.1-1.0.0 InfiniBand SW stack
  - Hard Drives: 2x 500GB 7.2 RPM 64MB Cache SATA 3.0Gb/s 3.5"

- Mellanox Connect-IB FDR InfiniBand adapters

- Mellanox SwitchX SX6036 InfiniBand VPI switch

- NVIDIA Tesla K20 GPUs (2 GPUs per node)

- NVIDIA CUDA 5.5 Development Tools and Display Driver 331.20

- GPUDirect RDMA (nvidia_peer_memory-1.0-0.tar.gz)

- MPI: Open MPI 1.7.4rc1, MVAPICH2-GDR 2.0b

- Application: HOOMD-blue (git master 28Jan14)

- Benchmark datasets: Lennard-Jones Liquid Benchmarks (256K and 512K Particles)
The Wilkes Cluster at University of Cambridge

- **The Wilkes Cluster**
  - The University of Cambridge in partnership with Dell, NVIDIA and Mellanox
    - deployed the UK’s fastest academic cluster, named Wilkes in November 2013
  - Produces a LINPACK performance of 240TF
    - on the Top500 position of 166 in the November 2013 list
  - Ranked most energy efficient air cooled supercomputer in the world
  - Ranked second in the worldwide Green500 ranking
    - Extremely high performance per watt of 3631 MFLOP/W
  - Factors behind the extreme energy efficiency:
    - Very high performance per watt provided by the NVIDIA K20 GPU
    - Industry leading energy efficiency obtained from the Dell T620 server
  - Interconnect network was architected by Mellanox
    - To achieve highest message rate possible for application scaling
      - Dual-rail Connect IB network providing a fully non-blocking network
      - Node to node bandwidth of over 100Gb/s
      - Message rate of 137 million messages per second
  - Architected to utilize the NVIDIA RDMA communication acceleration
    - Significantly increase the system's parallel efficiency
gdrcopy: A low-latency GPU memory copy library based on GPUDirect RDMA
- Offers the infrastructure to create user-space mappings of GPU memory
- Demonstrated further latency reduction by 55%
HOOMD-blue Performance – GPUDirect RDMA

- **GPUDirect RDMA unlocks performance between GPU and IB**
  - Demonstrated up to 102% of higher performance at 96 nodes
  - This new technology provides a direct P2P data path between GPU and IB
  - This provides a significant decrease in GPU-GPU communication latency
  - Complete offload CPU from all GPU communications across the network
  - MCA param to enable GPUDirect RDMA between 1 GPU and IB per node
    - `--mca btl_openib_want_cuda_gdr 1 -mca btl_openib_if_include mlx5_0:1`

### HOOMD-blue Performance
(LJ Liquid Benchmark, 512K Particles)

![Chart showing performance comparison between different numbers of nodes with and without GPUDirect RDMA]
HOOMD-blue Performance – Benefits of GPUDirect RDMA

**LJ Liquid Benchmark, 64K Particles**

- Red: Without GPUDirect RDMA
- Green: GPUDirect RDMA

**LJ Liquid Benchmark, 128K Particles**

- Red: Without GPUDirect RDMA
- Green: GPUDirect RDMA

**LJ Liquid Benchmark, 256K Particles**

- Red: Without GPUDirect RDMA
- Green: GPUDirect RDMA

**LJ Liquid Benchmark, 512K Particles**

- Red: Without GPUDirect RDMA
- Green: GPUDirect RDMA

*Higher is better*
HOOMD-blue Performance – Dual GPU-InfiniBand

- mpirun -np $NP -bind-to socket -display-map -report-bindings --map-by ppr:1:socket \\ --mca mtl ^mxm -mca coll_fca_enable 0 --mca btl openib,self --mca btl_openib_device_selection_verbose 1 \\ --mca btl_openib_warn_nonexistent_if 0 --mca btl_openib_if_include mlx5_0:1,mlx5_1:1 \\ --mca btl_smcuda_use_cuda_ipc 0 --mca btl_smcuda_use_cuda_ipc_same_gpu 1 --mca btl_openib_recommend_cuda_gdr 1 \\ hoomd lj_liquid_bmark_256000.hoomd

- mpirun -np $NP -ppn 2 -genvall -genv MV2_ENABLE_AFFINITY 1 -genv MV2_CPU_BINDING_LEVEL SOCKET \\ -genv MV2_CPU_BINDING_POLICY SCATTER -genv MV2_RAIL_SHARING_POLICY FIXED_MAPPING \\ -genv MV2_PROCESS_TO_RAIL_MAPPING mlx5_0:mlx5_1 \\ -genv MV2_USE_CUDA 1 -genv MV2_CUDA_IPC 0 -genv MV2_USE_GPUDIRECT 1 hoomd lj_liquid_bmark_256000.hoomd
HOOMD-blue Performance – Scalability

- Wilkes exceeds Titan in scalability performance with GPUDirect RDMA
  - Outperforms Titan by 111% at 64 nodes
- GPUDirect RDMA empowers Wilkes to surpass Titan on scalability
  - Titan has higher per-node performance but Wilkes outperforms in scalability
  - Titan: NVIDIA K20x GPUs at the PCIe Gen2 speed but at higher clock rate
  - Wilkes: NVIDIA K20 GPUs at PCIe Gen2, and FDR InfiniBand at Gen3 rate

HOOMD-blue Performance
(LJ Liquid Benchmark, 256K Particles)
HOOMD-blue Profiling – % Time Spent on MPI

- HOOMD-blue utilizes non-blocking communications in most data transfers
  - Change in network performance is seen between low node counts
  - 4 nodes: MPI_Waitall (75%), the rest are MPI_Bcast and MPI_Allreduce
  - 96 nodes: MPI_Bcast (35%), the rest are MPI_Allreduce, MPI_Waitall
Each rank engages in similar network communication
- Except for rank 0, which spends less time in MPI_Bcast
HOOMD-blue utilizes non-blocking and collectives for most data transfers
- 4 Nodes: MPI_Isend/MPI_Irecv are concentrated between 28KB to 229KB
- 96 Nodes: MPI_Isend/MPI_Irecv are concentrated between 64B to 16KB

GPUDirect RDMA is enabled for messages between 0B to 30KB
- MPI_Isend/_Irecv messages are able to take advantage of GPUDirect RDMA
- Messages fitted within the (tunable default of) 30KB window can be benefited
• Distribution of data transfers between the MPI processes
  – Non-blocking point-to-point data communications between processes are involved
HOOMD-blue – Summary

- **HOOMD-blue demonstrates good use of GPU and InfiniBand at scale**
  - FDR InfiniBand is the interconnect allows HOOMD-blue to scale
- **GPUDirect RDMA**
  - This new technology provides a direct P2P data path between GPU and IB
  - This provides a significant decrease in GPU-GPU communication latency
- **GPUDirect RDMA unlocks performance between GPU and IB**
  - Demonstrated up to 102% of higher performance at 96 nodes for 512K case
- **GPUDirect RDMA empowers Wilkes to surpass Titan on scalability**
  - Wilkes outperforms in scalability while Titan has higher per-node performance
  - Outperforms Titan by 111% at 64 nodes
- **FDR InfiniBand is the interconnect allows HOOMD-blue to scale**
  - Running in 1GbE would not scale beyond 1 node
Contact Us

Web: [www.hpcadvisorycouncil.com](http://www.hpcadvisorycouncil.com)
Email: [info@hpcadvisorycouncil.com](mailto:info@hpcadvisorycouncil.com)
Twitter: [www.twitter.com/hpccouncil](http://www.twitter.com/hpccouncil)
YouTube: [www.youtube.com/user/hpcadvisorycouncil](http://www.youtube.com/user/hpcadvisorycouncil)
Thank You
HPC Advisory Council