NEMO5
Performance Benchmarking
July 2017
• **NEMO5**
  - the 5th edition of NanoElectronics MOdeling (NEMO) Tools of the Klimeck group
  - incorporates the core concepts and insights gained from 15 years of development of:
    • NEMO-1D, NEMO-3D, NEMO-3D-Peta and OMEN
  - Licensing agreements for academic and commercial use are available; Free with restrictions for academic use
  - Software distribution and support is handled through the NEMO5 distribution and support group on nanoHUB

• **The core capabilities lie in the atomic-resolution calculation of nanostructure properties**
  - strain relaxation, phonon modes, electronic structure using the tight-binding model, self consistent Schroedinger-Poisson calculations, and quantum transport.
Test Cluster

- **Dell PowerEdge R720 32-node (640-core) “Jupiter” cluster**
  - Dual-Socket 10-Core Intel E5-2680v2 @ 2.80 GHz CPUs
  - Memory: 64GB memory, DDR4 1600 MHz
  - BIOS: Maximum Performance; Memory Snoop Mode: Home Snoop
  - OS: RHEL 6.5
  - MLNX_OFED_LINUX-4.0-2.0.0.1 InfiniBand SW stack
  - Mellanox ConnectX-4 EDR 100Gb/s InfiniBand Adapters
  - Mellanox Switch-IB SB7800 36-port EDR 100Gb/s InfiniBand Switch
  - Mellanox Spectrum SN2700 32-port 100Gb/s Ethernet Switch
  - Dell InfiniBand-Based Lustre Storage based on Dell PowerVault MD3460 and MD3420
  - Compilers: Intel Compilers 2016.4.258
  - MPI: Intel MPI 5.1.3
  - Python: 2.7.13
  - MPI Profiler: IPM
NEMO5 Performance – Interconnect

- InfiniBand provides 17% higher performance versus Ethernet RoCE
- InfiniBand provides 51% higher performance versus Ethernet TCP
- Ethernet RoCE provides 29% higher performance versus Ethernet TCP

Higher is better
NEMO5 Profiling – Time Spent by MPI Calls

- For the most time consuming MPI calls (as % of MPI time):
  - MPI_Send (67% of MPI time, 18% overall)
  - MPI_Barrier (29% of MPI time, 8% overall)
  - MPI_Recv (4% of MPI time, 1% overall)
For the most time consuming MPI calls is MPI_Send
- Some load imbalance observed in the time spent by the MPI tasks
InfiniBand consumed the least percentage time in overall runtime
- InfiniBand consumed around 24% of overall runtime, while TCP consumes up to 49%
NEMO5 Summary

- NEMO5 is the 5th edition of NanoElectronics MOdeling (NEMO) Tools
- InfiniBand provides 17% higher performance versus Ethernet RoCE
- InfiniBand provides 51% higher performance versus Ethernet TCP
- Ethernet RoCE provides 29% higher performance versus Ethernet TCP
- RDMA over Ethernet improved scalability on Ethernet network for NEMO5
- For the most time consuming MPI calls (as % of MPI time):
  - MPI_Send (67% of MPI time, 18% overall)
  - MPI_Barrier (29% of MPI time, 8% overall)
  - MPI_Recv (4% of MPI time, 1% overall)
- For the most time consuming MPI calls is MPI_Send
  - Some load imbalance observed in the time spent by the MPI tasks
Thank You