HPC ADVISORY COUNCIL

GROMACS
Performance Benchmark and Profiling

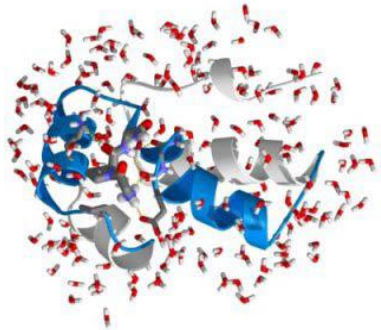
April 2017

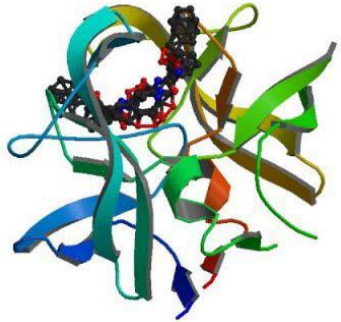
Note

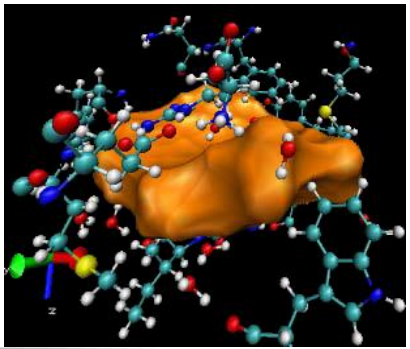
- **The following research was performed under the HPC Advisory Council activities**
  - Compute resource - HPC Advisory Council Cluster Center
- **The following was done to provide best practices**
  - GROMACS performance overview
  - Understanding GROMACS communication patterns
  - Ways to increase GROMACS productivity
- **For more info please refer to**
  - http://www.gromacs.org/

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 nonbonded interactions
  - Ongoing development to extend GROMACS with interfaces both to Quantum Chemistry and Bioinformatics/databases
  - An open source software released under the GPL







Objectives

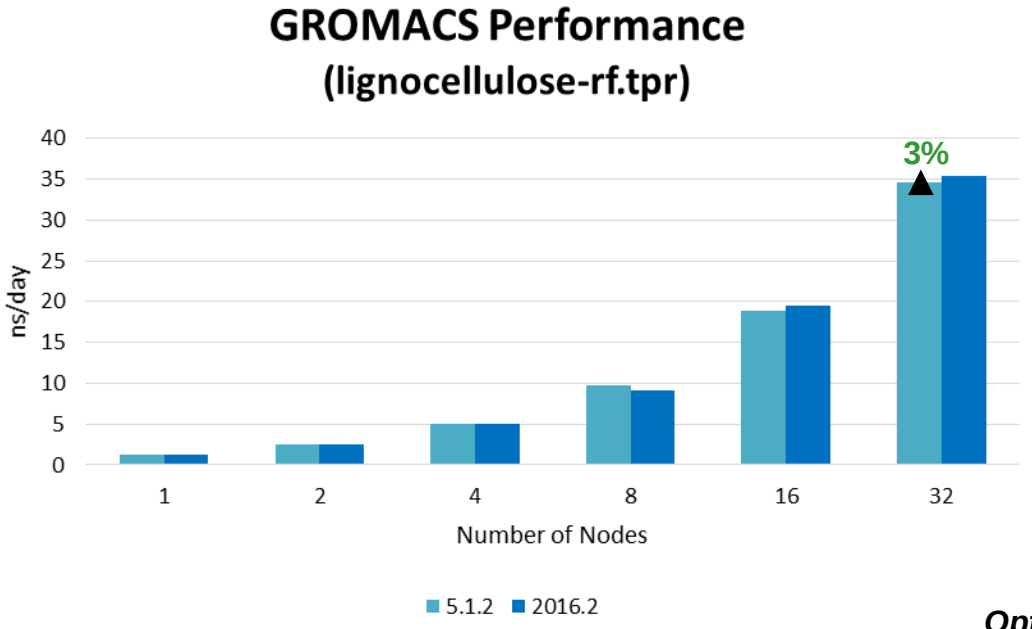
- **The presented research was done to provide best practices**
  - GROMACS performance benchmarking
    - MPI Library performance comparison
    - Interconnect performance comparison
    - CPUs comparison
    - Optimization tuning
- **The presented results will demonstrate**
  - The scalability of the compute environment/application
  - Considerations for higher productivity and efficiency

Test Cluster Configuration

- **HPE ProLiant DL360 Gen9 128-node (4096-core) "Hercules" cluster**
  - Dual-Socket 16-Core Intel E5-2697A v4 @ 2.60 GHz CPUs
  - Memory: 256GB memory, DDR4 2400 MHz, Memory Snoop Mode in BIOS sets to Home Snoop
  - OS: RHEL 7.2, MLNX_OFED_LINUX-3.4-2.0.0.0 InfiniBand SW stack
- **Mellanox ConnectX-4 EDR 100Gb/s InfiniBand Adapters**
- **Mellanox Switch-IB SB7800 36-port EDR 100Gb/s InfiniBand Switch**
- **Intel® Omni-Path Host Fabric Interface (HFI) 100Gb/s Adapter**
- **Intel® Omni-Path Edge Switch 100 Series**
- **MPI: Intel MPI 2017, Open MPI 2.02**
- **Application: GROMACS 2016.2 and 5.1.2**
- **Benchmarks:**
  - Benchmark datasets: lignocellulose3M_rff
  - http://www.prace-i.eu/UEABS/GROMACS/1.2/GROMACS_TestCaseB.tar.gz

GROMACS Performance – GROMACS versions

- **Small performance gain in the latest GROMACS version**
  - About 3% better performance seen on GROMACS version 2016.2 than 5.1.2

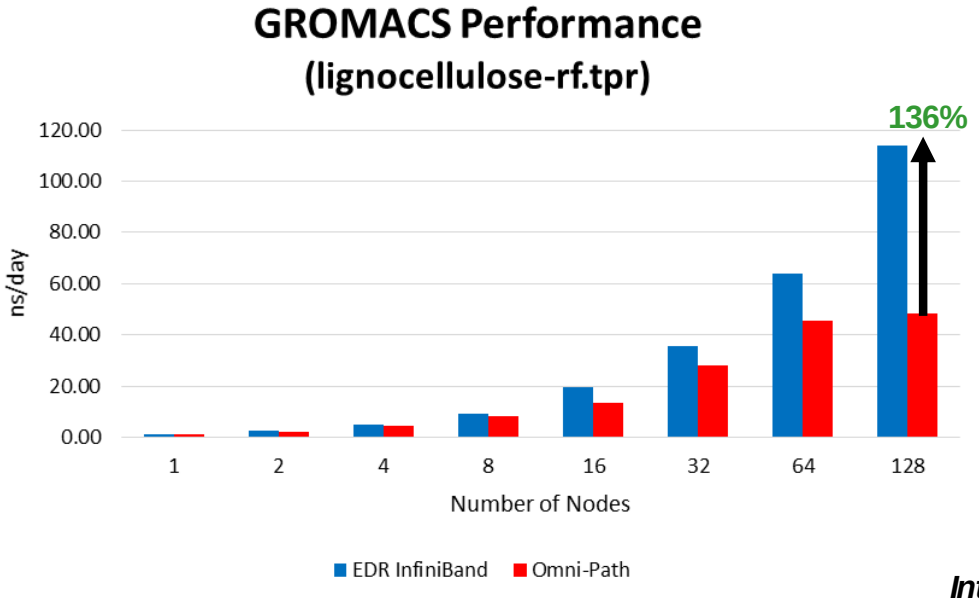


Higher is better

Optimized parameters used

GROMACS Performance – Interconnects

- **EDR InfiniBand enables higher scalability than Omni-Path for GROMACS**
  - InfiniBand delivers 136% better scaling versus Omni-Path for 128 nodes
  - 64 InfiniBand nodes delivers 33% higher performance compared to 128 Omni-Path nodes

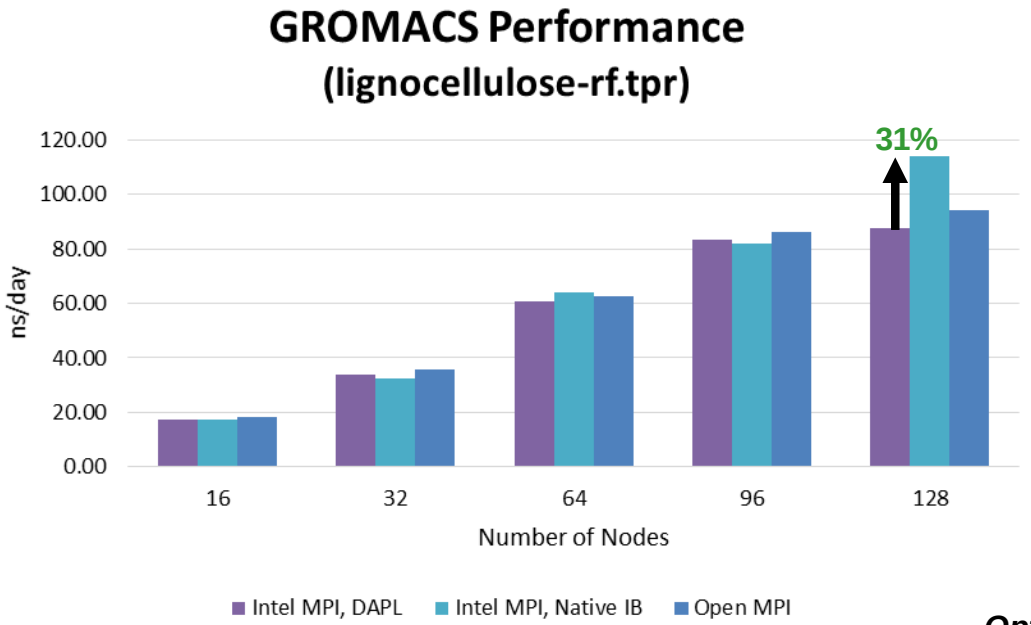


Higher is better

Intel MPI

GROMACS Performance – MPI

- **Intel MPI includes multiple transport providers for running on InfiniBand fabrics**
  - Native IB provides up to 31% better scaling than DAPL provider at 128 nodes

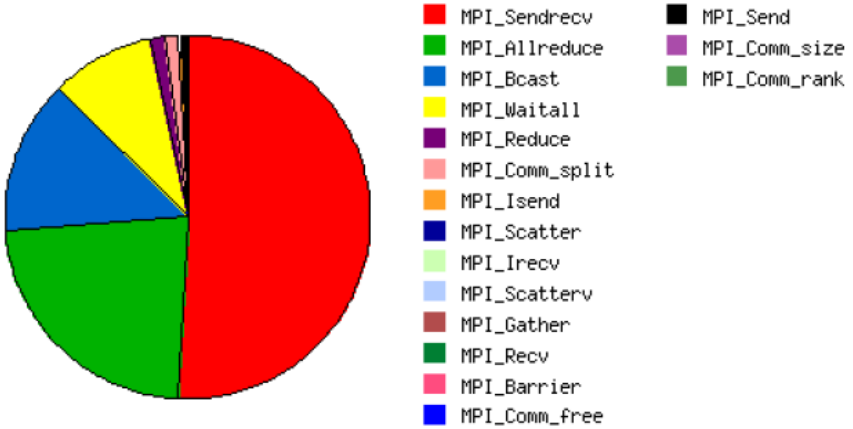


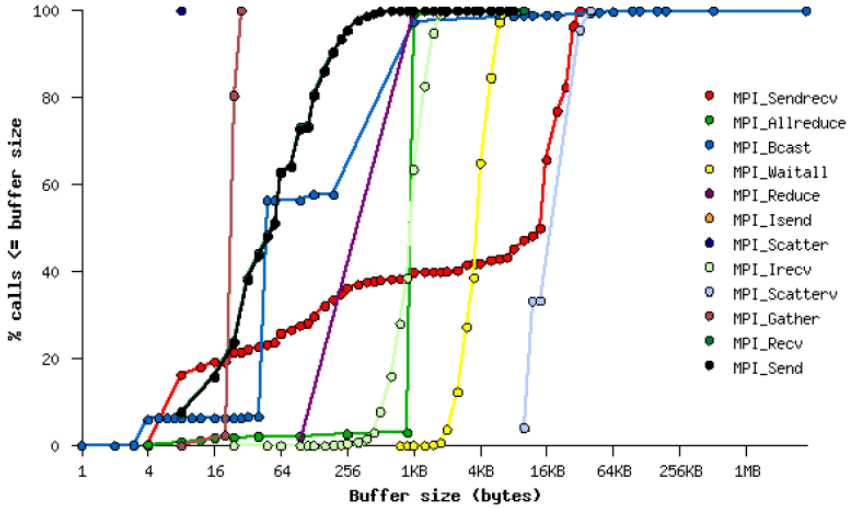
Higher is better

Optimized parameters used

GROMACS Profiling – % of MPI Calls

- **For the most time consuming MPI calls (as % of MPI time):**
  - MPI_Iprobe (51%), MPI_Allreduce (23%), MPI_Bcast (16%), MPI_Waitall (9%)
- **MPI communication accounts for 16% of overall wall clock time**





32 Nodes / 1024 Processes

GROMACS Summary

- **Small performance gain in the latest GROMACS version**
  - About 3% better performance seen on GROMACS version 2016.2 than 5.1.2
- **EDR InfiniBand enables higher scalability than Omni-Path for GROMACS**
  - InfiniBand delivers 136% better scaling versus Omni-Path for 128 nodes
  - 64 InfiniBand nodes delivers 33% higher performance compared to 128 Omni-Path nodes
- **Intel MPI includes multiple transport providers for running on InfiniBand fabrics**
  - Native IB transport provides up to 31% better scaling than DAPL provider at 128 nodes
- **For the most time consuming MPI calls (as % of MPI time):**
  - MPI_Iprobe (51%), MPI_Allreduce (23%), MPI_Bcast (16%), MPI_Waitall (9%)
  - MPI communciation accounts for 16% of overall wall clock time

Thank You

HPC Advisory Council