

OpenFOAM Performance With MPI Collectives Acceleration

Feb 2011

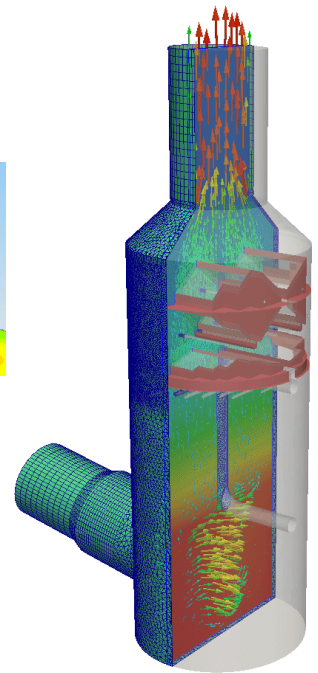
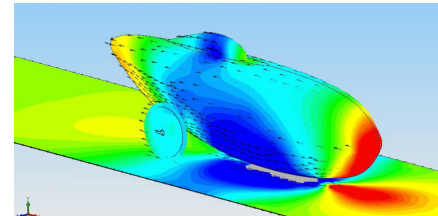


Open  FOAM

- **The following research was performed under the HPC Advisory Council HPC|Scale and HPC|works working group activities**
 - Participating vendors: HP, Intel, Mellanox
 - Compute resource - HPC Advisory Council Cluster Center
- **We would like to thank Mellanox for providing early access to its MPI Collectives Acceleration solution (FCA version 2.0)**
- **For more info please refer to**
 - <http://www.hp.com/go/hpc>
 - www.intel.com
 - www.mellanox.com
 - <http://www.open CFD.co.uk/openfoam>

- **Open FOAM utilization of MPI collectives operations**
 - Application profiling
 - Understanding OpenFOAM communication patterns
- **Preview on available MPI collectives accelerations**
- **First performance results with Open FOAM**
 - Utilizing MPI collectives accelerations

- **OpenFOAM® (Open Field Operation and Manipulation) CFD Toolbox in an open source CFD applications that can simulate**
 - Complex fluid flows involving
 - Chemical reactions
 - Turbulence
 - Heat transfer
 - Solid dynamics
 - Electromagnetics
 - The pricing of financial options
- **OpenFOAM support can be obtained from OpenCFD Ltd**

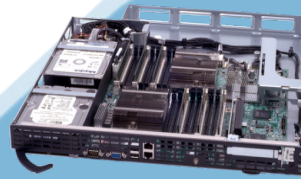


- **HP ProLiant SL2x170z G6 16-node cluster**
 - Six-Core Intel X5670 @ 2.93 GHz CPUs
 - Memory: 24GB per node
 - OS: CentOS5U5, OFED 1.5.1 InfiniBand SW stack
- **Mellanox ConnectX-2 adapters and switches**
- **MPI: Open MPI 1.4.3, Platform MPI 8.1.1**
- **Mellanox Fabric Collective Accelerator™ (FCA™) version 2.0**
- **Application: OpenFOAM 1.7.1**
- **Benchmark Workload**
 - Lid-driven cavity flow (2000x2000)

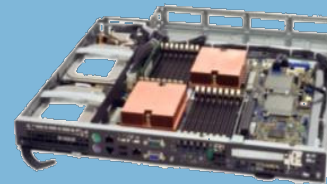
- **Solution-optimized for extreme scale out**



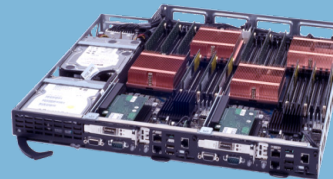
ProLiant z6000 chassis
Shared infrastructure
– fans, chassis, power



ProLiant SL160z G6 ProLiant SL165z G7
Large memory
-memory-cache apps



ProLiant SL170z G6
Large storage
-Web search and database apps



ProLiant SL2x170z G6
Highly dense
- HPC compute and
web front-end apps

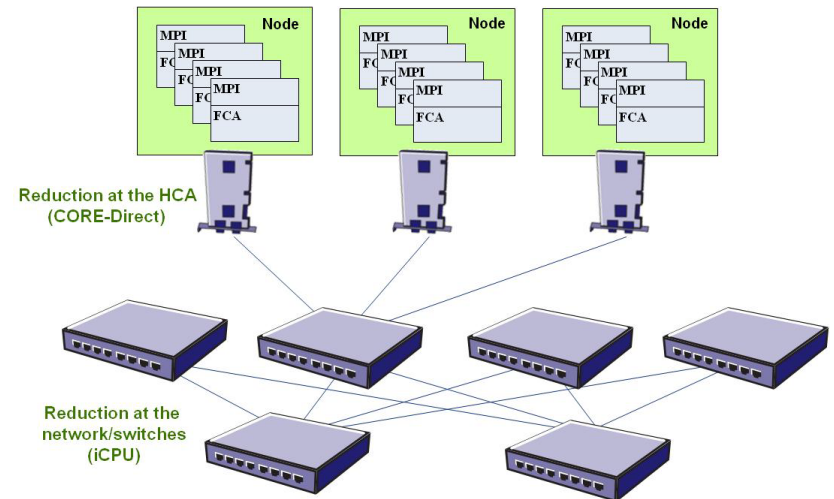
Save on cost and energy -- per node, rack and data center

Mix and match configurations

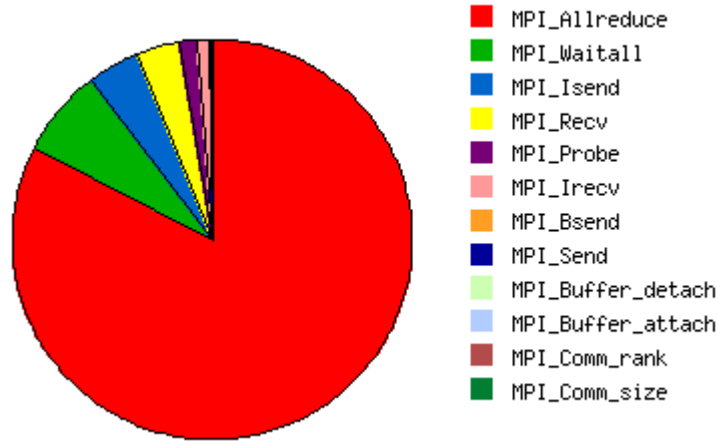
Deploy with confidence

* SPECpower_ssj2008
www.spec.org
17 June 2010, 13:28

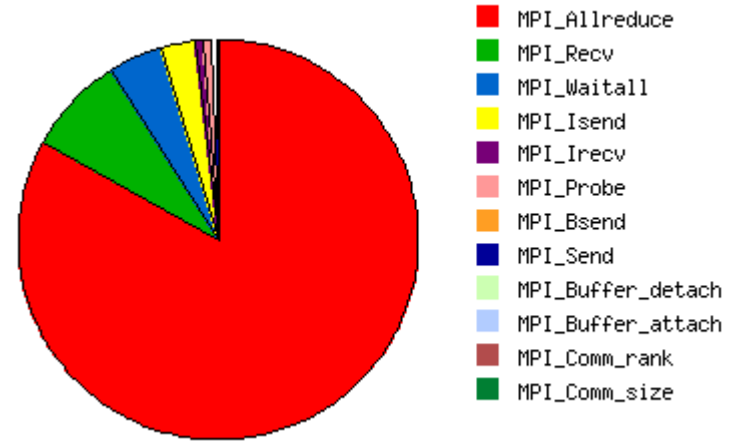
- **Mellanox Fabric Collectives Accelerator (FCA)**
 - Utilized hardware accelerations on the adapter (CORE-Direct)
 - Utilized managed switches capabilities (iCPU)
 - Accelerating MPI collectives operations
 - The world first complete solution for MPI collectives offloads
- **FCA 2.0 supports accelerations/offloading for**
 - MPI Barrier
 - MPI Broadcast
 - MPI AllReduce and Reduce
 - MPI AllGather and AllGatherV



- **MPI_Addreduce generates most communication overhead**



96 Processes (8 nodes)

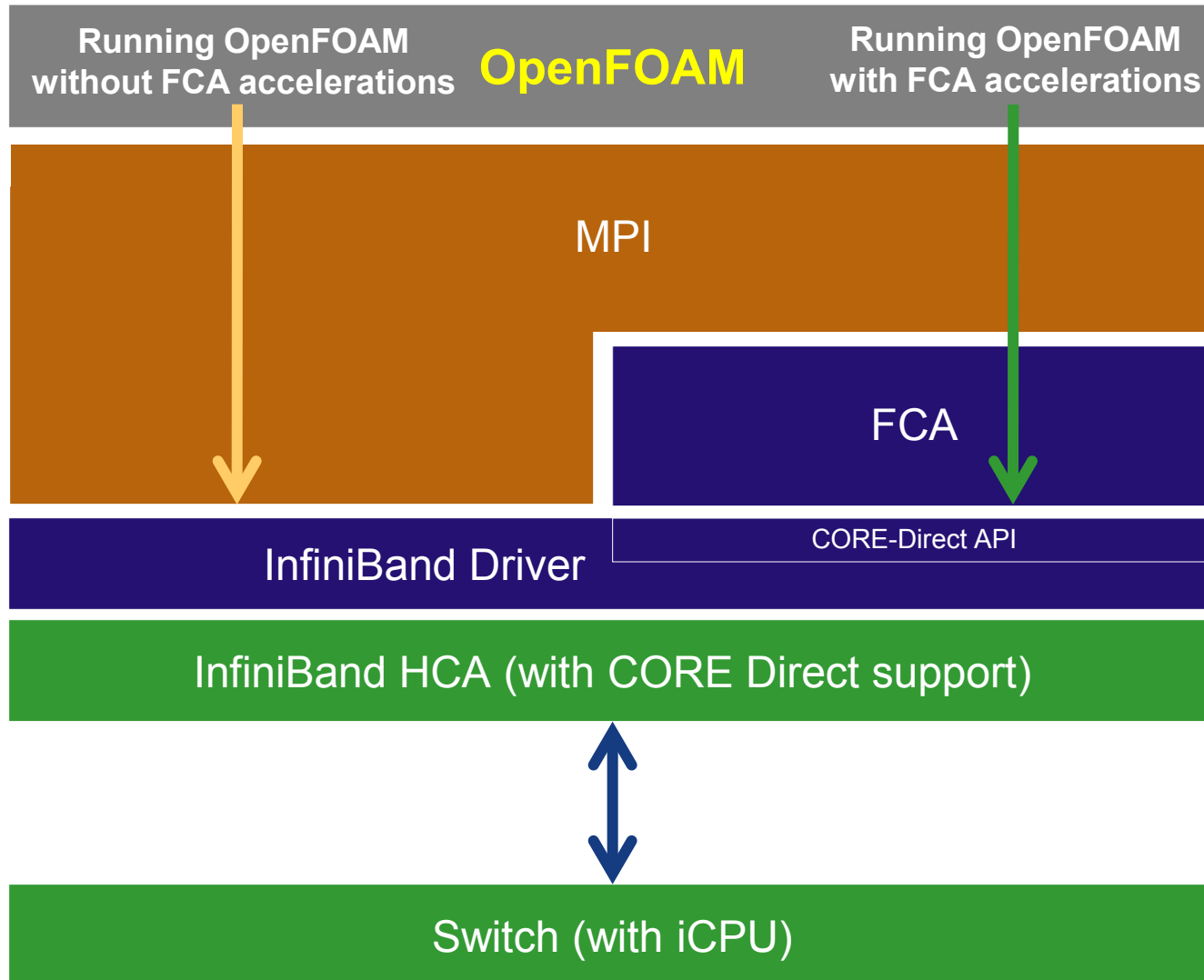


192 Processes (16 nodes)

12 Cores/Node

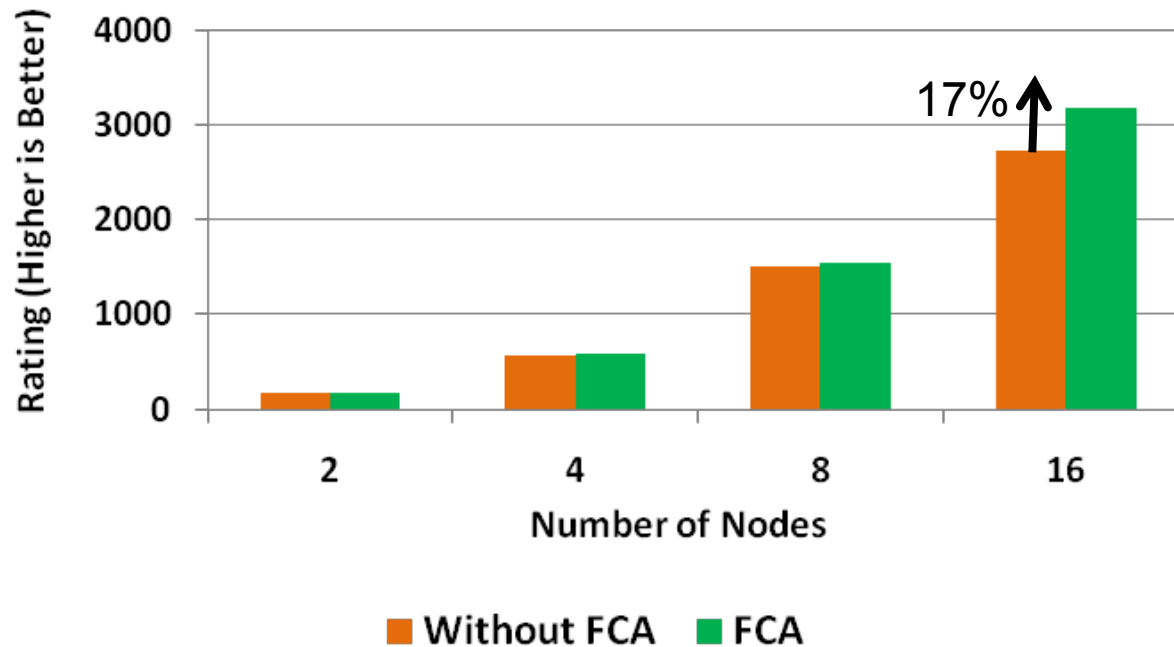
- **OpenFoam mainly utilizes MPI AllReduce collective operation**
- **MPI AllReduce is being used for data collection and reduction**
 - Data collocation as a results of parallel computations done in the cluster nodes
 - MPI AllReduce is being accelerated within FCA

Software Layers Overview



- **FCA accelerates OpenFOAM performance up to 17%**
 - At 16 nodes, 192 cores
 - Performance benefit increases with cluster size – higher benefit expected at larger scale

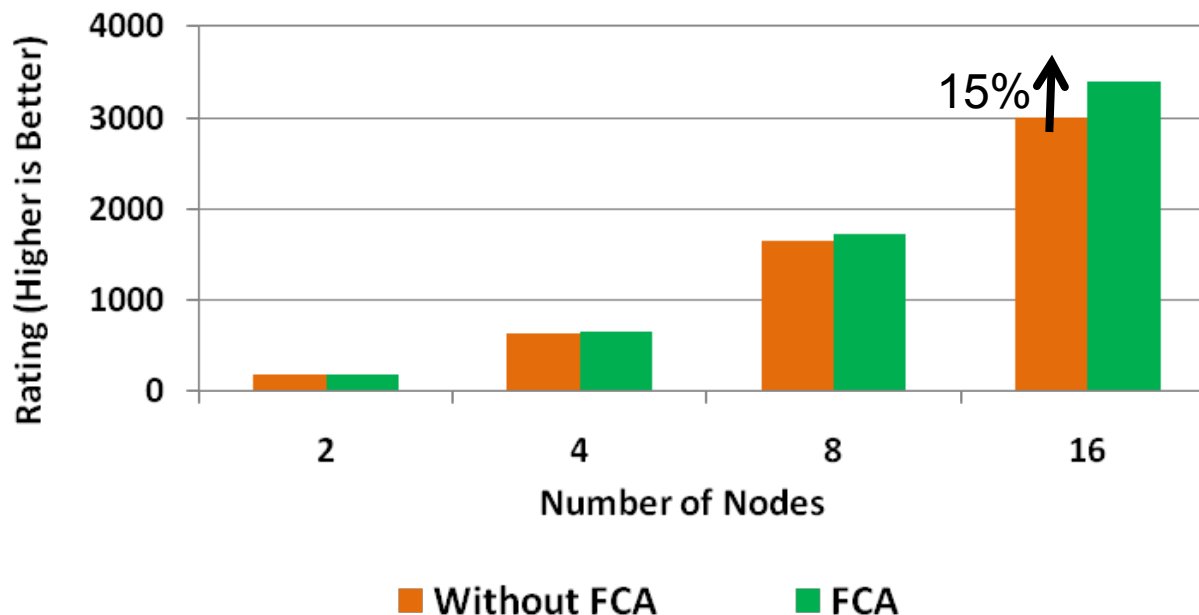
OpenFOAM Benchmark (Lid-driven cavity flow, Open MPI)



Higher is better

- **FCA accelerates OpenFOAM performance up to 15%**
 - At 16 nodes, 192 cores
 - Performance benefit increases with cluster size – higher benefit expected at larger scale

OpenFOAM Benchmark (Lid-driven cavity flow, PMPI)

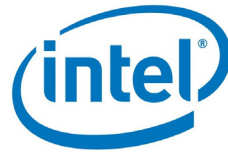


Higher is better

- **MPI Collectives accelerations can dramatically accelerate HPC applications performance**
 - Large scale systems will greatly benefit from such acceleration
 - The data presented here reviewed potential acceleration at small scale
- **FCA package has proven to accelerate application even at small scale**
 - Nearly 20% at 16 nodes for OpenFOAM CFD application
 - Higher performance boost expected at larger scale
- **OpenFOAM mainly utilizes the MPI AllReduce collective operation**
 - For data gathering/reduction
 - Other application with usage of more collectives operations expected to show greater benefit of FCA

Thank You

HPC Advisory Council



Open  FOAM



All trademarks are property of their respective owners. All information is provided "As-Is" without any kind of warranty. The HPC Advisory Council makes no representation to the accuracy and completeness of the information contained herein. HPC Advisory Council Mellanox undertakes no duty and assumes no obligation to update or correct any information presented herein