HPC·AI
ADVISORY COUNCIL
NETWORK OF EXPERTISE

MetaComp ICFD++
Performance Benchmarking and Profiling

Oct 2018

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**
  - MetaComp ICFD++ performance overview over Intel Skylake (SKL) based platforms
  - Understanding MetaComp ICFD++ communication patterns

- **More info on MetaComp ICFD++ Application**
  - http://www.metacomptech.com/index.php/features/icfd

MetaComp ICFD++

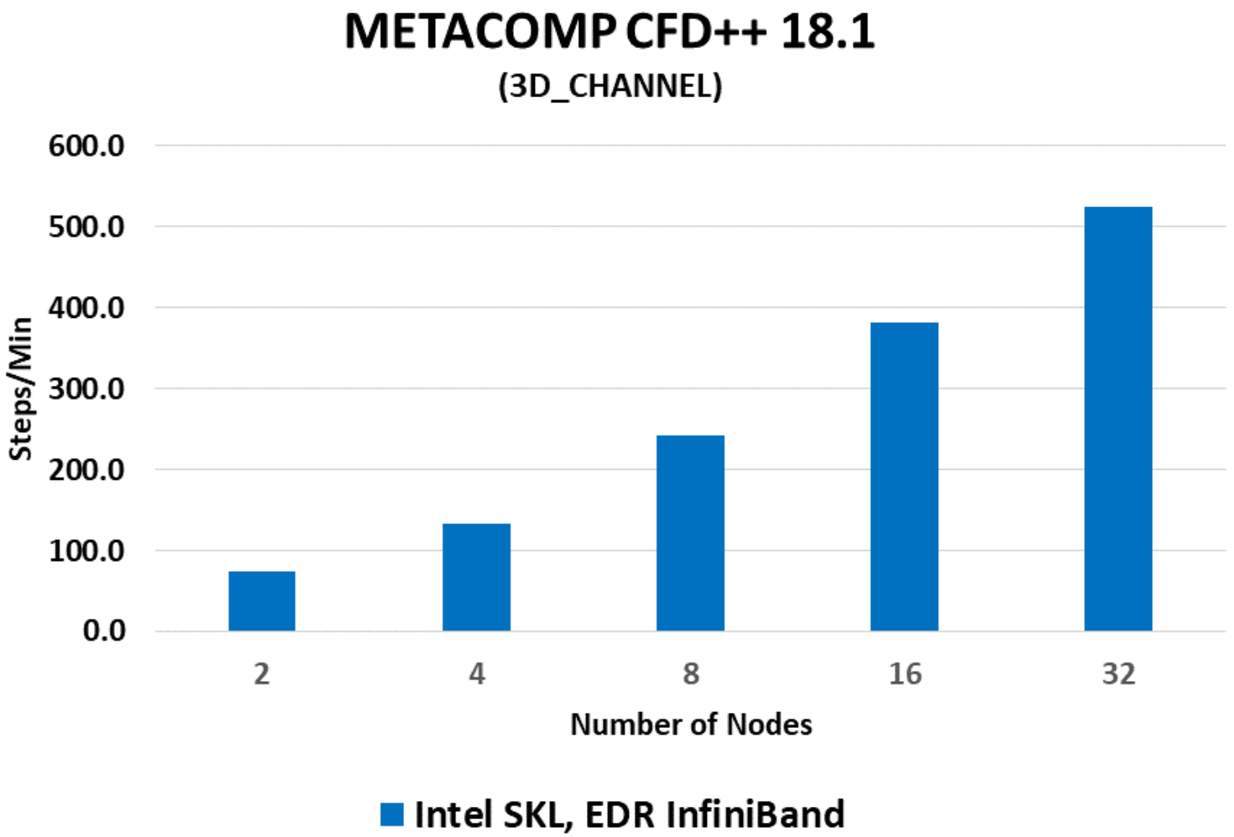
- **Computational Fluid Dynamics (CFD)**
  - Enables the study of the dynamics of things that flow
  - Enable better understanding of qualitative and quantitative physical phenomena in the flow which is used to improve engineering design
- **CFD brings together a number of different disciplines**
  - Fluid dynamics, mathematical theory of partial differential systems, computational geometry, numerical analysis, Computer science
- **MetaComp ICFD++ is a part of MetaComp's CFD software suite**
  - ICFD++ can be used to simulate compressible and incompressible fluids and flows, unsteady and steady flows, large range of speed regimes including low speeds through subsonic, transonic, supersonic and hypersonic speeds, laminar and turbulent flows, various equations of state

Cluster Configuration

- **Helios Cluster**
  - Supermicro SYS-6029U-TR4 / Foxconn Groot 1A42USF00-600-G 32-node cluster
  - Dual Socket Intel(R) Xeon(R) Gold 6138 CPU @ 2.00GHz
  - Mellanox ConnectX-5 EDR 100Gb/s InfiniBand/VPI adapters
  - Mellanox Switch-IB 2 SB7800 36-Port 100Gb/s EDR InfiniBand switch
  - Memory: 192GB DDR4 2677MHz RDIMMs per node
  - 1TB 7.2K RPM SSD 2.5" hard drive per node

- **Software**
  - OS: RHEL 7.5, MLNX_OFED 4.4
  - MPI: HPC-X 2.2
  - MetaComp ICFD++ 18.1

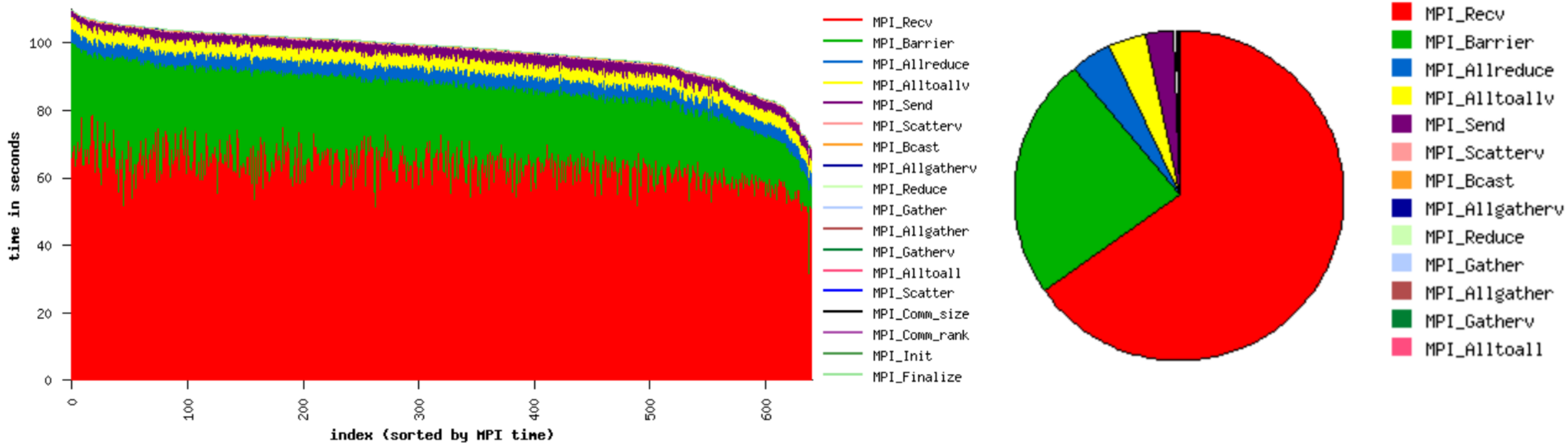
MetaComp ICFD++ Performance



Higher is better

MetaComp CDF++ Application Profile (16 nodes Intel SKL)

- **15.83% MPI and WallClock of 616 seconds**



MetaComp CDF++ Application Profile (16 nodes Intel SKL)

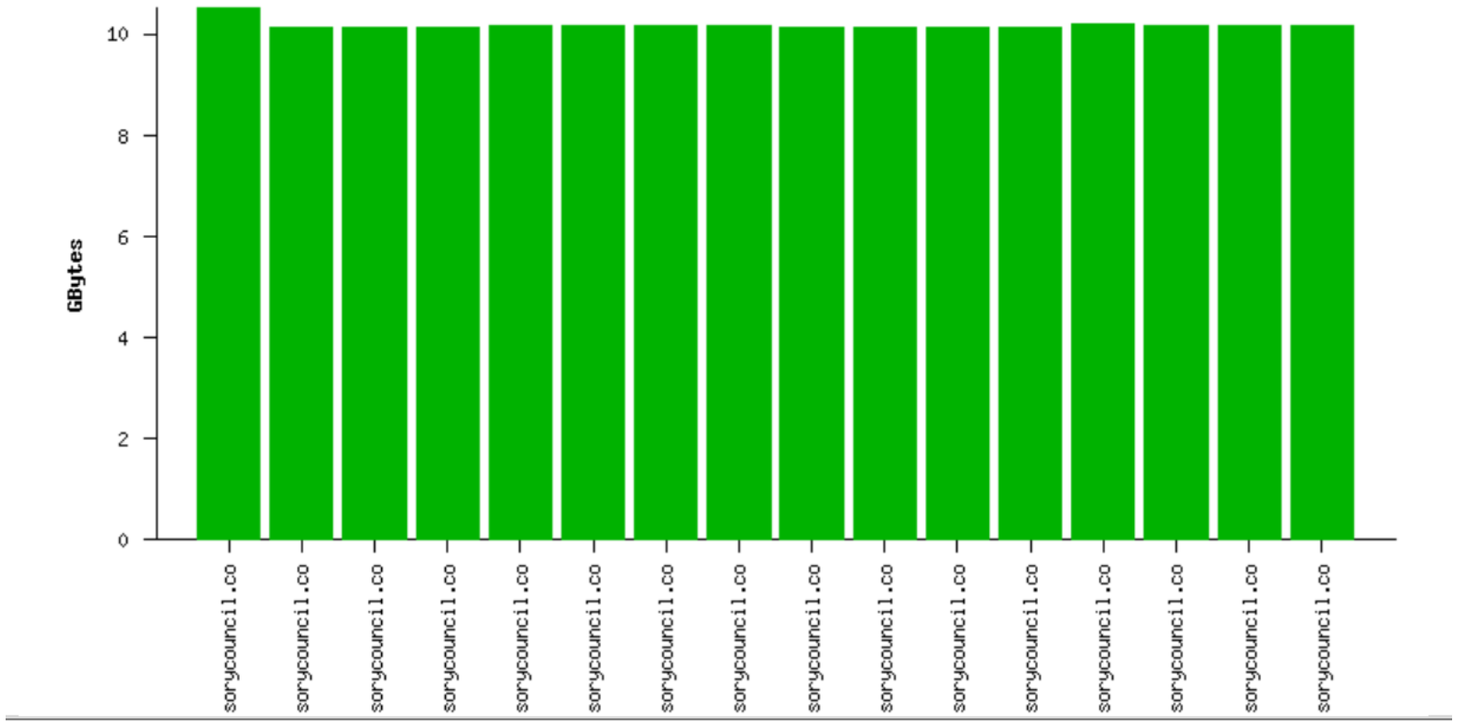
HPC-AI ADVISORY COUNCIL
NETWORK OF EXPERTISE

- **Communication pattern**

Communication Event Statistics (% detail, --- error)									
	Comm Size	Buffer Size	Ncalls	Total Time	Avg Time	Min Time	Max Time	%MPI	%Wall
MPI_Recv	0	4	339125648	2.286226e+04	6.741531e-05	0.000000e+00	5.296400e+00	36.64	5.80
MPI_Barrier	640	0	30751360	1.484294e+04	4.826759e-04	4.768400e-06	3.636400e-01	23.79	3.77
MPI_Recv	0	10240	20525888	7.472028e+03	3.640295e-04	9.536700e-07	7.681100e+00	11.98	1.90
MPI_Alltoallv	640	0	8320	2.315986e+03	3.636400e-01	2.620000e-03	3.565600e+00	3.71	0.59
MPI_Recv	0	16384	31920000	1.024091e+03	3.208305e-05	9.536700e-07	7.217200e-03	1.64	0.26
MPI_Recv	0	20480	28256000	9.652150e+02	3.415965e-05	1.907300e-06	6.856000e-03	1.55	0.24
MPI_Allreduce	640	4	23056000	9.453686e+02	4.100315e-05	4.768400e-06	2.702800e-02	1.52	0.24
MPI_Recv	0	32768	20912000	8.283904e+02	3.961316e-05	3.814700e-06	5.526100e-03	1.33	0.21
MPI_Recv	0	24576	21776000	7.891171e+02	3.623793e-05	1.907300e-06	6.243000e-03	1.26	0.20
MPI_Recv	0	28672	17856000	7.001990e+02	3.921365e-05	2.861000e-06	6.178900e-03	1.12	0.18
MPI_Recv	0	8192	22704224	6.423066e+02	2.829018e-05	0.000000e+00	1.089400e+00	1.03	0.16
MPI_Recv	0	12288	19024000	5.476966e+02	2.878977e-05	9.536700e-07	5.672900e-03	0.88	0.14
MPI_Recv	0	14336	17872000	4.946790e+02	2.767899e-05	9.536700e-07	5.768800e-03	0.79	0.13
MPI_Recv	0	40960	9792000	4.524363e+02	4.620469e-05	4.768400e-06	6.209100e-03	0.73	0.11
MPI_Allreduce	640	8	15365760	4.481250e+02	2.916387e-05	1.907300e-06	9.109000e-03	0.72	0.11
MPI_Recv	0	4096	42195345	4.011041e+02	9.505885e-06	0.000000e+00	9.176200e-01	0.64	0.10
MPI_Recv	0	5120	19632173	3.882852e+02	1.977800e-05	0.000000e+00	2.682600e-01	0.62	0.10
MPI_Allreduce	640	24	10242560	3.540626e+02	3.456778e-05	5.006800e-06	1.124800e-02	0.57	0.09
MPI_Recv	0	6144	12848153	3.276638e+02	2.550279e-05	0.000000e+00	5.742500e-01	0.53	0.08

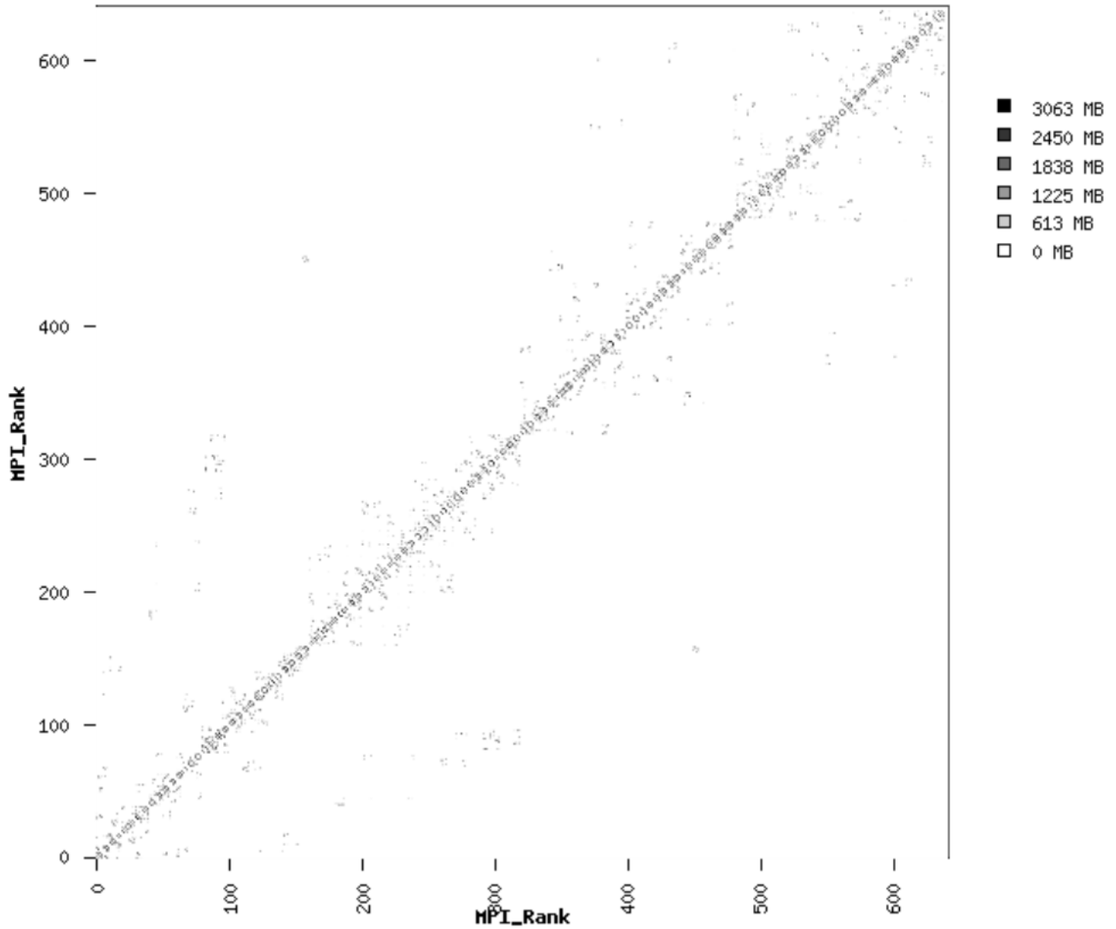
MetaComp CDF++ Application Profile (8 nodes Intel SKL)

- **Memory usage: ~10GB per node**



MetaComp ICFD++ Application Profile (16 nodes Intel SKL)

- **Communication is done mainly between near ranks**



MetaComp ICFD++ Summary

- **MetaComp performance testing over Intel SKL based platform**
  - An range of 38% scaling was achieved from 16 to 32 nodes

- **MetaComp profiling on “3D_CHANNEL”**
  - MPI communication accounts for 15.83% of overall wall clock time at 8 nodes
  - MPI_Recv is 66% of MPI, MPI_Barrier is 28% of MPI and MPI_Allreduce is 6% of MPI
  - Most communication is done between ranks that are close to each other

HPC·AI ADVISORY COUNCIL
NETWORK OF EXPERTISE

Thank You