



MILC

Performance Benchmark and Profiling

Dec 2020

- **The following research was performed under the HPC Advisory Council activities**
 - Compute resource – AMD internal system
- **The following was done to provide best practices**
 - MILC performance overview over Intel based platforms
 - Understanding MILC communication patterns
- **More info on MILC**
 - https://github.com/milc-qcd/milc_qcd

- **The MIMD Lattice Computation (MILC) represents part of a set of codes used to study quantum chromodynamics (QCD), the theory of the strong interactions of subatomic physics**
- **It performs simulations of four dimensional SU(3) lattice gauge theory on MIMD parallel machines**
- **"Strong interactions" are responsible for binding quarks into protons and neutrons and holding them all together in the atomic nucleus**
- **The MILC collaboration has produced application codes to study several different QCD research areas**

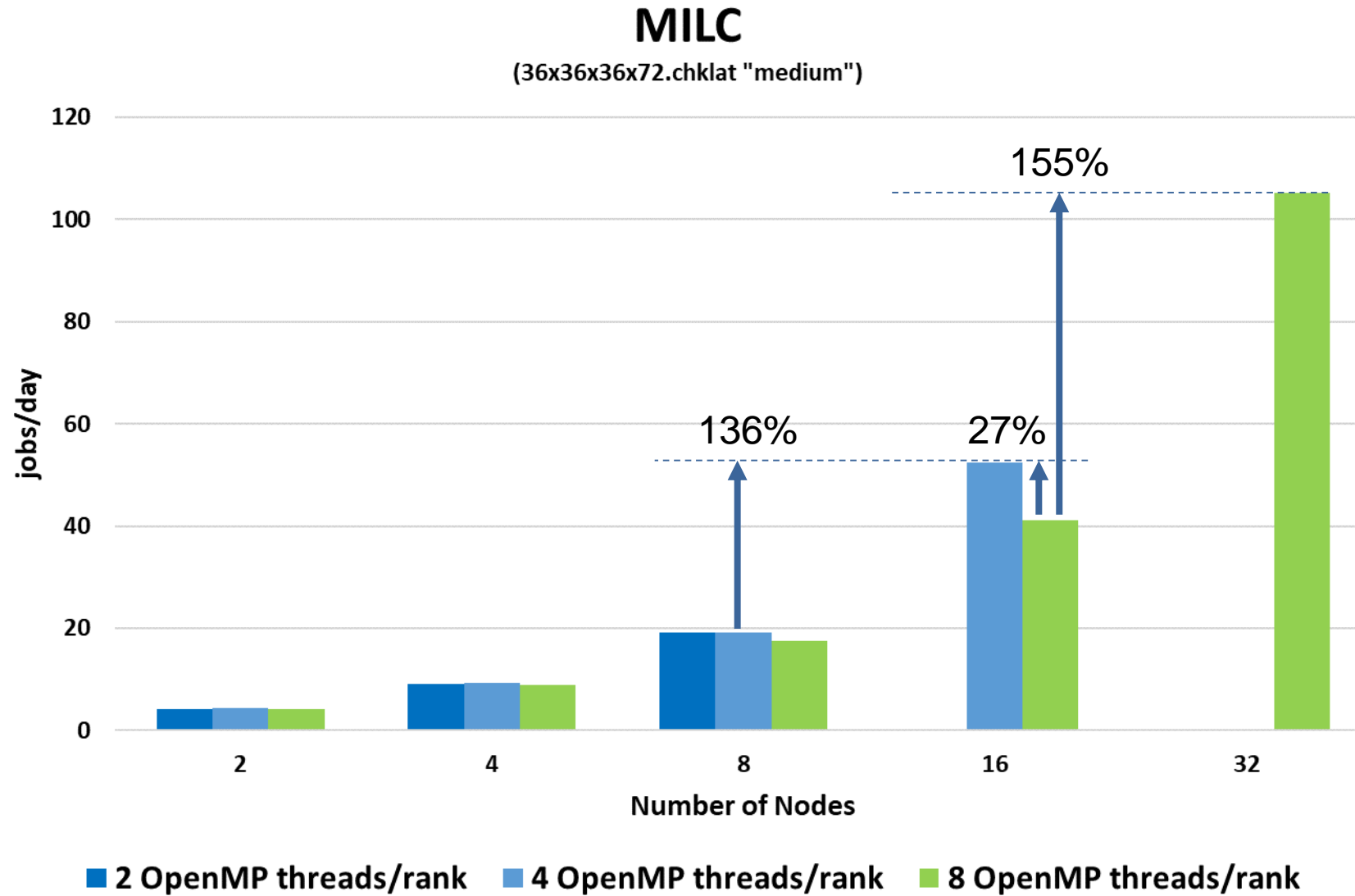
- **Atlantis cluster**

- AMD EPYC 7742 (Rome) @ 2.25 GHz (64 cores/socket, 2 sockets/node)
- ConnectX-6 HDR InfiniBand
- Quantum Switch HDR InfiniBand
- Memory: 512GB DDR4 2677MHz RDIMMs per node
- NFS Storage

- **Software**

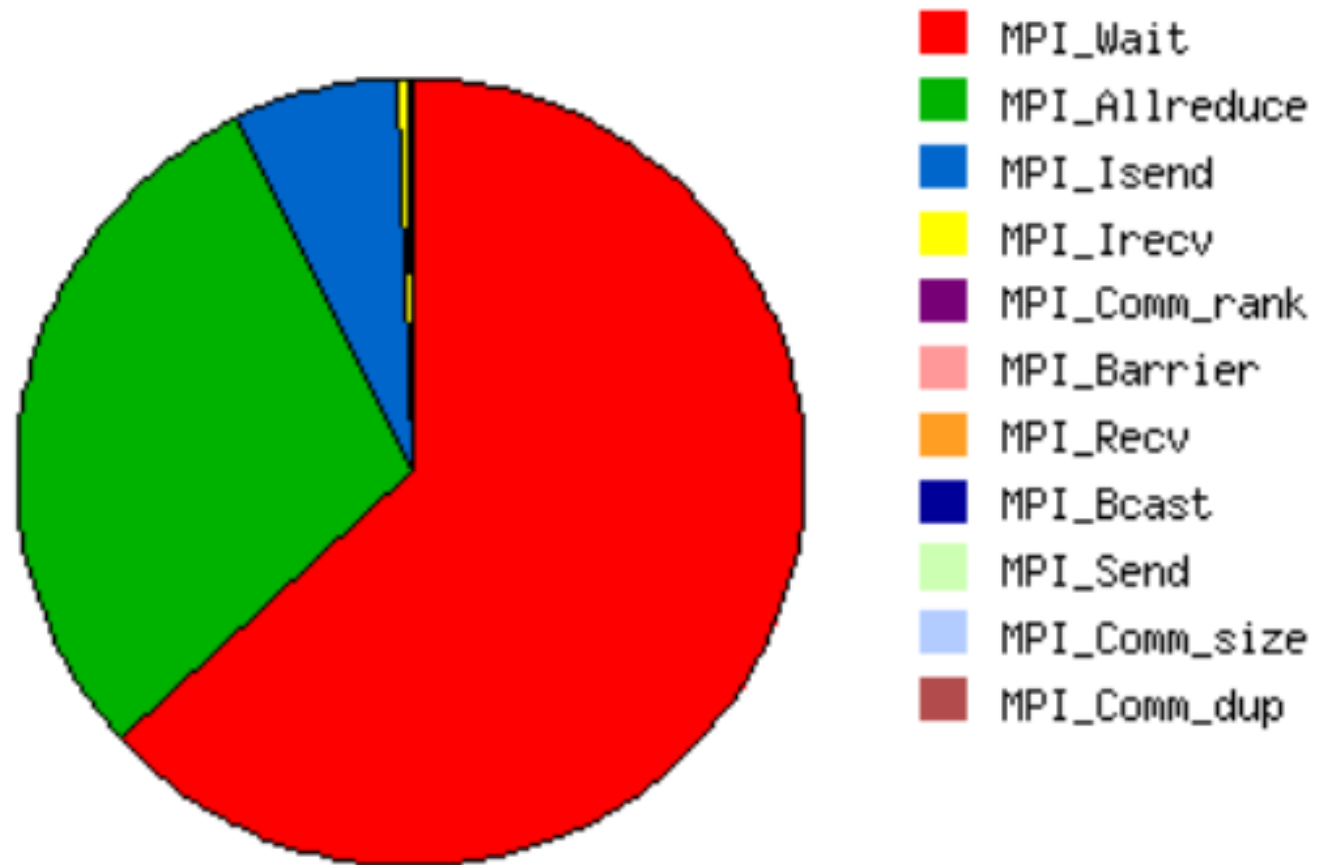
- OS: RHEL 8.1, MLNX_OFED 5.0.2
- MPI: HPC-X 2.7
- MLIC: develop branch of https://github.com/milc-qcd/milc_qcd, commit 77d89f04bdc8fb55ebd40d555cb1f54c4b39d105 (May 29 1 2020)

MILC Performance – Super Linear Scalability



Note: 2 OpenMP threads/rank on 16/32 nodes and 4 OpenMP threads/rank on 32 could not be tested due to the problem decomposition in this case.

- **55% MPI Communication**



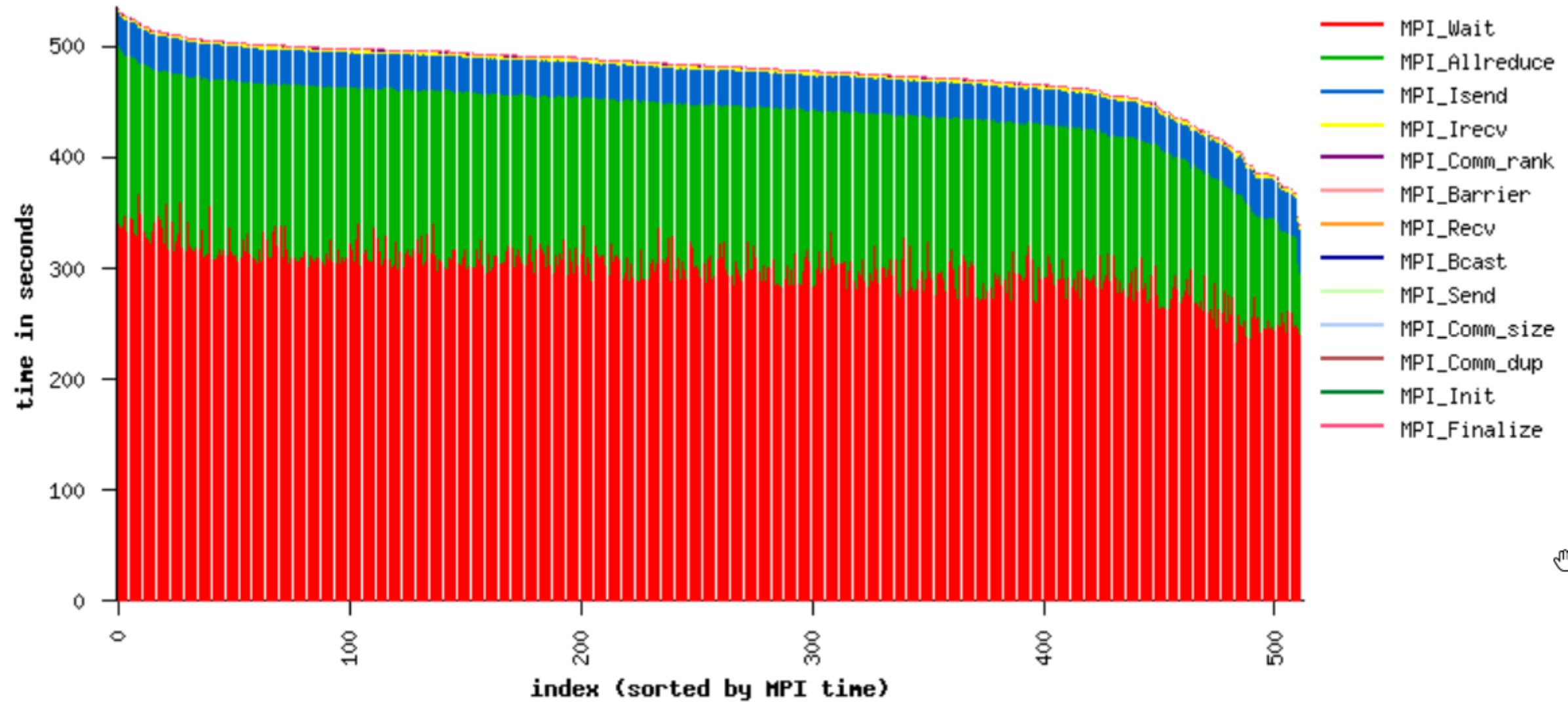
MILC Profile on 32 Nodes

- 63% of MPI Communication spent on MPI_Wait
- 29% MPI Allreduce 8 bytes
- Async send and receive communication

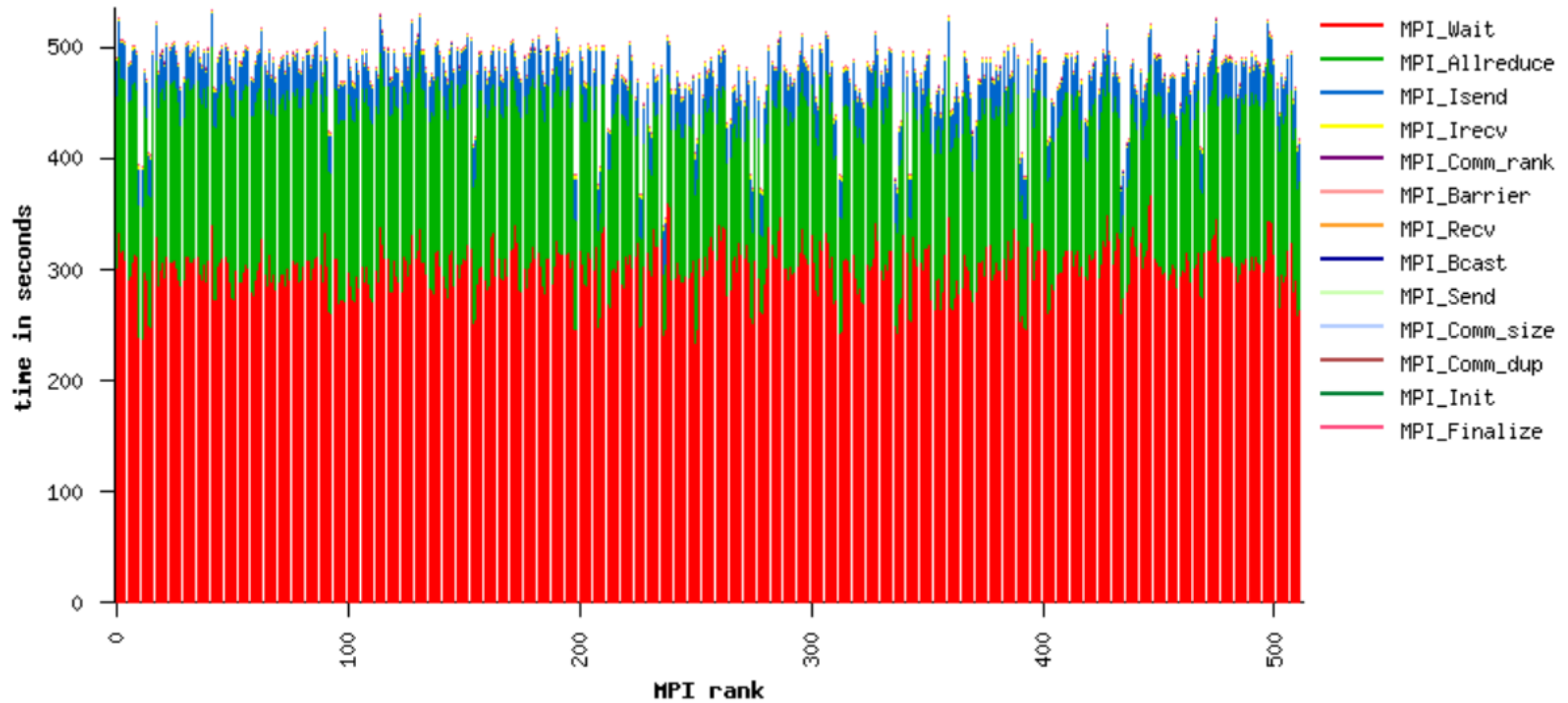
Communication Event Statistics (% detail, --- error)									
	Comm Size	Buffer Size	Ncalls	Total Time	Avg Time	Min Time	Max Time	%MPI	%Wall
MPI_Wait	0	0	11154931712	1.539791e+05	1.380368e-05	0.000000e+00	2.436600e-01	63.24	35.27
MPI_Allreduce	512	8	634250240	7.160776e+04	1.129014e-04	3.814700e-06	2.439000e-01	29.41	16.40
MPI_Isend	0	49152	5092916736	1.396905e+04	2.742838e-06	0.000000e+00	1.611100e-02	5.74	3.20
MPI_Isend	0	98304	445880320	2.264142e+03	5.077914e-06	0.000000e+00	2.187000e-03	0.93	0.52
MPI_Irecv	0	98304	445880320	6.359448e+02	1.426268e-06	0.000000e+00	1.997900e-03	0.26	0.15
MPI_Irecv	0	49152	5092916736	5.826193e+02	1.143980e-07	0.000000e+00	1.364000e-03	0.24	0.13
MPI_Allreduce	512	32	715776	1.049305e+02	1.465969e-04	5.006800e-06	1.406000e-02	0.04	0.02

MILC Profile on 32 Nodes

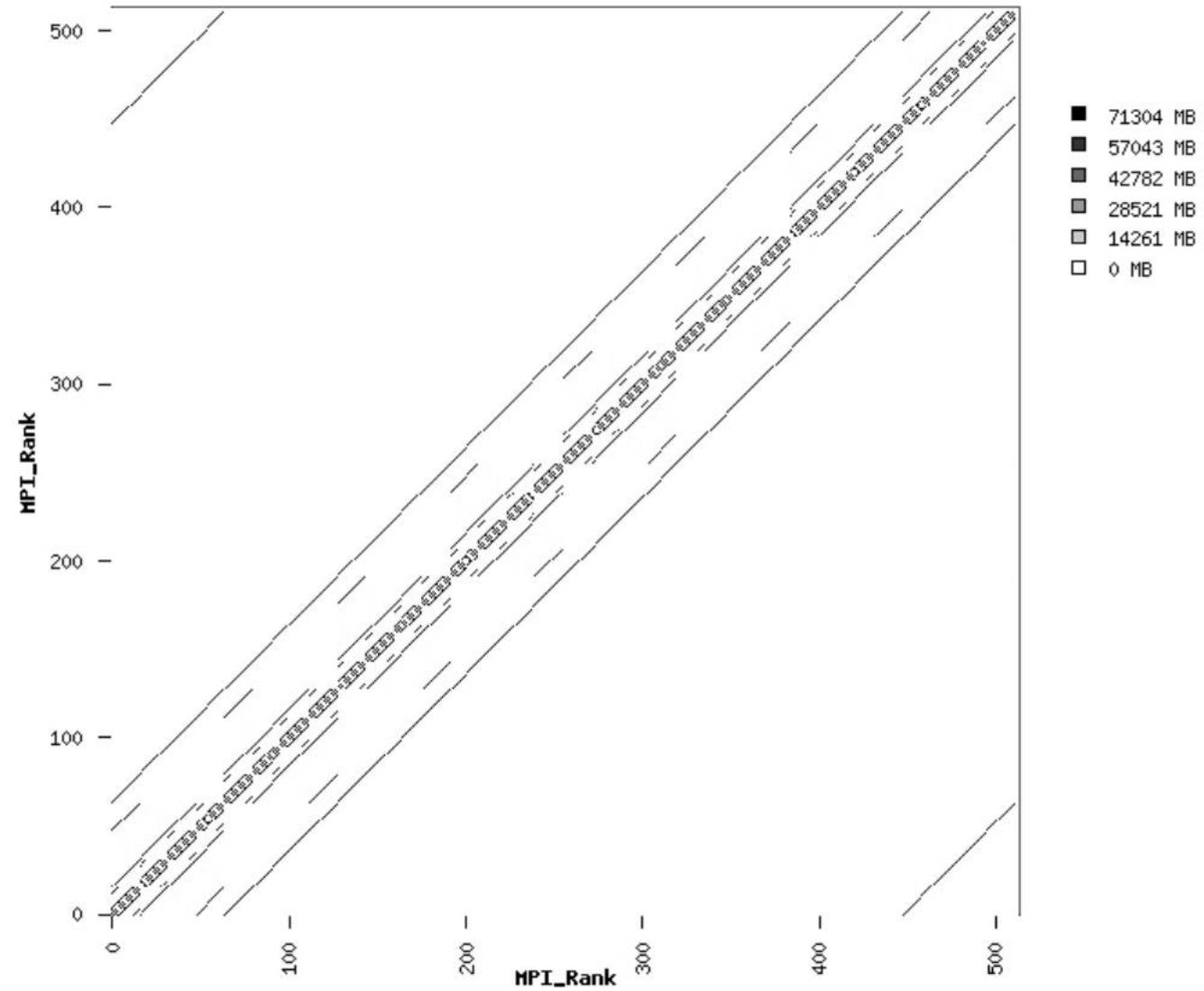
- 25% imbalance



MILC Profile on 32 Nodes Helios Cluster

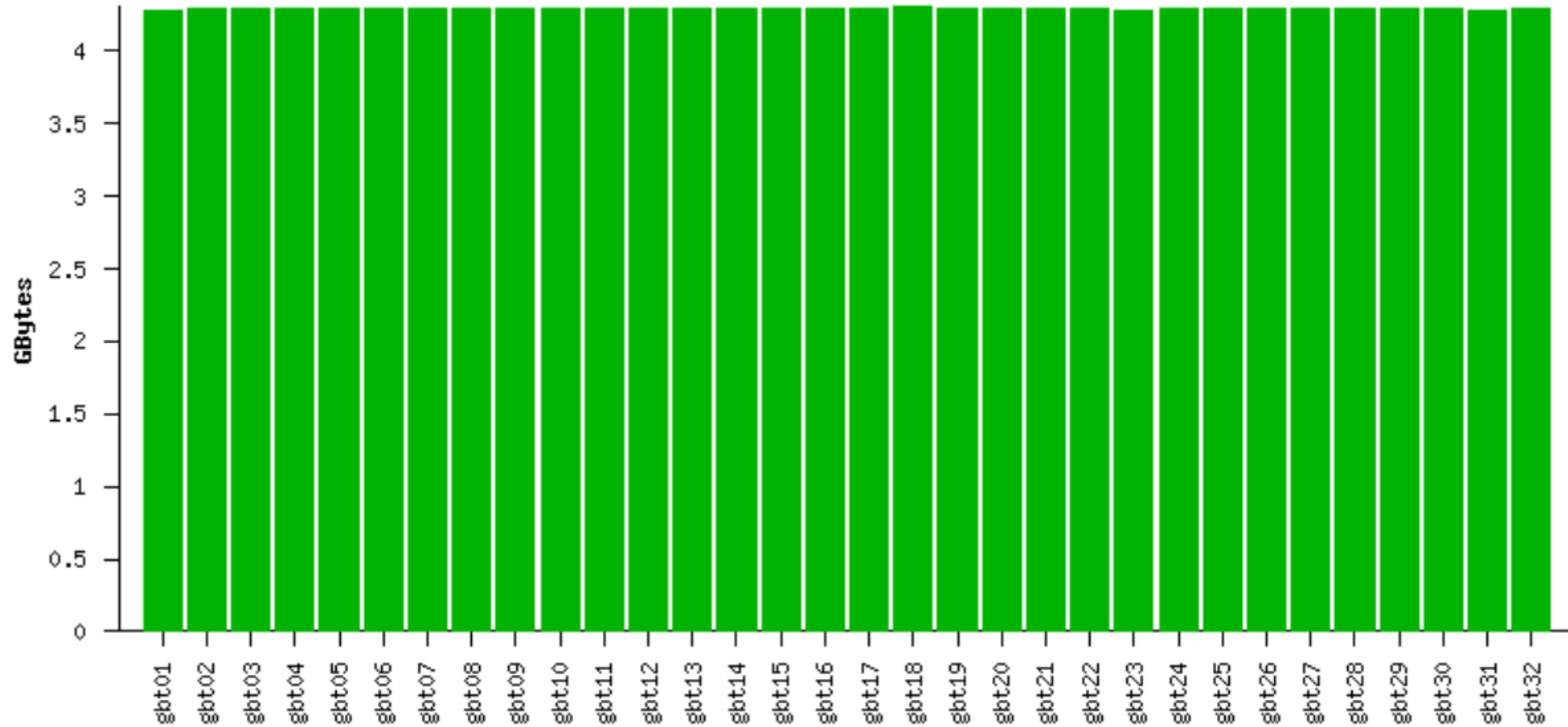


- **4D Communication Matrix**



MILC Profile on 32 Nodes Helios Cluster

- **Memory footprint**



- **Super linear scaling was achieved**
 - 1.55X 16 to 32 nodes on 8 OpenMP threads
 - 1.36X 8 to 16 nodes on 4 OpenMP threads
- **4 OpenMP threads per rank generated best performance**
 - 27% better performance achieved when using 4 OpenMP threads comparing to 8 OpenMP threads on 16 nodes.
 - 10% better performance achieved when using 4 OpenMP threads comparing to 8 OpenMP threads on 8 nodes.
- **MILC Profile**
 - Async P2P communication with 8 byte MPI Allreduce collective
 - 4D communication matrix

Thank You

