More Results

1) Data Scaling on Cori: Table I shows the profiles of LAMMPS simulation and analysis (1000 time steps) on 64 nodes (32 ranks per node, 2 threads per rank) of Cori for different system sizes (column 1). In each row (columns 2 – 4), the data for default and NUMA-aware mappings are shown in the first and second sub-rows, respectively. The total execution time for NUMA-aware mapping is lower than the default in all cases. The reason for this can be explained from columns 3 – 4. The L3 cache misses and the MPI_Send times are higher for the default mapping. MPI_Send is used in both simulation and analysis code paths and is one of the most time consuming and called functions. Overall, the NUMA-aware mapping results in up to 10% lower execution times than the default mapping.

TABLE I: Profiles of 1000-step simulation and analysis using default and NUMA-aware mappings on 2048 ranks of Cori.

#Atoms	Wallclock (s)	L3 miss	MPI_Send (s)
12M	111.1	2.2×10^8	53.0
	104.4	2.1×10^8	49.8
21M	184.9	4.1×10^8	87.5
	180.7	3.8×10^8	82.2
34M	314.4	6.9×10^8	170.4
	282.9	6.5×10^8	148.8
51M	439.5	12.7×10^8	225.3
	412.3	11.1×10^8	205.5