PRODUCT BRIEF

SUPERMICRO AND AMD TOGETHER INCREASE PERFORMANCE OF HPC WORKLOADS USING THE AMD INSTINCT™ MI355X ACCELERATOR

Table of Contents

Executive Summary

High-performance computing (HPC) workloads increasingly rely on accelerators such as the AMD Instinct MI355X to achieve breakthrough performance, scalability, and energy efficiency. These accelerators are particularly beneficial in compute-intensive applications where massive parallelism enables faster iteration and higher-resolution modeling. Applications in quantum chemistry and materials science workloads, such as Chroma, GROMACS, and NAMD, achieve significant performance gains through mixed-precision matrix operations and optimized BLAS libraries. Overall, accelerators like the AMD Instinct MI355X transform traditional CPU-bound HPC clusters into GPU-accelerated supercomputing platforms, delivering order-of-magnitude improvements in performance per watt and faster time-to-solution across scientific and industrial domains.

Key Solution Benchmarking Components and Setup

Hardware – AMD Instinct GPU (AS -4126GS-NMR-LCC - Tray, Front & Real)

Figure 1 - Supermicro Liquid Cooled AS -4126GS-TNMR

Firmware Version of Supermicro Server

Model Number	CPLD version	BIOS Version	BMC Version
AS -4126GS-NMR-LCC	F2.65.0E	BIOS Date: 09/12/2025 Ver 1.7	01.03.03.01

Hardware Specifications of the Server

CPU Model	CPU Count	MEM (GB)	DIMM PN	DIMM Count
AMD EPYC 9575F 64-Core Processor	2	3072	MTC40F2047S1RC64BB1	24

Detailed Hardware Version Per Node

Item Name	Model Name	Qty
System	AS -4126GS-NMR-LCC	1
Motherboard	H14DSG-OD	1
Processor	AMD EPYC 9575F 64-Core Processor	2
Memory	MTC40F2047S1RC64BB1	24

GPU	AMD Instinct MI355X	8
Disk	Micron_7450_MTFDKCC3T8TFR	8
NIC cards	POLLARA-1Q400P	8
NIC cards	MCX75310AAC-NEAT	1
Power Supply	PWS-6K61G-2R	4
Fans	NA	4

HPC Benchmark Performance Test

The following HPC benchmark results summary was generated using Chroma, GROMACS, and NAMD. Reference scripts and configurations are published in AMD InfinityHub: https://github.com/amd/InfinityHub-CI.

HPC Benchmark Summary — AMD Instinct™ MI355X

Benchmarking	Workload / Problem	Primary Metric	Result	What It Indicates
Chroma	QUDA BICGSTAB Clover Solver (24×24×24×128 lattice)	Time to Solution	**100.7 – 116.8 s**	Fast lattice QCD solve dominated by GPU compute (≈99.9% compute time)
		Sustained Performance	**11.1 – 12.9 TFLOPS**	Efficient HBM3 utilization for memory-bound LQCD kernels
		GPU Memory Usage	**~1.24 GB per GPU**	Excellent memory efficiency for large lattice volumes
GROMACS	ADH-Dodec	Simulation Throughput	**309.8 ns/day**	High-throughput MD for small/medium biomolecular systems
		Wall Time	**5.58 s**	Excellent strong-scaling efficiency
GROMACS	Cellulose-NVE	Simulation Throughput	**125.9 ns/day**	Sustained performance on a more communication-intensive system
		Wall Time	**13.73 s**	Performance is impacted mainly by PME imbalance, not GPU compute
GROMACS	STMV Virus	Simulation Throughput	**63.4 ns/day**	Real-world large biomolecular workload with heavy PME cost
		Wall Time	**27.24 s**	Demonstrates MI355X scalability under challenging MD conditions
NAMD	Large-scale MD (1M steps)	Steady-state Throughput	**37.8 ns/day**	Consistent long-run performance with stable timestep execution

Benchmarking	Workload / Problem	Primary Metric	Result	What It Indicates
		Time per Step	**~0.00457 s/step**	Low jitter and strong GPU acceleration efficiency
		Total Runtime	**~4575 s (1.27 hrs)**	Sustained performance over an extended production run

Significance of the Results

These benchmark results demonstrate a double-digit performance uplift over CPU-only systems where the workloads can take advantage of GPU's compute acceleration, enabling workloads that previously required months of simulation time to be completed in days with GPU acceleration.

For additional context, the following table lists a few estimated metrics based on published results from OpenBencharking.org:

Benchmark	Workload / Problem	Primary Metric	MI355X Result	High End CPU Result (Est.)	MI355X Uplift vs CPU (Est)
Chroma	QUDA BICGSTAB Clover Solver	Time to solution	**100.7 – 116.8 s**	~1,000 – 1,600 s	**~9× – 14× faster**
GROMACS	ADH-Dodec	Simulation throughput	**309.8 ns/day**	~30 – 45 ns/day (GROMACS Benchmark - OpenBenchmarking.org)	**~7× – 10×**
GROMACS	Cellulose-NVE	Simulation throughput	**125.9 ns/day**	~12 – 18 ns/day	**~7× – 10×**
GROMACS	STMV Virus	Simulation throughput	**63.4 ns/day**	~6 – 9 ns/day	**~7× – 10×**
NAMD	Large-scale MD (1M steps)	Simulation throughput	**37.8 ns/day**	~4 – 8 ns/day (NAMD Benchmark - OpenBenchmarking.org)	**~6× – 9×**

Summary

The AMD Instinct MI355X is a compelling solution for High-Performance Computing (HPC) applications due to its massive memory capacity, high double-precision (FP64) performance, and the open-source ROCm software ecosystem. These features enable it to handle extensive, complex scientific simulations and data-intensive tasks efficiently and at scale.

Appendix:

Chroma Benchmarking:

Metric(s)	Results
Time to Solution (seconds per Dirac solve)	Lattice initialized: problem size = 24 24 24 128 layout size = 12 24 24 128 logical machine size = 1 1 1 8 subgrid size = 24 24 24 16 total number of nodes = 8 total volume = 1769472 subgrid volume = 221184 **QUDA_BICGSTAB_CLOVER_SOLVER: time=116.838694 s** **Performance=11094.2525735327 GFLOPS Total Time (incl. load gauge)=116.838847 s** QUDA_BICGSTAB_CLOVER_SOLVER: solution 0 : 10000 iterations. Relative Rsd = 1.820395e-07 QUDA_BICGSTAB_CLOVER_SOLVER: solution 1 : 0 iterations. Relative Rsd = 1.670204e-07 QUDA_BICGSTAB_CLOVER_SOLVER: solution 2 : 0 iterations. Relative Rsd = 1.645771e-07 QUDA_BICGSTAB_CLOVER_SOLVER: solution 3 : 0 iterations. Relative Rsd = 1.629515e-07 QUDA_BICGSTAB_CLOVER_SOLVER: solution 4 : 0 iterations. Relative Rsd = 1.752275e-07 QUDA_BICGSTAB_CLOVER_SOLVER: solution 5 : 0 iterations. Relative Rsd = 1.718917e-07 QUDA_BICGSTAB_CLOVER_SOLVER: solution 6 : 0 iterations. Relative Rsd = 1.87469e-07 QUDA_BICGSTAB_CLOVER_SOLVER: solution 7 : 0 iterations. Relative Rsd = 1.490381e-07 QUDA_BICGSTAB_CLOVER_SOLVER: solution 8 : 0 iterations. Relative Rsd = 1.809485e-07 QUDA_BICGSTAB_CLOVER_SOLVER: solution 9 : 0 iterations. Relative Rsd = 1.814808e-07 QUDA_BICGSTAB_CLOVER_SOLVER: solution 10 : 0 iterations. Relative Rsd = 1.619684e-07 **QUDA_BICGSTAB_CLOVER_SOLVER: solution 11 : 0 iterations. Relative Rsd = 1.559617e-07** initQuda-endQuda Total time = 137.838 secs QUDA Total time = 117.183 secs download = 0.158 secs (0.135%), with 14 calls at 1.129e+04 us per call upload = 0.006 secs (0.005%), with 12 calls at 5.185e+02 us per call init = 0.139 secs (0.119%), with 306 calls at 4.540e+02 us per call preamble = 0.025 secs (0.022%), with 26 calls at 9.806e+02 us per call compute = 116.787 secs (99.662%), with 28 calls at 4.171e+06 us per call epilogue = 0.000 secs (0.000%), with 2 calls at 2.800e+01 us per call free = 0.000 secs (0.000%), with 136 calls at 3.750e-01 us per call total accounted = 117.115 secs (99.942%) total missing = 0.068 secs (0.058%) Device memory used = 1239.4 MiB Pinned device memory used = 0.0 MiB Managed memory used = 0.0 MiB Page-locked host memory used = 30.6 MiB Total host memory used >= 30.7 MiB initQuda Total time = 0.129 secs init = 0.129 secs (99.997%), with 2 calls at 6.455e+04 us per call

```
    total accounted     =    0.129 secs ( 99.997%)
    total missing       =    0.000 secs (  0.003%)

     loadGaugeQuda Total time =    0.137 secs
          download    =    0.136 secs ( 99.351%),     with      1 calls at 1.361e+05 us per call
            init      =    0.001 secs (  0.499%),     with      7 calls at 9.757e+01 us per call
         compute      =    0.000 secs (  0.093%),     with      3 calls at 4.233e+01 us per call
    total accounted     =    0.137 secs ( 99.942%)
    total missing       =    0.000 secs (  0.058%)

     loadCloverQuda Total time =    0.033 secs
          download    =    0.014 secs ( 40.451%),     with      1 calls at 1.353e+04 us per call
            init      =    0.000 secs (  0.879%),     with      4 calls at 7.350e+01 us per call
    total accounted     =    0.014 secs ( 41.329%)
    total missing       =    0.020 secs ( 58.671%)

   invertMultiSrcQuda Total time =   116.838 secs
          download    =    0.008 secs (  0.007%),     with     12 calls at 7.011e+02 us per call
           upload     =    0.006 secs (  0.005%),     with     12 calls at 5.185e+02 us per call
            init      =    0.009 secs (  0.008%),     with    293 calls at 3.019e+01 us per call
        preamble      =    0.025 secs (  0.022%),     with     26 calls at 9.806e+02 us per call
         compute      =  116.786 secs ( 99.956%),     with     25 calls at 4.671e+06 us per call
        epilogue      =    0.000 secs (  0.000%),     with      2 calls at 2.800e+01 us per call
            free      =    0.000 secs (  0.000%),     with    112 calls at 3.125e-01 us per call
    total accounted     =  116.836 secs ( 99.998%)
    total missing       =    0.002 secs (  0.002%)

         endQuda Total time =    0.046 secs
            free      =    0.000 secs (  0.024%),     with     24 calls at 4.583e-01 us per call
    total accounted     =    0.000 secs (  0.024%)
    total missing       =    0.046 secs ( 99.976%)

    initQuda-endQuda Total time =   137.838 secs

           QUDA Total time =   117.183 secs
          download    =    0.158 secs (  0.135%),     with     14 calls at 1.129e+04 us per call
           upload     =    0.006 secs (  0.005%),     with     12 calls at 5.185e+02 us per call
            init      =    0.139 secs (  0.119%),     with    306 calls at 4.540e+02 us per call
        preamble      =    0.025 secs (  0.022%),     with     26 calls at 9.806e+02 us per call
         compute      =  116.787 secs ( 99.662%),     with     28 calls at 4.171e+06 us per call
        epilogue      =    0.000 secs (  0.000%),     with      2 calls at 2.800e+01 us per call
            free      =    0.000 secs (  0.000%),     with    136 calls at 3.750e-01 us per call
    total accounted     =  117.115 secs ( 99.942%)
    total missing       =    0.068 secs (  0.058%)

Device memory used = 1239.4 MiB
Pinned device memory used = 0.0 MiB
Managed memory used = 0.0 MiB
Page-locked host memory used = 30.6 MiB
Total host memory used >= 30.7 MiB

Lattice initialized:
  problem size = 24 24 24 128
```

	layout size = 12 24 24 128 logical machine size = 1 1 1 8 subgrid size = 24 24 24 16 total number of nodes = 8 total volume = 1769472 subgrid volume = 221184 **QUDA_BICGSTAB_CLOVER_SOLVER: time=100.730744 s** **Performance=12868.3451558513 GFLOPS Total Time (incl. load gauge)=100.730801 s** QUDA_BICGSTAB_CLOVER_SOLVER: solution 0 : 10000 iterations. Relative Rsd = 1.820395e-07 QUDA_BICGSTAB_CLOVER_SOLVER: solution 1 : 0 iterations. Relative Rsd = 1.670204e-07 QUDA_BICGSTAB_CLOVER_SOLVER: solution 2 : 0 iterations. Relative Rsd = 1.645771e-07 QUDA_BICGSTAB_CLOVER_SOLVER: solution 3 : 0 iterations. Relative Rsd = 1.629515e-07 QUDA_BICGSTAB_CLOVER_SOLVER: solution 4 : 0 iterations. Relative Rsd = 1.752275e-07 QUDA_BICGSTAB_CLOVER_SOLVER: solution 5 : 0 iterations. Relative Rsd = 1.718917e-07 QUDA_BICGSTAB_CLOVER_SOLVER: solution 6 : 0 iterations. Relative Rsd = 1.87469e-07 QUDA_BICGSTAB_CLOVER_SOLVER: solution 7 : 0 iterations. Relative Rsd = 1.490381e-07 QUDA_BICGSTAB_CLOVER_SOLVER: solution 8 : 0 iterations. Relative Rsd = 1.809485e-07 QUDA_BICGSTAB_CLOVER_SOLVER: solution 9 : 0 iterations. Relative Rsd = 1.814808e-07 QUDA_BICGSTAB_CLOVER_SOLVER: solution 10 : 0 iterations. Relative Rsd = 1.619684e-07 QUDA_BICGSTAB_CLOVER_SOLVER: solution 11 : 0 iterations. Relative Rsd = 1.559617e-07

GROMACS ADH-Dodec Benchmarking

Metric(s)	Results
Time to Solution (seconds per Dirac solve)	Dynamic load balancing report: DLB was off during the run due to a low measured imbalance. Average load imbalance: 5.6%. The balanceable part of the MD step is 55%, load imbalance is computed from this. Part of the total run time spent waiting due to load imbalance: 3.1%. Average PME mesh/force load: 1.065 Part of the total run time spent waiting due to PP/PME imbalance: 2.3 % Core t (s) Wall t (s) (%) **Time: 352.590 5.579 6320.4** **(ns/day) (hour/ns)**

	Performance: 309.784 0.077

GROMACS Cellulose-NVE Benchmarking

Metric(s)	Results
Time to Solution (seconds per Dirac solve)	Dynamic load balancing report: DLB was off during the run due to low measured imbalance. Average load imbalance: 12.2%. The balanceable part of the MD step is 45%, load imbalance is computed from this. Part of the total run time spent waiting due to load imbalance: 5.5%. Average PME mesh/force load: 1.761 Part of the total run time spent waiting due to PP/PME imbalance: 20.1 % NOTE: 5.5 % of the available CPU time was lost due to load imbalance in the domain decomposition. Dynamic load balancing was automatically disabled, but it might be beneficial to manually turn it on (option -dlb yes) You can also consider manually changing the decomposition (option -dd); e.g. by using fewer domains along the box dimension in which there is considerable inhomogeneity in the simulated system. NOTE: 20.1 % performance was lost because the PME ranks had more work to do than the PP ranks. You might want to increase the number of PME ranks or increase the cut-off and the grid spacing.

	Core t (s) Wall t (s) (%) **Time: 866.930 13.732 6313.3** **(ns/day) (hour/ns)** Performance: 125.852 0.191

GROMACS STMV Benchmarking

Metric(s)	Results
Time to Solution (seconds per Dirac solve)	Dynamic load balancing report: DLB was off during the run due to low measured imbalance. Average load imbalance: 10.1%. The balanceable part of the MD step is 40%, load imbalance is computed from this. Part of the total run time spent waiting due to load imbalance: 4.0%. Average PME mesh/force load: 2.530 Part of the total run time spent waiting due to PP/PME imbalance: 33.6 % NOTE: 33.6 % performance was lost because the PME ranks had more work to do than the PP ranks. You might want to increase the number of PME ranks or increase the cut-off and the grid spacing. **Core t (s) Wall t (s) (%)** **Time: 1727.529 27.242 6341.4**

	(ns/day) (hour/ns) **Performance: 63.437 0.378** GROMACS reminds you: "Step On the Brakes" (2 Unlimited)

NAMD Benchmarking

Metric(s)	Results
Time to Solution (seconds per Dirac solve)	TIMING: 500 CPU: 3.68424, 0.0049911/step Wall: 3.7041, 0.00578063/step, 29.8929 ns/days, 1.60493 hours remaining, 0.000000 MB of memory in use. PERFORMANCE: 500 averaging 29.8929 ns/day, 0.00578063 sec/step with standard deviation 0 TIMING: 1000 CPU: 6.0327, 0.00469694/step Wall: 6.19253, 0.00497685/step, 34.7207 ns/days, 1.38108 hours remaining, 0.000000 MB of memory in use. PERFORMANCE: 1000 averaging 32.1265 ns/day, 0.00537874 sec/step with standard deviation 0.000568358 TIMING: 1500 CPU: 8.29485, 0.00452429/step Wall: 8.54739, 0.00470972/step, 36.6901 ns/days, 1.30629 hours remaining, 0.000000 MB of memory in use. PERFORMANCE: 1500 averaging 33.5161 ns/day, 0.00515574 sec/step with standard deviation 0.000557416 TIMING: 2000 CPU: 10.568, 0.00454637/step Wall: 10.9348, 0.0047749/step, 36.1892 ns/days, 1.32371 hours remaining, 0.000000 MB of memory in use. PERFORMANCE: 2000 averaging 34.1466 ns/day, 0.00506053 sec/step with standard deviation 0.000493356 TIMING: 2500 CPU: 12.7787, 0.0044214/step Wall: 13.1805, 0.00449136/step, 38.4738 ns/days, 1.24448 hours remaining, 0.000000 MB of memory in use. PERFORMANCE: 2500 averaging 34.9324 ns/day, 0.0049467 sec/step with standard deviation 0.000497333 TIMING: 3000 CPU: 15.0566, 0.0045558/step Wall: 15.4713, 0.00458154/step, 37.7166 ns/days, 1.26883 hours remaining, 0.000000 MB of memory in use.

	PERFORMANCE: 3000 averaging 35.3675 ns/day, 0.00488584 sec/step with standard deviation 0.000469143 TIMING: 3500 CPU: 17.2926, 0.00447184/step Wall: 17.7068, 0.00447104/step, 38.6487 ns/days, 1.23761 hours remaining, 0.000000 MB of memory in use. PERFORMANCE: 3500 averaging 35.8018 ns/day, 0.00482658 sec/step with standard deviation 0.000456061 TIMING: 4000 CPU: 19.6098, 0.00463455/step Wall: 20.0466, 0.00467966/step, 36.9257 ns/days, 1.29471 hours remaining, 0.000000 MB of memory in use. PERFORMANCE: 4000 averaging 35.9385 ns/day, 0.00480821 sec/step with standard deviation 0.000425414 TIMING: 4500 CPU: 21.8623, 0.00450488/step Wall: 22.3149, 0.00453657/step, 38.0905 ns/days, 1.25449 hours remaining, 0.000000 MB of memory in use. PERFORMANCE: 4500 averaging 36.1655 ns/day, 0.00477803 sec/step with standard deviation 0.00040811 TIMING: 5000 CPU: 24.1253, 0.00452615/step Wall: 24.6064, 0.00458296/step, 37.7049 ns/days, 1.26668 hours remaining, 0.000000 MB of memory in use. PERFORMANCE: 5000 averaging 36.3138 ns/day, 0.00475852 sec/step with standard deviation 0.000389683 TIMING: 5500 CPU: 26.3827, 0.00451464/step Wall: 26.8787, 0.00454453/step, 38.0238 ns/days, 1.25543 hours remaining, 0.000000 MB of memory in use. PERFORMANCE: 5500 averaging 36.4628 ns/day, 0.00473907 sec/step with standard deviation 0.000375274 TIMING: 6000 CPU: 28.6379, 0.00451051/step Wall: 29.1345, 0.0045116/step, 38.3012 ns/days, 1.2457 hours remaining, 0.000000 MB of memory in use. PERFORMANCE: 6000 averaging 36.6093 ns/day, 0.00472012 sec/step with standard deviation 0.000363785 TIMING: 6500 CPU: 30.8953, 0.00451478/step Wall: 31.4009, 0.0045328/step, 38.1221 ns/days, 1.25093 hours remaining, 0.000000 MB of memory in use. PERFORMANCE: 6500 averaging 36.7214 ns/day, 0.00470571 sec/step with standard deviation 0.000352151

	TIMING: 7000 CPU: 33.1689, 0.00454709/step Wall: 33.6751, 0.00454839/step, 37.9914 ns/days, 1.2546 hours remaining, 0.000000 MB of memory in use. PERFORMANCE: 7000 averaging 36.8093 ns/day, 0.00469447 sec/step with standard deviation 0.000340938 TIMING: 7500 CPU: 35.386, 0.00443439/step Wall: 35.9011, 0.00445212/step, 38.8129 ns/days, 1.22743 hours remaining, 0.000000 MB of memory in use. PERFORMANCE: 7500 averaging 36.9364 ns/day, 0.00467831 sec/step with standard deviation 0.000334442 TIMING: 8000 CPU: 37.6216, 0.00447117/step Wall: 38.1376, 0.00447292/step, 38.6325 ns/days, 1.23254 hours remaining, 0.000000 MB of memory in use. PERFORMANCE: 8000 averaging 37.038 ns/day, 0.00466548 sec/step with standard deviation 0.000327156 TIMING: 8500 CPU: 39.8555, 0.0044678/step Wall: 40.3805, 0.00448584/step, 38.5212 ns/days, 1.23547 hours remaining, 0.000000 MB of memory in use. PERFORMANCE: 8500 averaging 37.1221 ns/day, 0.00465491 sec/step with standard deviation 0.00031975 TIMING: 9000 CPU: 42.0972, 0.00448332/step Wall: 42.6334, 0.0045057/step, 38.3514 ns/days, 1.24032 hours remaining, 0.000000 MB of memory in use. PERFORMANCE: 9000 averaging 37.1883 ns/day, 0.00464662 sec/step with standard deviation 0.00031219 TIMING: 9500 CPU: 44.346, 0.00449772/step Wall: 44.8856, 0.00450448/step, 38.3618 ns/days, 1.23936 hours remaining, 0.000000 MB of memory in use. PERFORMANCE: 9500 averaging 37.2483 ns/day, 0.00463914 sec/step with standard deviation 0.000305142 TIMING: 10000 CPU: 46.5922, 0.00449227/step Wall: 47.1325, 0.00449388/step, 38.4523 ns/days, 1.23582 hours remaining, 0.000000 MB of memory in use. PERFORMANCE: 10000 averaging 37.3067 ns/day, 0.00463188 sec/step with standard deviation 0.000298774 TIMING: 10500 CPU: 48.8778, 0.00457128/step Wall: 49.4178, 0.00457054/step, 37.8074 ns/days, 1.25626 hours remaining, 0.000000 MB of memory in use.

PERFORMANCE: 10500 averaging 37.3302 ns/day, 0.00462896 sec/step with standard deviation 0.000291516

TIMING: 11000 CPU: 51.1127, 0.00446977/step Wall: 51.6534, 0.00447109/step, 38.6483 ns/days, 1.22831 hours remaining, 0.000000 MB of memory in use.

PERFORMANCE: 11000 averaging 37.3882 ns/day, 0.00462178 sec/step with standard deviation 0.000286475

TIMING: 11500 CPU: 53.3346, 0.00444388/step Wall: 53.8752, 0.00444375/step, 38.886 ns/days, 1.22018 hours remaining, 0.000000 MB of memory in use.

PERFORMANCE: 11500 averaging 37.4509 ns/day, 0.00461404 sec/step with standard deviation 0.000282339

TIMING: 12000 CPU: 55.5846, 0.0044999/step Wall: 56.1252, 0.00449992/step, 38.4007 ns/days, 1.23498 hours remaining, 0.000000 MB of memory in use.

PERFORMANCE: 12000 averaging 37.4896 ns/day, 0.00460928 sec/step with standard deviation 0.000277114

TIMING: 12500 CPU: 57.8421, 0.00451509/step Wall: 58.3845, 0.00451867/step, 38.2413 ns/days, 1.2395 hours remaining, 0.000000 MB of memory in use.

PERFORMANCE: 12500 averaging 37.5191 ns/day, 0.00460566 sec/step with standard deviation 0.000271884

TIMING: 13000 CPU: 60.0728, 0.0044614/step Wall: 60.6146, 0.00446018/step, 38.7428 ns/days, 1.22283 hours remaining, 0.000000 MB of memory in use.

PERFORMANCE: 13000 averaging 37.5647 ns/day, 0.00460006 sec/step with standard deviation 0.000267914

……..

TIMING: 1000000 CPU: 4568.79, 0.00457358/step Wall: 4569.49, 0.00457478/step, 37.7723 ns/days, 0 hours remaining, 0.000000 MB of memory in use.

PERFORMANCE: 1000000 averaging 37.8228 ns/day, 0.00456868 sec/step with standard deviation 4.42148e-05

ETITLE: TS BOND ANGLE DIHED IMPRP ELECT VDW BOUNDARY MISC KINETIC TOTAL TEMP POTENTIAL TOTALAVG TEMPAVG PRESSURE GPRESSURE VOLUME PRESSAVG GPRESSAVG

	ENERGY: 1000000 37450.7364 103598.7394 81350.5560 4991.3843 -4020282.4336 288953.9743 0.0000 0.0000 638151.2365 -2865785.8066 288.9875 -3503937.0432 -2865804.9533 288.9870 285.5127 320.9117 10194598.5131 300.7918 336.1243 WRITING EXTENDED SYSTEM TO OUTPUT FILE AT STEP 1000000 WRITING COORDINATES TO OUTPUT FILE AT STEP 1000000 The last position output (seq=-2) takes 0.008 seconds, 0.000 MB of memory in use WRITING VELOCITIES TO OUTPUT FILE AT STEP 1000000 The last velocity output (seq=-2) takes 0.008 seconds, 0.000 MB of memory in use ========== WallClock: 4574.600586 CPUTime: 4572.560547 Memory: 0.000000 MB

Industry References:

Metrics			
Kernel Performance (Dslash & Clover)	**Dslash Mflops**	**Achieved bandwidth**	• **If bandwidth is ≥70% of peak, the GPU is performing near optimal.** • **For MI355X (HBM3 ~3.9 TB/s peak), 1960 GB/s (~50%) is typical for LQCD unless fused kernels are used.**
MPI	**MPI: Halo exchange: bytes transferred**		**⊠ MPI wait < 20% = very good scaling**

<table>
<tr>
<td></td>
<td>MPI: Halo time per iteration =

MPI: Communication/computation overlap = %

MPI: Total MPI wait = %</td>
<td></td>
<td>⊠ MPI wait > 30% = NIC or topology bottleneck

⊠ Multi-GPU systems with NVLink/XGMI generally show much lower wait time</td>
</tr>
<tr>
<td>Solver</td>
<td>
<table>
<tr><th>Field</th><th>Meaning</th></tr>
<tr><td>iter</td><td>CG iteration count</td></tr>
<tr><td>r</td><td>Current solver residual (convergence)</td></tr>
<tr><td>time</td><td>Total elapsed time so far</td></tr>
</table>
</td>
<td></td>
<td>Solver loop

What matters

• Total iterations (1214) → should match across hardware if using same precision.

• Total time scaling with GPU performance reveals hardware differences.

• Residual decreasing steadily confirms solver stability.

HPC insight:
You compare time-to-solution, not GFLOPS.
LQCD is memory- and communication-bound—not compute bound.
<table>
<tr><th>Field</th><th>Meaning</th></tr>
<tr><td>iterations</td><td>How hard the linear system is.</td></tr>
<tr><td>solve time</td><td>This is THE key metric: smaller = better hardware efficiency.</td></tr>
<tr><td>Mflops</td><td>Not peak FLOPS; sustained FLOPS of the actual run.</td></tr>
<tr><td>bandwidth</td><td>Real achieved HBM throughput.</td></tr>
</table>
</td>
</tr>
</table>

			residual	Must meet convergence tolerance to validate correctness.

Learn More

Learn more about the Supermicro AMD Servers: https://www.supermicro.com/en/products/aplus

SUPERMICRO

As a global leader in high performance, high efficiency server technology and innovation, we develop and provide end-to-end green computing solutions to the data center, cloud computing, enterprise IT, big data, HPC, and embedded markets. Our Building Block Solutions® approach allows us to provide a broad range of SKUs, and enables us to build and deliver application-optimized solutions based upon your requirements.

Learn more at www.supermicro.com

AMD

For more than 50 years AMD has driven innovation in high-performance computing, graphics and visualization technologies. Billions of people, leading Fortune 500 businesses and cutting-edge scientific research institutions around the world rely on AMD technology daily to improve how they live, work and play. AMD employees are focused on building leadership high-performance and adaptive products that push the boundaries of what is possible.

Learn more at www.amd.com