Systematic benchmark on various CPU platforms

Jen-Chang “James” CHEN (PhD)
Senior research fellow, NSCC Singapore
james@nscc.sg

National Supercomputing Centre

Outlines

- Meaning of benchmark
- Background study from the fields
- Proposed “systematic benchmark”
- Results of base benchmark on various CPU platforms
  - Intel Xeon E5, AMD EPYC, Power 8 & 9 and NEC VE
- Summary

Meaning of performance benchmark

- Provide the performance metrics for super-computer and HPC
  - Portable
  - Representative
  - Reproducible

- Propose the internal benchmark sets in NSCC Singapore for
  - System architecture
  - Tender evaluation
  - Industrial R&D

Source: https://blog.thunderquote.com/2016/12/01/procurement-benchmarks/

Background study-SPEC benchmarks

- **Standard Performance Evaluation Corporation**, www.spec.org
- **SPEC CPU® 2017**
  - SPECspeed 2017 Integer, SPECspeed 2017 Floating Point, SPECrate 2017 Integer, and SPECrate 2017 Floating Point
- March 1, 2018, SPEC offers High Performance Group (HPG) benchmark suites
  - **SPEC ACCEL Benchmark Suite**
    - Exercises the performance of the accelerator, host CPU, memory transfer between host and accelerator, support libraries and drivers, and compilers. (OpenCL 1.1, OpenACC 1.0, and OpenMP 4.5 APIs)
  - **SPEC MPI Benchmark Suites** (MPIM2007 and MPIL2007**)**
    - Evaluates the performance of MPI based program across a wide range of cluster and SMP hardware. (in compliance with MPI 2.1 standard)
  - **SPEC OMP2012 Benchmark Suite**
    - Designs for measuring performance using applications based on the OpenMP 3.1 standard for shared-memory parallel processing.

Background study-NAS Parallel benchmark

- NAS Parallel Benchmarks (NPB)
  - a small set of programs designed to help evaluate the performance of parallel supercomputers. The benchmarks are derived from computational fluid dynamics (CFD).

Version	Benchmarks Included	Problem Classes	Programming Models Used	Major Changes
NPB 3.3.1	IS, EP, CG, MG, FT, BT, BT-IO, SP, LU, UA, DC, DT	S,W,A,B,C,D,E	MPI, OpenMP, serial	added Class E
NPB 3.3.1-MZ	BT-MZ, SP-MZ, LU-MZ	S,W,A,B,C,D,E,F	MPI+OpenMP, OpenMP, serial	nested OpenMP version
GridNPB 3.1	ED, HC, VP, MB	S,W,A,B	Globus, Java, serial	added Globus version
NPB 3.0	IS, EP, CG, MG, FT, BT, SP, LU	S,W,A,B,C	OpenMP, HPF, Java	new programming paradigms
NPB 2.4.1	IS, EP, CG, MG, FT, BT, BT-IO, SP, LU	S,W,A,B,C,D	MPI	added BT-IO, Class D
NPB 2.3	IS, EP, CG, MG, FT, BT, SP, LU	S,W,A,B,C	MPI, serial	added CG, serial version

Notation:
IS - Integer Sort
EP - Embarrassingly Parallel
CG - Conjugate Gradient
MG - Multi-Grid on a sequence of meshes
FT - discrete 3D fast Fourier Transform
BT - Block Tri-diagonal solver
SP - Scalar Penta-diagonal solver
LU - Lower-Upper Gauss-Seidel solver

Source: https://www.nas.nasa.gov/publications/npb.html

Background study-CORAL-2 benchmark

- RFP for CORAL-2 on 9/April/2018
- Two tiers with four category benchmarks

Tier-1

Category	Programs	Features
Skeleton Benchmarks	CORAL MPI, Memory, ML/DL micro-benchmark, I/O suite, CLOMP, Pynamic, RAJA performance suite	MPI, Memory, FFT, different precisions of GEMM , DL algorithms, file I/O, thread and functionality test
Scalable Science Benchmarks	HACC, Nekbone, QMCPack, LAMMPS	Cosmology, CFD, Quantum physics, CMD
Throughput Benchmarks	AMG, Kripke, Quicksilver, Pennant	Access pattern, latency and bandwidth of memory and network latency
Data Science and DL Benchmarks	Big data analytic suite, Deep learning suite	Big data application and Deep learning algorithms

Tier-2

Category	Programs	Features
Scalable Science Benchmarks	ACME, VPIC	Climate and particle-in-cell simulation
Throughput Benchmarks	AMG, Laghos, LAMMPS	Multigrid solver, finite element and MD
Data Science and DL Benchmarks	Parallel integer sort, Havoq	BigSort of Int8 and graph analysis algorithms

Proposed systematic benchmarks utilized in NSCC

- Base benchmark
  - Represents the fundamental math algorithms for the basic computing unit.
  - HPL, HPCG, STREAM and HPCC
- Application benchmark
  - Consists of CPU and GPU categories.
  - The selected FIVE applications in each category are all based on the one-year statistics of usage.
  - CPU
    - CMD, Materials simulation and quantum physics and CFD
  - GPU
    - CMD and micro-magnetic simulation
    - No ML/DL in the top 5 list.

Base benchmark

Software	Description	URLs
HPL	Solves a (random) dense linear system in double precision (64 bits) arithmetic on distributed-memory computers.	http://www.netlib.org/benchmark/hpl/
HPCG	Performs a fixed number of symmetric Gauss-Seidel preconditioned conjugate gradient iterations using double precision (64 bit) floating point values	https://github.com/hpcg-benchmark/hpcg/
STREAM	Measures sustainable memory bandwidth (in MB/s) and the corresponding computation rate for simple vector kernels.	https://www.cs.virginia.edu/stream/
HPCC	A benchmark suite that measures a range memory access patterns. It contains HPL, DGEMM, STREAM, PTRANS, RandomAccess, FFT and Communication bandwidth & latency.	http://icl.cs.utk.edu/hpcc/

TOP 5 software applications in CPU node

Rank	Software	Total core-hours	Total # of jobs	Total # of users
1	Gromacs	59,459,581.6	31,203	128
2	VASP	25,986,373.5	242,363	143
3	LAMMPS	18,223,540.4	86,675	88
4	OpenFOAM	12,665,607.3	25,221	60
5	Quantum Espresso	7,984,759.7	28,379	82

Accounting period: 1/April/2017 – 31/March/2018

Benchmark cases for CPU node

Software	Benchmark cases	Application
Gromacs	http://www.prace-ri.eu/ueabs/#GROMACS	CMD
VASP	https://github.com/egplar/vasptest	Materials & quantum physics
LAMMPS	https://lammps.sandia.gov/bench.html	CMD
OpenFOAM	https://openfoamwiki.net/index.php/Benchmarks	CFD
Quantum Espresso	http://www.prace-ri.eu/ueabs/#Quantum_Espresso	Materials & quantum physics

TOP 5 software applications in GPU node

Rank	Software	Total GPU-hours	Total # of jobs	Total # of users
1	Gromacs	396,612.6	15,131	66
2	NAMD	92,017.1	645	6
3	AMBER	73,321.4	4,342	15
4	LAMMPS	67,147.9	2,164	33
5	Mumax3	36,931.2	7,457	10

Accounting period: 1/April/2017 – 31/March/2018

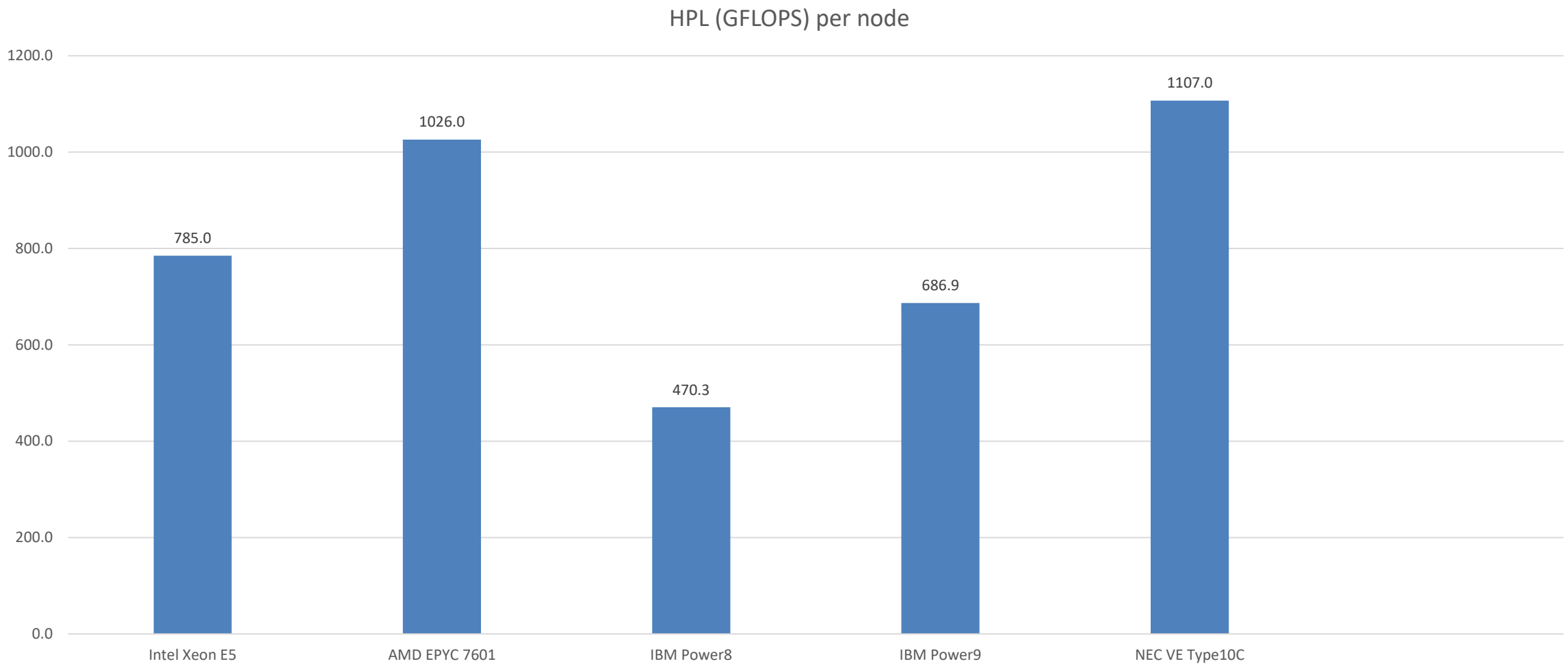
Benchmark cases for GPU node

Software	Benchmark cases	Application
Gromacs	http://www.prace-ri.eu/ueabs/#GROMACS	CMD
NAMD	https://www.ks.uiuc.edu/Research/namd/performance.html	CMD
AMBER	http://ambermd.org/gpus/benchmarks.htm	CMD
LAMMPS	https://lammps.sandia.gov/bench.html	CMD
Mumax3	https://github.com/mumax/mumax.github.io/blob/master/bench.svg	Micro-magnetic sim.

CPU specification

Model	CORES	Memory type	Base frequency (GHz)	Memory bandwidth (GB/s)	TDP (W)
Intel Xeon E5-2690 V3	12	DDR4-2133	2.60	68	135
AMD EPYC 7601	32	DDR4-2666	2.20	170	180
IBM Power 8	10	DDR4-1600	2.06	48	190
IBM Power 9	20	DDR4-2666	2.30	240	190
NEC VE Type 10C	8	HBM2	1.40	750	260

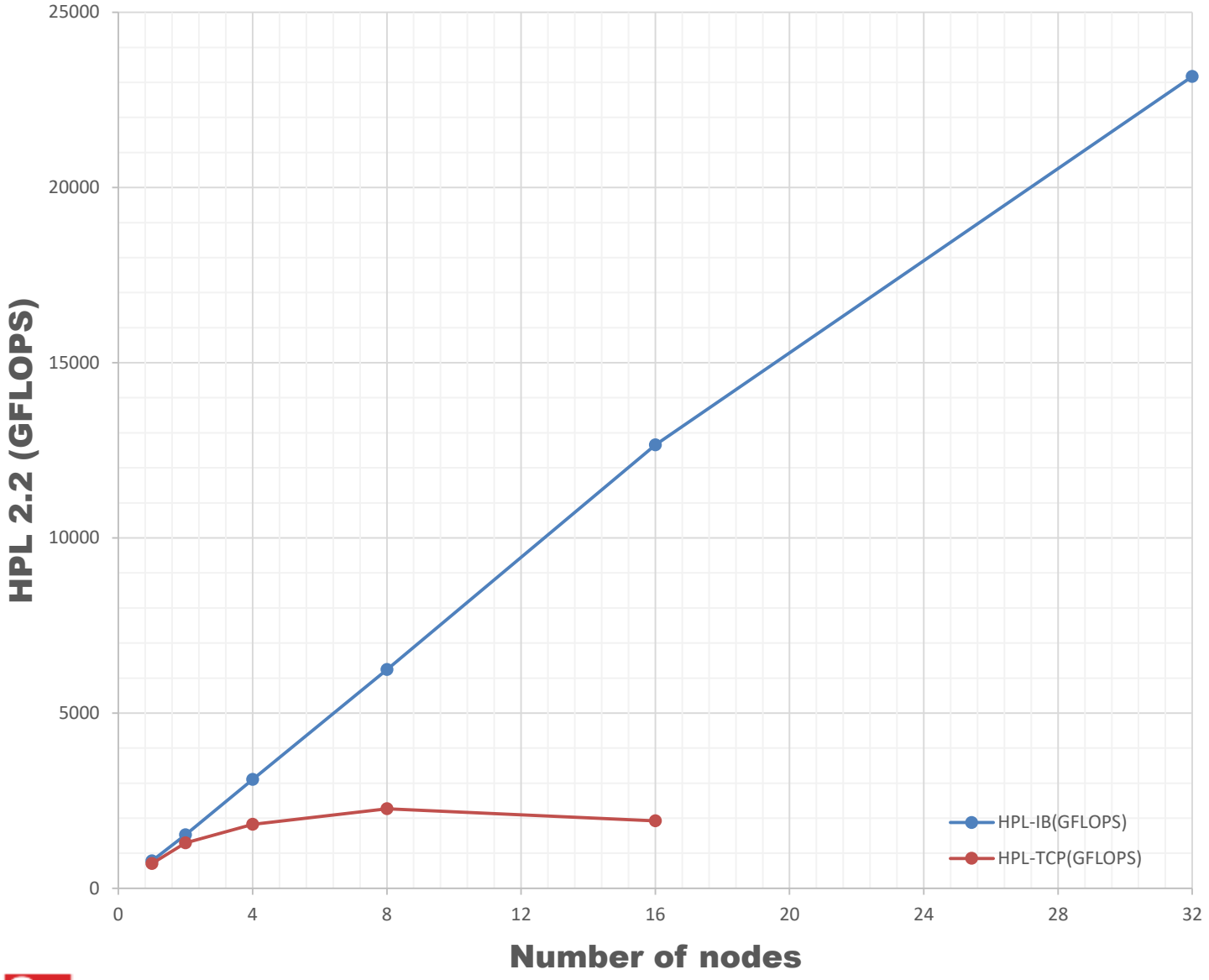
Base benchmark: HPL (CPU only)



Software: Intel (CentOS 6.9 + Intel compiler suite), AMD (CentOS 7 + PGI suite), IBM (RHEL 7 + IBM AT + IBM ESSL), NEC (CentOS 7 + NEC development suite)

National Supercomputing Centre

Multi-node HPL benchmark (Intel)



Platform/Compiler	HPL 2.2 R_{max} (GFLOPS)	Number of nodes
ASPIRE1/Intel 2016	785.0	1 node
ASPIRE1/Intel 2016	1,524.0	2 nodes
ASPIRE1/Intel 2016	3,107.0	4 nodes
ASPIRE1/Intel 2016	6,239.0	8 nodes
ASPIRE1/Intel 2016	12,650.0	16 nodes
ASPIRE1/Intel 2016	23,170.0	32 nodes

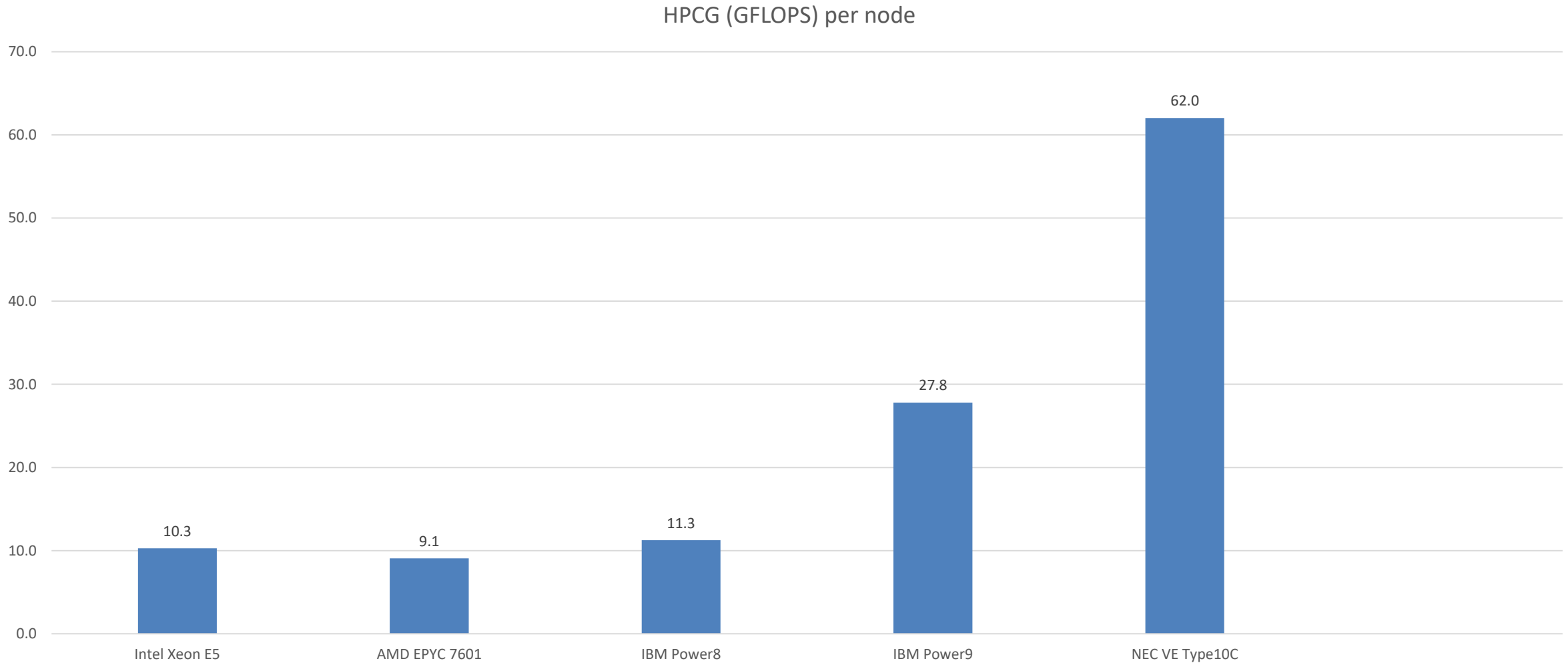
Note: Each node has 24 cores!

Source: http://www.netlib.org/benchmark/hpl/
Version: 2.2

Number of nodes	Speedup-IB (Pn/P1)	Speedup-TCP (Pn/P1)
1	1.0	1.0
2	1.9	1.8
4	4.0	2.6
8	7.9	3.2
16	16.0	2.7
32	29.5	-

Note: Binary, input and platform (ASPIRE1) are the same.
TCP setting: mpirun -env I_MPI_FABRICS tcp

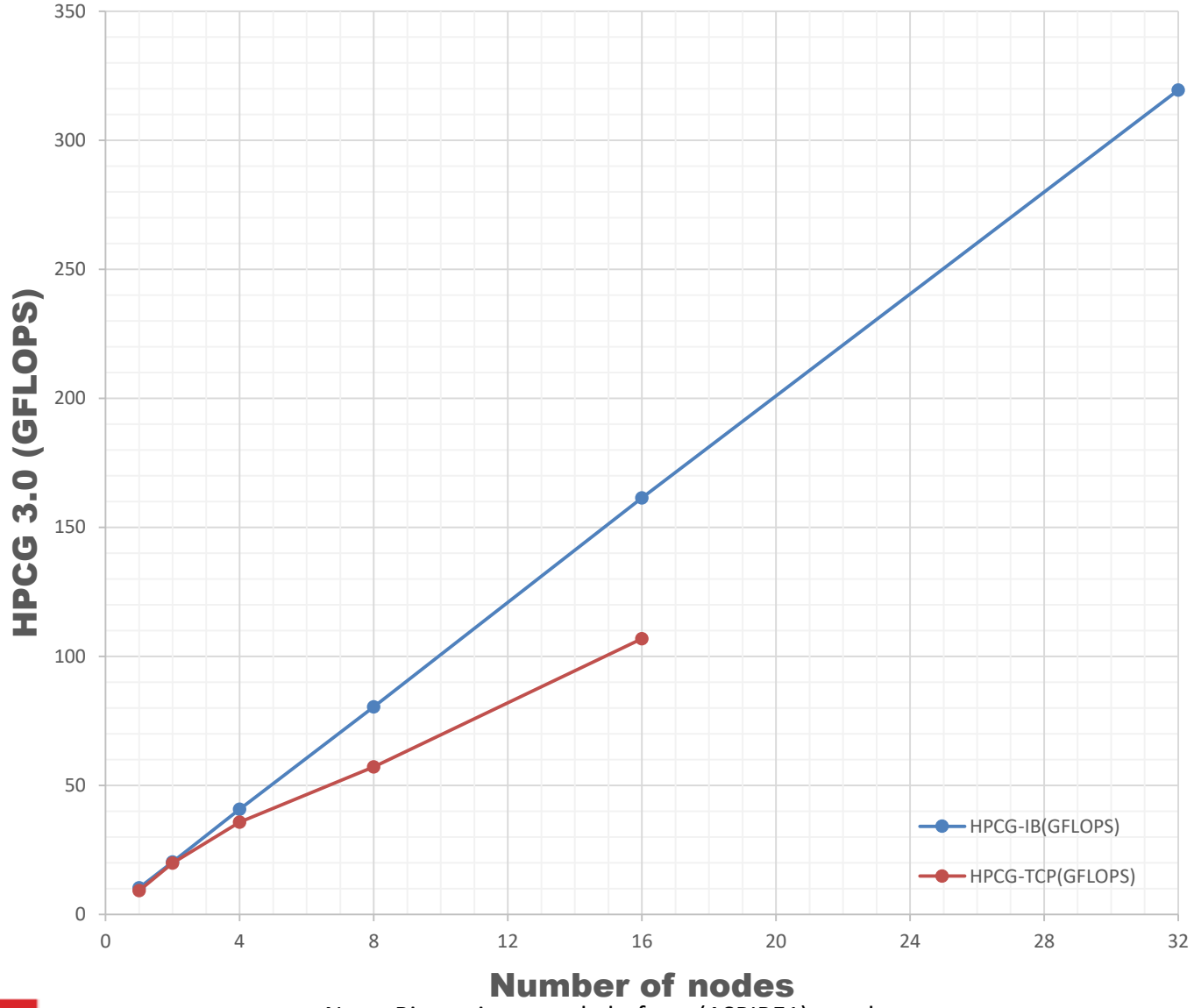
Base benchmark: HPCG (CPU only)



Software: Intel (CentOS 6.9 + Intel compiler suite), AMD (CentOS 7 + GNU 7.3), IBM (RHEL 7 + IBM AT), NEC (CentOS 7 + NEC development suite)

National Supercomputing Centre

Multi-node HPCG benchmark (Intel)



Platform/Compiler	HPCG3.0 (GFLOPS)	Number of nodes
ASPIRE1/Intel 2016	10.30	1 node
ASPIRE1/Intel 2016	20.30	2 node
ASPIRE1/Intel 2016	40.70	4 nodes
ASPIRE1/Intel 2016	80.40	8 nodes
ASPIRE1/Intel 2016	161.40	16 nodes
ASPIRE1/Intel 2016	319.40	32 nodes

Note: Each node has 24 cores.

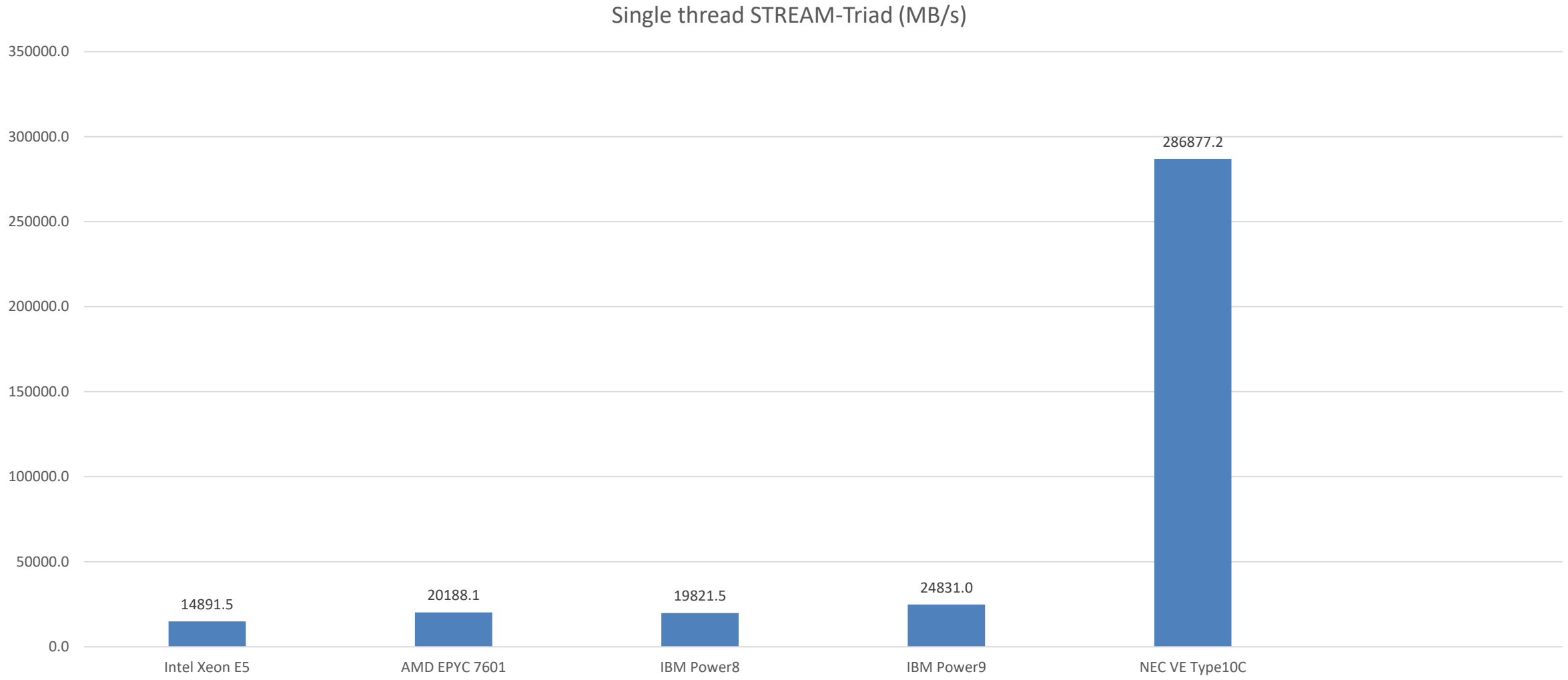
Source: https://github.com/hpcg-benchmark/hpcg/
Version: Revision: 3.0 Date: November 11, 2015
Input: Problem size is 128 x 128 x 128 (3,660 secs)

Number of nodes	Speedup-IB (Pn/P1)	Speedup-TCP (Pn/P1)
1	1.0	1.0
2	2.0	2.2
4	4.0	3.9
8	7.8	6.2
16	15.7	11.5
32	31.0	-

Note: Binary, input and platform (ASPIRE1) are the same.
TCP setting: mpirun -env I_MPI_FABRICS tcp

National Supercomputing Centre

Base benchmark: STREAM (CPU only)



Software: Intel (CentOS 6.9 + GNU), AMD (CentOS 7 + GNU 7.3), IBM (RHEL 7 + IBM AT), NEC (CentOS 7 + NEC GNU)

National Supercomputing Centre

Application (Haswell vs EPYC)-Quantum Espresso

Case: Au surface (ausurf.in)-pw.x

Source: http://www.prace-ri.eu/UEABS/Quantum_Espresso/QuantumEspresso_TestCaseA.tar.gz

Compiler	Wall Time (Secs)	Remarks
GNU	4320	OpenMPI 3.0.0, FFTW 3.3.6, OpenBLAS 0.2.20, ScaLAPACK 2.0.2 (MPI=64; OMP=1)
	2940	OpenMPI 3.0.0, FFTW 3.3.6, OpenBLAS 0.2.20, ScaLAPACK 2.0.2 (MPI=8; OMP=8)
	3060	OpenMPI 3.0.0, FFTW 3.3.6, OpenBLAS 0.2.20, ScaLAPACK 2.0.2 (MPI=8; OMP=8) mpirun --map-by numa --report-bindings
Intel	3473.35	OpenMPI 3.0.0, Intel MKL 18.0.1.163 (MPI=64; OMP=1)
PGI	4560	OpenMPI 3.0.0, FFTW 3.3.7, OpenBLAS 0.2.20, ScaLAPACK 2.0.2 (MPI=64; OMP=1)
ASPIRE1 Intel Xeon E5-2690 v3	2148.78	24 cores, quantum-espresso/5.4.0/parallel

Application (Haswell vs EPYC)-Quantum Espresso

Case: H_2O 256 molecules (cp.in)-cp.x

Source: http://www.prace-ri.eu/UEABS/Quantum_Espresso/QuantumEspresso_TestCaseA.tar.gz

Compiler	Wall Time (Secs)	Remarks Modify cp.in for the following settings nstep = 20, iprint = 10, isave = 10,
GNU	1920.00	OpenMPI 3.0.0, FFTW 3.3.6, OpenBLAS 0.2.20, ScaLAPACK 2.0.2 (MPI=8; OMP=8)
	1920.00	OpenMPI 3.0.0, FFTW 3.3.6, OpenBLAS 0.2.20, ScaLAPACK 2.0.2 (MPI=8; OMP=8) mpirun --map-by numa --report-bindings
	2615.05	OpenMPI 3.0.0, FFTW 3.3.6, OpenBLAS 0.2.20, ScaLAPACK 2.0.2 (MPI=64; OMP=1)
Intel	2022.66	OpenMPI 3.0.0, Intel MKL 18.0.1.163 (MPI=64; OMP=1)
PGI	2454.97	OpenMPI 3.0.0, FFTW 3.3.7, OpenBLAS 0.2.20, ScaLAPACK 2.0.2 (MPI=64; OMP=1)
ASPIRE1 Intel Xeon E5-2690 v3	851.96	24 cores, quantum-espresso/5.4.0/parallel

Application (Haswell vs EPYC)-Gromacs

Gromacs 2016.4

Case: lignocellulose

Source: http://www.prace-ri.eu/UEABS/GROMACS/1.2/GROMACS_TestCaseB.tar.gz

Compiler	Performance (ns/day)	Remarks
GNU	1.55	OpenMPI 3.0.0, FFTW 3.3.6, OpenBLAS 0.2.20, (MPI=64; OMP=1)
Intel	1.25	OpenMPI 3.0.0, Intel MKL 18.0.1.163 (MPI=64; OMP=1)
PGI	-	GROMACS recommend against PGI because the performance with C++ is very bad.
ASPIRE1 CPU Intel Xeon E5-2690 v3	0.89	24 cores, gromacs/5.1.2/gcc493/impi
ASPIRE1 CPU+GPU Intel Xeon E5-2690 v3 K40	0.79 1.32	gromacs/5.1.2/gcc493/impi_cuda MPI=2 OMP=12 GPU=1 (one node) MPI=2 OMP=12 GPU=2 (one node)

Other benchmarks utilized in NSCC

Software	Type	URLs
OFED Perftest	Network	https://community.mellanox.com/docs/DOC-2802 https://github.com/linux-rdma/perftest
IPER	Network	https://iperf.fr/
IOZONE	File I/O	http://www.iozone.org/
IOR	File I/O	https://github.com/LLNL/ior
MDTEST	File I/O	https://github.com/LLNL/mdtest
BBCP	Date transfer	http://www.slac.stanford.edu/~abh/bbcp/
Rsync	Date transfer	https://rsync.samba.org/
tf_cnn_benchmark	DL, CNN	https://github.com/tensorflow/benchmarks/tree/master/scripts/tf_cnn_benchmarks

Summary

- HPL, HPCG and STREAM of Triad provide the basic metrics for the performance.
- RDMA is the must based on the results of mulit-node HPL and HPCG benchmarks.
- The latest CPU architecture increases memory bandwidth and therefore HPCG and STREAM performances increase a lot.
  - Indicates the suitable platform for memory intensive workload.
- Correlation between base and application benchmarks are ongoing.
  - Software stack and libraries
  - Optimization

Source: http://www.scoutingforbeer.org.uk/wp-content/uploads/2016/04/thank-you-1400x800-c-default.gif