SPEChpc™ 2021 Large Result

NVIDIA Corporation

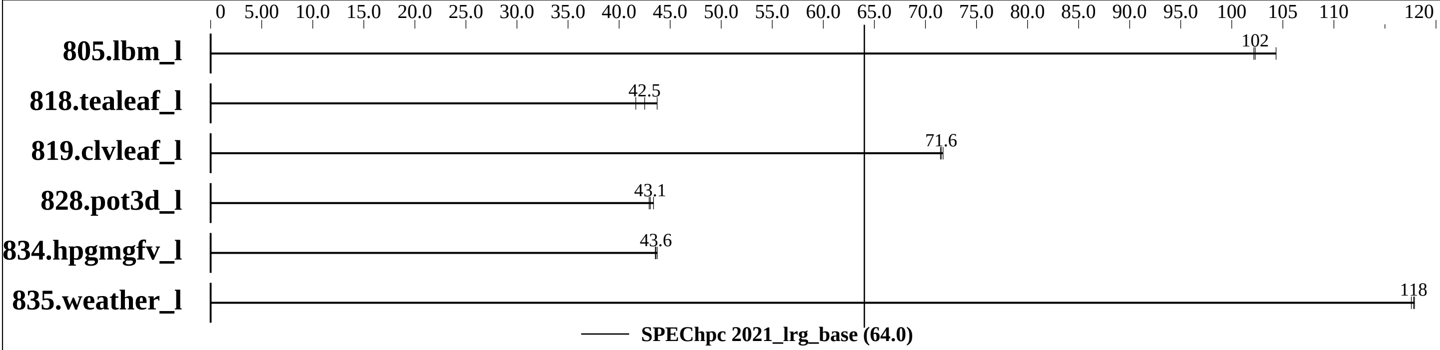
Selene: NVIDIA DGX SuperPOD
(AMD EPYC 7742 2.25 GHz, Tesla A100-SXM-80 GB)

SPEChpc 2021_lrg_base = 64.0

SPEChpc 2021_lrg_peak = Not Run

hpc2021 License: 019
Test Sponsor: NVIDIA Corporation
Tested by: NVIDIA Corporation

Test Date: Sep-2022
Hardware Availability: Jul-2020
Software Availability: Mar-2022



Results Table

Benchmark	Base									Peak								
	Model	Ranks	Thrds/Rnk	Seconds	Ratio	Seconds	Ratio	Seconds	Ratio	Model	Ranks	Thrds/Rnk	Seconds	Ratio	Seconds	Ratio	Seconds	Ratio
805.lbm_l	ACC	2048	8	26.7	102	26.1	104	**26.6**	**102**									
818.tealeaf_l	ACC	2048	8	33.2	43.7	**34.1**	**42.5**	34.8	41.6									
819.clvleaf_l	ACC	2048	8	**29.3**	**71.6**	29.3	71.7	29.4	71.5									
828.pot3d_l	ACC	2048	8	106	42.9	105	43.4	**106**	**43.1**									
834.hpgmgfv_l	ACC	2048	8	76.9	43.5	**76.8**	**43.6**	76.6	43.7									
835.weather_l	ACC	2048	8	29.1	118	29.1	118	**29.1**	**118**									

SPEChpc 2021_lrg_base = 64.0

SPEChpc 2021_lrg_peak = Not Run

Results appear in the order in which they were run. Bold underlined text indicates a median measurement.

NVIDIA Corporation

Selene: NVIDIA DGX SuperPOD
(AMD EPYC 7742 2.25 GHz, Tesla A100-SXM-80 GB)

SPEChpc 2021_lrg_base = 64.0

SPEChpc 2021_lrg_peak = Not Run

hpc2021 License: 019
Test Sponsor: NVIDIA Corporation
Tested by: NVIDIA Corporation
Test Date: Sep-2022
Hardware Availability: Jul-2020
Software Availability: Mar-2022

Hardware Summary

Type of System:	SMP
Compute Node:	DGX A100
Interconnects:	Multi-rail InfiniBand HDR fabric DDN EXAScalar file system
Compute Nodes Used:	128
Total Chips:	256
Total Cores:	16384
Total Threads:	32768
Total Memory:	256 TB
Max. Peak Threads:	--

Software Summary

Compiler:	C/C++/Fortran: Version 22.3 of NVIDIA HPC SDK for Linux
MPI Library:	OpenMPI Version 4.1.2rc4
Other MPI Info:	HPC-X Software Toolkit Version 2.10
Other Software:	None
Base Parallel Model:	ACC
Base Ranks Run:	2048
Base Threads Run:	8
Peak Parallel Models:	Not Run
Minimum Peak Ranks:	--
Maximum Peak Ranks:	--
Max. Peak Threads:	--
Min. Peak Threads:	--

Node Description: DGX A100

Hardware

Number of nodes:	128
Uses of the node:	compute
Vendor:	NVIDIA Corporation
Model:	NVIDIA DGX A100 System
CPU Name:	AMD EPYC 7742
CPU(s) orderable:	2 chips
Chips enabled:	2
Cores enabled:	128
Cores per chip:	64
Threads per core:	2
CPU Characteristics:	Turbo Boost up to 3400 MHz
CPU MHz:	2250
Primary Cache:	32 KB I + 32 KB D on chip per core
Secondary Cache:	512 KB I+D on chip per core
L3 Cache:	256 MB I+D on chip per chip (16 MB shared / 4 cores)
Other Cache:	None
Memory:	2 TB (32 x 64 GB 2Rx8 PC4-3200AA-R)
Disk Subsystem:	OS: 2TB U.2 NVMe SSD drive Internal Storage: 30TB (8x 3.84TB U.2 NVMe SSD drives)
Other Hardware:	None
Accel Count:	8
Accel Model:	Tesla A100-SXM-80 GB
Accel Vendor:	NVIDIA Corporation
Accel Type:	GPU
Accel Connection:	NVLINK 3.0, NVSWITCH 2.0 600 GB/s
Accel ECC enabled:	Yes
Accel Description:	See Notes
Adapter:	NVIDIA ConnectX-6 MT28908
Number of Adapters:	8
Slot Type:	PCIe Gen4

(Continued on next page)

Software

Accelerator Driver:	NVIDIA UNIX x86_64 Kernel Module 470.103.01
Adapter:	NVIDIA ConnectX-6 MT28908
Adapter Driver:	InfiniBand: 5.4-3.4.0.0
Adapter Firmware:	InfiniBand: 20.32.1010
Adapter:	NVIDIA ConnectX-6 MT28908
Adapter Driver:	Ethernet: 5.4-3.4.0.0
Adapter Firmware:	Ethernet: 20.32.1010
Operating System:	Ubuntu 20.04 5.4.0-121-generic
Local File System:	ext4
Shared File System:	Lustre
System State:	Multi-user, run level 3
Other Software:	None

Node Description: DGX A100

Hardware (Continued)

Data Rate:	200 Gb/s
Ports Used:	1
Interconnect Type:	InfiniBand / Communication
Adapter:	NVIDIA ConnectX-6 MT28908
Number of Adapters:	2
Slot Type:	PCIe Gen4
Data Rate:	200 Gb/s
Ports Used:	2
Interconnect Type:	InfiniBand / FileSystem

Interconnect Description: Multi-rail InfiniBand HDR fabric

Hardware

Vendor:	NVIDIA
Model:	N/A
Switch Model:	NVIDIA Quantum QM8700
Number of Switches:	164
Number of Ports:	40
Data Rate:	200 GB/s per port
Firmware:	MLNX-OS v3.10.2202
Topology:	Full three-level fat-tree
Primary Use:	Inter-process communication

Software

: --

Interconnect Description: DDN EXAScalar file system

Hardware

Vendor:	NVIDIA
Model:	N/A
Switch Model:	NVIDIA Quantum QM8700
Number of Switches:	26
Number of Ports:	40
Data Rate:	200 GB/s per port
Firmware:	MLNX-OS v3.10.2202
Topology:	Full three-level fat-tree
Primary Use:	Global storage

Software

: --

Compiler Invocation Notes

```
Binaries built and run within a NVHPC SDK 22.3 CUDA 11.0 Ubuntu 20.04
 Container available from NVIDIA GPU Cloud (NGC):
  https://ngc.nvidia.com/catalog/containers/nvidia:nvhpc
  https://catalog.ngc.nvidia.com/orgs/nvidia/containers/nvhpc/tags
```

Submit Notes

```
The config file option 'submit' was used.
 MPI startup command:
   srun command was used to start MPI jobs.

 Individual Ranks were bound to the NUMA nodes, GPUs and NICs using this "wrapper.GPU" bash-script for the case of 1 rank per GPU

   ln -s -f libnuma.so.1 /usr/lib/x86_64-linux-gnu/libnuma.so
   export LD_LIBRARY_PATH+=:/usr/lib/x86_64-linux-gnu
   export LD_RUN_PATH+=:/usr/lib/x86_64-linux-gnu
   declare -a NUMA_LIST
   declare -a  GPU_LIST
   declare -a  NIC_LIST
   NUMA_LIST=($NUMAS)
   GPU_LIST=($GPUS)
   NIC_LIST=($NICS)
   export UCX_NET_DEVICES=${NIC_LIST[$SLURM_LOCALID]}:1
   export OMPI_MCA_btl_openib_if_include=${NIC_LIST[$SLURM_LOCALID]}
   export CUDA_VISIBLE_DEVICES=${GPU_LIST[$SLURM_LOCALID]}
   numactl -l -N ${NUMA_LIST[$SLURM_LOCALID]} $*

 and this "wrapper.MPS" bash-script for the oversubscribed case.

   ln -s -f libnuma.so.1 /usr/lib/x86_64-linux-gnu/libnuma.so
   export LD_LIBRARY_PATH+=:/usr/lib/x86_64-linux-gnu
   export LD_RUN_PATH+=:/usr/lib/x86_64-linux-gnu
   declare -a NUMA_LIST
   declare -a  GPU_LIST
   declare -a  NIC_LIST
   NUMA_LIST=($NUMAS)
   GPU_LIST=($GPUS)
   NIC_LIST=($NICS)
   NUM_GPUS=${#GPU_LIST[@]}
   RANKS_PER_GPU=$((SLURM_NTASKS_PER_NODE / NUM_GPUS))
   GPU_LOCAL_RANK=$((SLURM_LOCALID / RANKS_PER_GPU))
   export UCX_NET_DEVICES=${NIC_LIST[$GPU_LOCAL_RANK]}:1
   export OMPI_MCA_btl_openib_if_include=${NIC_LIST[$GPU_LOCAL_RANK]}
   set +e
   nvidia-cuda-mps-control -d 1>&2
   set -e
   export CUDA_VISIBLE_DEVICES=${GPU_LIST[$GPU_LOCAL_RANK]}
   numactl -l -N ${NUMA_LIST[$GPU_LOCAL_RANK]} $*
   if [ $SLURM_LOCALID -eq 0 ]
   then
       echo 'quit' | nvidia-cuda-mps-control 1>&2
   fi
```

General Notes

```
Full system details documented here:
https://images.nvidia.com/aem-dam/Solutions/Data-Center/gated-resources/nvidia-dgx-superpod-a100.pdf

Environment variables set by runhpc before the start of the run:
SPEC_NO_RUNDIR_DEL = "on"
```

Platform Notes

```
 Detailed A100 Information from nvaccelinfo
 CUDA Driver Version:           11040
 NVRM version:                  NVIDIA UNIX x86_64 Kernel Module 470.7.01
 Device Number:                 0
 Device Name:                   NVIDIA A100-SXM-80 GB
 Device Revision Number:        8.0
 Global Memory Size:            85198045184
 Number of Multiprocessors:     108
 Concurrent Copy and Execution: Yes
 Total Constant Memory:         65536
 Total Shared Memory per Block: 49152
 Registers per Block:           65536
 Warp Size:                     32
 Maximum Threads per Block:     1024
 Maximum Block Dimensions:      1024, 1024, 64
 Maximum Grid Dimensions:       2147483647 x 65535 x 65535
 Maximum Memory Pitch:          2147483647B
 Texture Alignment:             512B
 Clock Rate:                    1410 MHz
 Execution Timeout:             No
 Integrated Device:             No
 Can Map Host Memory:           Yes
 Compute Mode:                  default
 Concurrent Kernels:            Yes
 ECC Enabled:                   Yes
 Memory Clock Rate:             1593 MHz
 Memory Bus Width:              5120 bits
 L2 Cache Size:                 41943040 bytes
 Max Threads Per SMP:           2048
 Async Engines:                 3
 Unified Addressing:            Yes
 Managed Memory:                Yes
 Concurrent Managed Memory:     Yes
 Preemption Supported:          Yes
 Cooperative Launch:            Yes
   Multi-Device:                Yes
 Default Target:                cc80
```

Compiler Version Notes

```
==============================================================================
 CC  805.lbm_l(base) 818.tealeaf_l(base) 834.hpgmgfv_l(base)
------------------------------------------------------------------------------
nvc 22.3-0 64-bit target on x86-64 Linux -tp zen2-64
NVIDIA Compilers and Tools
Copyright (c) 2022, NVIDIA CORPORATION & AFFILIATES.  All rights reserved.
------------------------------------------------------------------------------

==============================================================================
 FC  819.clvleaf_l(base) 828.pot3d_l(base) 835.weather_l(base)
------------------------------------------------------------------------------
nvfortran 22.3-0 64-bit target on x86-64 Linux -tp zen2-64
NVIDIA Compilers and Tools
Copyright (c) 2022, NVIDIA CORPORATION & AFFILIATES.  All rights reserved.
------------------------------------------------------------------------------
```

Base Compiler Invocation

C benchmarks:
`mpicc`

Fortran benchmarks:
`mpif90`

Base Portability Flags

805.lbm_l: `-DSPEC_OPENACC_NO_SELF`

Base Optimization Flags

C benchmarks:
`-fast -DSPEC_ACCEL_AWARE_MPI -acc=gpu -gpu=cuda11.0 -gpu=cc80 -Mstack_arrays -Mfprelaxed -Mnouniform -tp=zen2`

Fortran benchmarks:
`-DSPEC_ACCEL_AWARE_MPI -fast -acc=gpu -gpu=cuda11.0 -gpu=cc80 -Mstack_arrays -Mfprelaxed -Mnouniform -tp=zen2`

Base Other Flags

C benchmarks (except as noted below):
`-Ispecmpitime -w`

834.hpgmgfv_l: `-Ispecmpitime  -w`

Fortran benchmarks (except as noted below):
`-w`

819.clvleaf_l: `-Ispecmpitime -w`

The flags file that was used to format this result can be browsed at
http://www.spec.org/hpc2021/flags/nv2021_flags_v1.0.3.2022-11-03.html

You can also download the XML flags source by saving the following link:
http://www.spec.org/hpc2021/flags/nv2021_flags_v1.0.3.2022-11-03.xml

Tested with SPEChpc2021 v1.1.7 on 2022-09-27 11:50:30-0400.
Report generated on 2022-11-03 14:04:16 by hpc2021 PS formatter v1.0.3.
Originally published on 2022-11-02.