



HPCG BENCHMARK ON CORAL SYSTEMS

Massimiliano Fatica, Everett Phillips

▶ GPUs in HPC:

- ▶ Flops: DP, SP, HP

- ▶ Memory BW*:

 - ▶ P100 16GB: 559 GB/s TRIAD STREAM (76.4%) (Peak: 715 MHz x 4096 bit= 732 GB/s)

 - ▶ V100 16GB: 851 GB/s TRIAD STREAM (94.8%) (Peak: 877 MHz x 4096 bit= 898 GB/s)

 - ▶ V100 32GB: 826 GB/s TRIAD STREAM (92.1%) (Peak: 877 MHz x 4096 bit= 898 GB/s)

- ▶ Power efficiency

▶ HPCG is all about:

- ▶ local memory bandwidth

- ▶ network (at large scale)

* CUDA/DRIVER: 10.0.130/410.48, GB=10⁹ bytes

HPCG ON GPU

- ▶ HPCG GPU implementation in CUDA:
 - ▶ Details in “A CUDA implementation of the HPCG benchmark”, PBMS@SC14
 - ▶ Improved Symmetric Gauss-Seidel (Kumahata et al, SC 16 HPCG BOF)
- ▶ GPUs are the fastest processor to run HPCG

CORAL SYSTEMS

▶ SUMMIT at ORNL:

- ▶ 4608 nodes with 2 IBM Power9 and 6 Nvidia V100 GPUs
- ▶ Dual rail EDR Infiniband in fat-tree topology
- ▶ #1 system in Nov18 Top500 (143.5 PF)
- ▶ #1 system in Nov18 HPCG (2.926 PF on 4356 nodes)

▶ SIERRA at LLNL:

- ▶ 4238 nodes with 2 IBM Power9 and 4 Nvidia V100 GPUs
- ▶ Dual rail EDR Infiniband in fat-tree topology
- ▶ #2 system in Nov18 Top500 (94.6 PF)
- ▶ #2 system in Nov18 HPCG (1.796 PF on 4096 nodes)



Coral systems are the first to break the PF barrier in HPCG

▶ Current reporting does not capture relevant information:

- ▶ No mentions of the size of the machine or memory subsystem

June 2018 HPCG Results							
Rank	Site	Computer	Cores	HPL Rmax (Pflop/s)	TOP500 Rank	HPCG (Pflop/s)	Fraction of Peak
1	DOE/SC/ORNL USA	Summit – AC922, IBM POWER9 22C 3.07GHz, dual-rail Mellanox EDR Infiniband, NVIDIA Volta V100 IBM	2,392,000	122.300	1	2.926	1.5%
2	DOE/NNSA/LLNL USA	Sierra – S922LC, Power9 180224C 3.1GHz, Mellanox EDR, NVIDIA Tesla V100 IBM / NVIDIA / Mellanox	835,584	71.610	3	1.796	1.5%
3	RIKEN Advanced Institute for Computational Science Japan	K computer – , SPARC64 VIIIfx 2.0GHz, Tofu interconnect Fujitsu	705,024	10.510	16	0.603	5.3%
4	DOE/NNSA/LANL/SNL USA	Trinity – Cray XC40, Intel Xeon E5-2698 v3 300160C 2.3GHz, Aries Cray	979,072	14.137	9	0.546	1.8%
5	Swiss National Supercomputing Centre (CSCS) Switzerland	Piz Daint – Cray XC50, Intel Xeon E5- 2690v3 12C 2.6GHz, Cray Aries, NVIDIA Tesla P100 16GB Cray	361,760	19.590	6	0.486	1.9%
6	National Supercomputing Center in Wuxi China	Sunway TaihuLight – Sunway MPP, SW26010 260C 1.45GHz, Sunway NRCPC	10,649,600	93.015	2	0.481	0.4%
129	University of Tartu Estonia	Vedur – DL, AMD Opteron 6276 2.3 GHz, Mellanox 4x QDR Infiniband HPE	32			0.00010	4.4%
130	EXPRESS Rack 6 The International Space Station (ISS)	Spaceborne Computer – Apollo 40, Intel Xeon E5-2620 v4 8C 2.1GHz, FDR Infiniband HPE	32	0.001		0.000034	2.9%

FLOPS WRONG METRIC

GM200 chip: DP:SP=1:32, SP=6.72 TF, DP=0.21TF, 384 bit at 6 GHz

GK210 chip: DP:SP=1:3 , SP=4.37 TF, DP=1.45TF, 384 bit at 5 GHz

1 x K80 (2 GK210 GPUs), ECC enabled, clk=875

1.8% of peak

2x1x1 process grid

256x256x256 local domain

SpMV = 49.1 GF (309.1 GB/s Effective) 24.5 GF_per (154.6 GB/s Effective)

SymGS = 62.2 GF (480.2 GB/s Effective) 31.1 GF_per (240.1 GB/s Effective)

total = 58.7 GF (445.3 GB/s Effective) 29.4 GF_per (222.7 GB/s Effective)

final = 55.1 GF (417.5 GB/s Effective) **27.5 GF_per** (208.8 GB/s Effective)

2 x M40 (2 GM200 GPUs), ECC enabled, clk=1114

18% of peak

2x1x1 process grid

256x256x256 local domain

SpMV = 69.4 GF (437.2 GB/s Effective) 34.7 GF_per (218.6 GB/s Effective)

SymGS = 83.7 GF (645.7 GB/s Effective) 41.8 GF_per (322.8 GB/s Effective)

total = 79.6 GF (603.7 GB/s Effective) 39.8 GF_per (301.9 GB/s Effective)

final = 74.2 GF (562.7 GB/s Effective) **37.1 GF_per** (281.4 GB/s Effective)

HPCG LIST

System	HPCG Flops (TF)	#Proc	HPCG/ proc	BW/proc GB/s	Total BW GB/S	Eff. (F/B)
Summit	2950000	26136 (4356*6)	112 GF	851 (898)	22241736 (23470128)	.131 .124
Sierra	1796000	16384 (4096*4)	109.6 GF	851 (898)	13942784 (14712832)	.128 .122
K	603000	82944	7.2 GF	43 (64)	3566592 (5308416)	.167 .112
Piz Daint	486000	5320	91 GF	559 (732)	2973880 (3894240)	.162 .124