HPCG UPDATE: ISC'15

Jack Dongarra
Michael Heroux
Piotr Luszczek

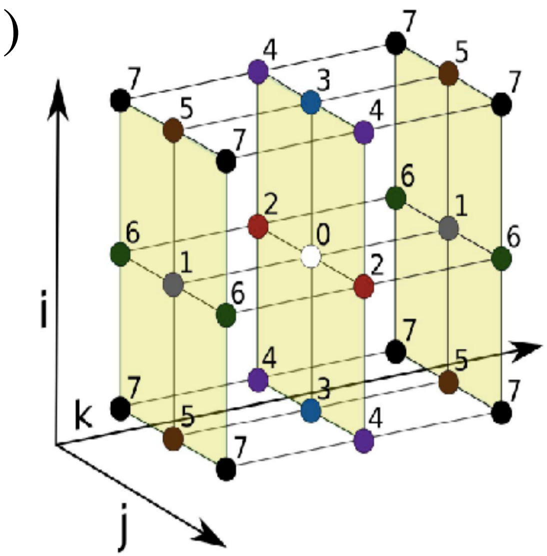
HPCG Snapshot

- High Performance Conjugate Gradient (HPCG).
- Solves $Ax=b$, A large, sparse, b known, x computed.
- An optimized implementation of PCG contains essential computational and communication patterns that are prevalent in a variety of methods for discretization and numerical solution of PDEs

- Patterns:
  - Dense and sparse computations.
  - Dense and sparse collective.
  - Multi-scale execution of kernels via MG (truncated) V cycle.
  - Data-driven parallelism (unstructured sparse triangular solves).
- Strong verification (via spectral properties of PCG).

Model Problem Description

- Synthetic discretized 3D PDE (FEM, FVM, FDM).
- Zero Dirichlet BCs, Synthetic RHS s.t. solution = 1.
- Local domain: $(n_x \times n_y \times n_z)$
- Process layout: $(np_x \times np_y \times np_z)$
- Global domain: $(n_x * np_x) \times (n_y * np_y) \times (n_z * np_z)$
- Sparse matrix:
  - 27 nonzeros/row interior.
  - 8 – 18 on boundary.
  - Symmetric positive definite.



27-point stencil operator

Merits of HPCG

- Includes major communication/computational patterns.
  - Represents a minimal collection of the major patterns.
- Rewards investment in:
  - High-performance collective ops.
  - Local memory system performance.
  - Low latency cooperative threading.
- Detects/measures variances from bitwise reproducibility.
- Executes kernels at several (tunable) granularities:
  - nx = ny = nz = 104 gives
  - nlocal = 1,124,864; 140,608; 17,576; 2,197
  - ComputeSymGS with multicoloring adds one more level:
    - 8 colors.
    - Average size of color = 275.
    - Size ratio (largest:smallest): 4096
  - Provide a “natural” incentive to run a big problem.

HPL vs. HPCG: Bookends

- Some see HPL and HPCG as “bookends” of a spectrum.
  - Applications teams know where their codes lie on the spectrum.
  - Can gauge performance on a system using both HPL and HPCG numbers.

HPCG Status

Special Issue: International Journal of High Performance Computer Applications

1. Reference HPCG.
2. Intel.
3. Nvidia.
4. NUDT.
5. Riken.
6. Coming a little later: IBM.

- Discussion and results from vendor optimizations.
- Articles in final review.
- Some highlights…

Rewards investment high performance collectives.

"Edison spends only 1.9% of the total time in all-reduce while SuperMUC, Occigen, and Stampede spend 12.9%,5.9%, and 22.0%, respectively. We believe this difference primarily comes from that Edison uses a low-diameter high-radix Aries network with Dragonfly topology."

Intel HPCG Paper

Collectives futures

- “Addressing the bottleneck in collective communications will be also an important challenge as the collectives are shown to often take well above 10% of the total time. Even though high-radix Dragonfly topology considerably speedups the collectives, we envision that continued innovation in network infrastructure will be necessary due to ever increasing concurrency in high performance computing systems.”

Impact broader set of computations

"The optimizations described in this paper are not limited to the HPCG benchmark and can be also applicable to other problems and sparse solvers as exemplified by our evaluation with unstructured matrices shown in [our previous report]."

Looking toward next generation memories

“We expect challenges and opportunities laid out for HPCG in the next few years. One of the significant challenges will be effective use of emerging memory technologies and the accompanied diversification of memory hierarchy.”

Detecting FP Variations (Reproducibility)

Residual=4.25079640861055785883e-08 (0x1.6d240066fda73p-25)
Residual=4.25079640861032293954e-08 (0x1.6d240066fd910p-25)
Residual=4.25079640861079079289e-08 (0x1.6d240066fdbd3p-25)
Residual=4.25079640861054528568e-08 (0x1.6d240066fda60p-25)
Residual=4.25079640861068491377e-08 (0x1.6d240066fdb33p-25)
Residual=4.25079640861059094605e-08 (0x1.6d240066fdaa5p-25)

“The code correctly identified small variations in the residuals, caused by the network off-loading collectives. There is a small improvement in performance but the off-loading collectives introduce a small non-reproducibility.”

Vendor improvement: Intel 4X

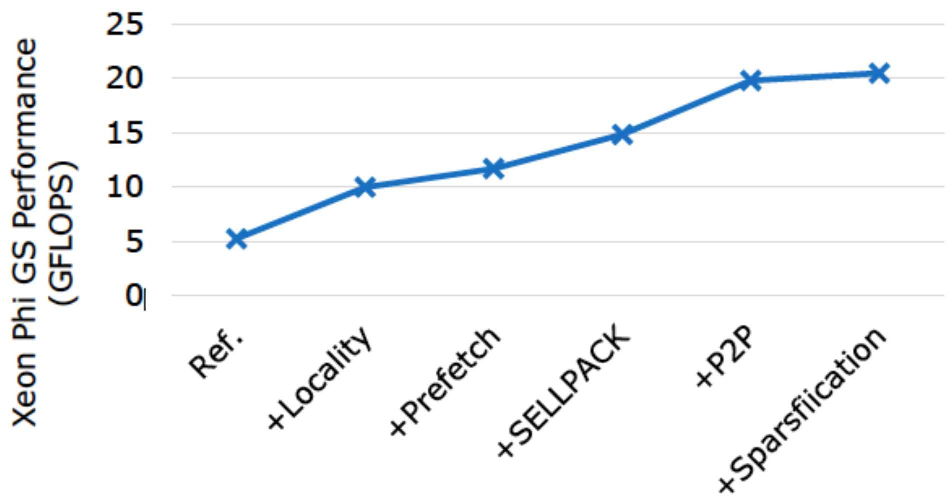


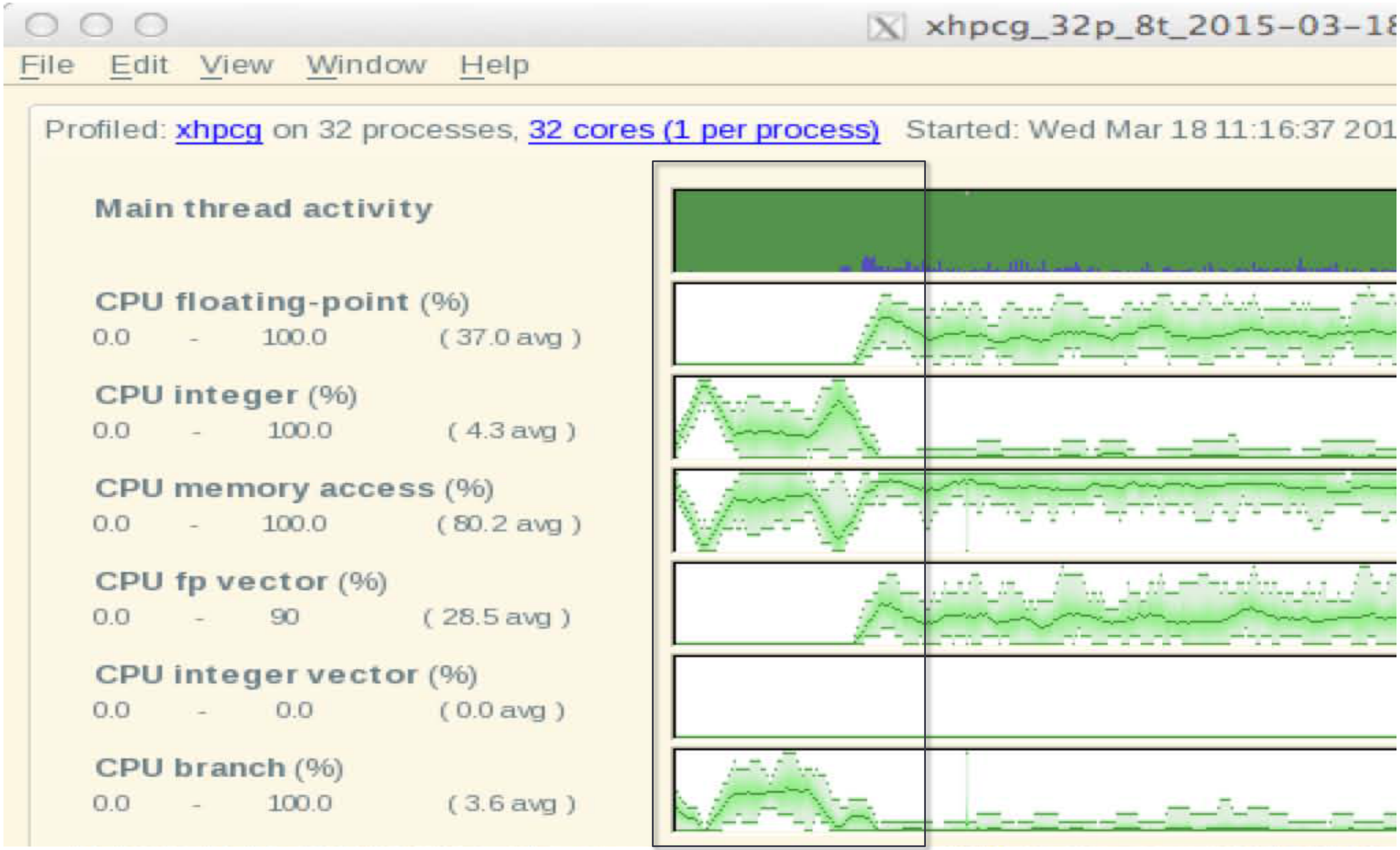
Fig. 5: The impact of optimizations on the Xeon Phi performance of SymGS parallelized with task scheduling.

- `Ref.`: the reference implementation ran with 240 MPI ranks
- `+Locality`: storage layout optimization for locality (Section IV-A1)
- `+Prefetch`: software prefetches
- +SELLPACK: vectorization-friendly matrix storage format [43]
- +P2P: point-to-point synchronization instead of barriers
- `+Sparsification`: eliminating unnecessary synchronization [10]

Next (and last) Major Version 3.X

- Concern: Too much like STREAMS.
  - Not true, from previous results.
  - Still: Interested in mixing in address/integer/logic instructions.
- Approach:
  - Time problem generation.
  - Include this time as part of overhead.
  - Overhead: Generation + Vendor optimization costs.

HPCG 2.4 Profile (Allinea output)



Other Items

- Reference version on GitHub:
  - https://github.com/hpcg-benchmark/hpcg
  - Website: hpcg-benchark.org, includes results auto-upload from yaml.
  - Mail list hpcg.benchmark@gmail.com
- Next event: SC'15:
  - 40 entries so far, expect more.
  - Release of HPCG 3.0.
  - Transition from version 2.4 to 3.0 is under discussion.

Summary

- HPCG is
  - Addressing original goals.
  - Rewarding vendor investment in features we care about.
- HPCG has traction.
  - Original goal of top 50 systems seems reachable, and more.
- Biggest challenge (my bias):
  - Pre-mature conclusions based on incomplete analysis of reference version.
  - IJHPCA papers should dispel these concerns.
- Version 3.X will (hopefully) be the final major version.
- HPL and HPCG make a nice set of bookends.
  - Anyone got a (wood) router?

HPCG RANKINGS JULY 2015

HPCG Highlights

- 40 Systems:
  - Up from 25 at SC’14 and 15 at ISC’14.
  - Most entries from the very top of the TOP500 list.
- New supercomputers (also coming to TOP500) are:
  - KAUST Shaheen II
  - Moscow State: Lomonosov 2
- Strong showing from Japan and NEC SX machines:
  - Achieve over 10% of peak performance with HPCG
- Updated results from TACC with larger scale of the system tested.
- IBM BlueGene machines make their first appearance on the list.

And The Winners Are…

HPCG

PRESENTED AT ISC High Performance JULY 15, 2015

NUMBER 3

SYSTEM **Titan**
DOE
Oak Ridge National Laboratory
USA

ACHIEVED **0.322** Pflop/s

JACK DONGARRA

MICHAEL HEROUX

PIOTR LUSZCZEK

IN COLLABORATION WITH ICL UT INNOVATIVE COMPUTING LABORATORY THE UNIVERSITY of TENNESSEE

Sandia National Laboratories

SPONSORED BY DEPARTMENT OF ENERGY UNITED STATES OF AMERICA

HPCG

PRESENTED AT ISC High Performance JULY 15, 2015

NUMBER 2

SYSTEM **K computer**
RIKEN Advanced Institute for Computational Science
JAPAN

ACHIEVED **0.461** Pflop/s

JACK DONGARRA

MICHAEL HEROUX

PIOTR LUSZCZEK

IN COLLABORATION WITH ICL UT INNOVATIVE COMPUTING LABORATORY THE UNIVERSITY of TENNESSEE

Sandia National Laboratories

SPONSORED BY DEPARTMENT OF ENERGY UNITED STATES OF AMERICA

HPCG

PRESENTED AT ISC High Performance JULY 15, 2015

NUMBER 1

SYSTEM **Tianhe-2**
National Super Computer
Center in Guangzhou
CHINA

ACHIEVED **0.580** Pflop/s

JACK DONGARRA

MICHAEL HEROUX

PIOTR LUSZCZEK

IN COLLABORATION WITH ICL UT INNOVATIVE COMPUTING LABORATORY THE UNIVERSITY of TENNESSEE

Sandia National Laboratories

SPONSORED BY DEPARTMENT OF ENERGY UNITED STATES OF AMERICA

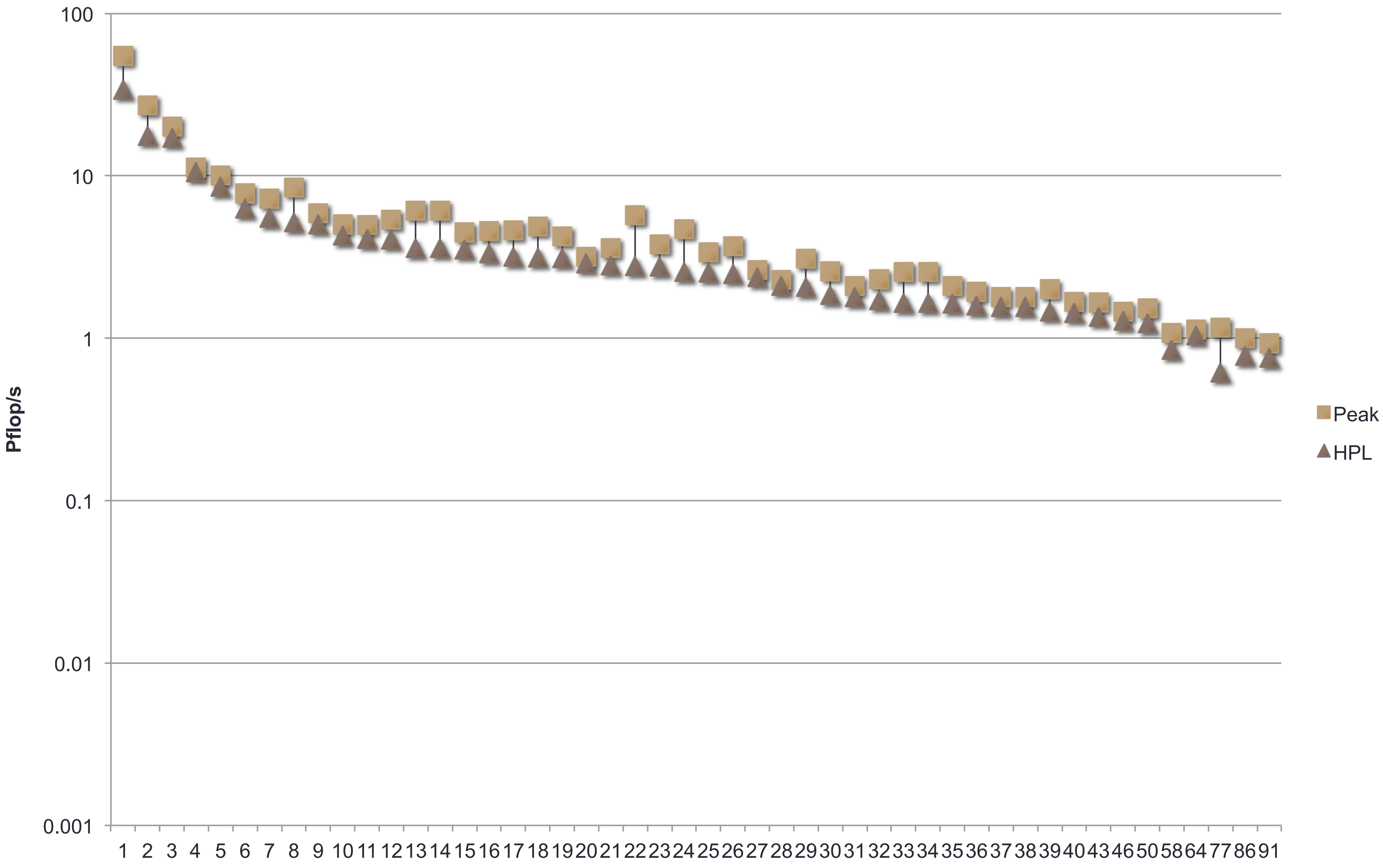
HPCG Results, July 2015

Rank	Site	Computer	Cores	HPL Rmax (Pflops)	HPL Rank	HPCG (Pflops)	HPCG /HPL	% of Peak
1	NSCC / Guangzhou	Tianhe-2 NUDT, Xeon 12C 2.2GHz + Intel Xeon Phi 57C + Custom	3,120,000	33.9	1	.580	1.7%	1.1%
2	RIKEN Advanced Inst for Comp Sci	K computer Fujitsu SPARC64 VIIIfx 8C + Custom	705,024	10.5	4	.461	4.4%	4.1%
3	DOE/OS Oak Ridge Nat Lab	Titan, Cray XK7 AMD 16C + Nvidia Kepler GPU 14C + Custom	560,640	17.6	2	.322	1.8%	1.2%
4	DOE/OS Argonne Nat Lab	Mira BlueGene/Q, Power BQC 16C 1.60GHz + Custom	786,432	8.59	5	.167	1.9%	1.7%
5	NASA Ames	Pleiades, SGI ICE X, Intel 2.6,2.8,2.5 GHz+ IB	186,288	4.09	11	.132	3.2%	2.7%
6	Swiss CSCS	Piz Daint, Cray XC30, Xeon 8C + Nvidia Kepler 14C + Custom	115,984	6.27	6	.125	2.0%	1.6%
7	KAUST	Shaheen II,Cray XC40, Xeon 16C 2.3GHz + Custom	196,608	5.54	7	.114	2.1%	1.6%
8	Texas Advanced Computing Center	Stampede, Dell Intel 8c + Intel Xeon Phi 61c + IB	522,080	5.17	8	.097	1.9%	1.0%
9	Leibniz Rechenzentrum	SuperMUC, Intel 8C + IB	147,456	2.90	20	.0833	2.9%	2.6%
10	EPSRC/University of Edinburgh	ARCHER - Cray XC30, Xeon 12C 2.7GHz + Custom	118,080	1.64	34	.0808	4.9%	3.2%

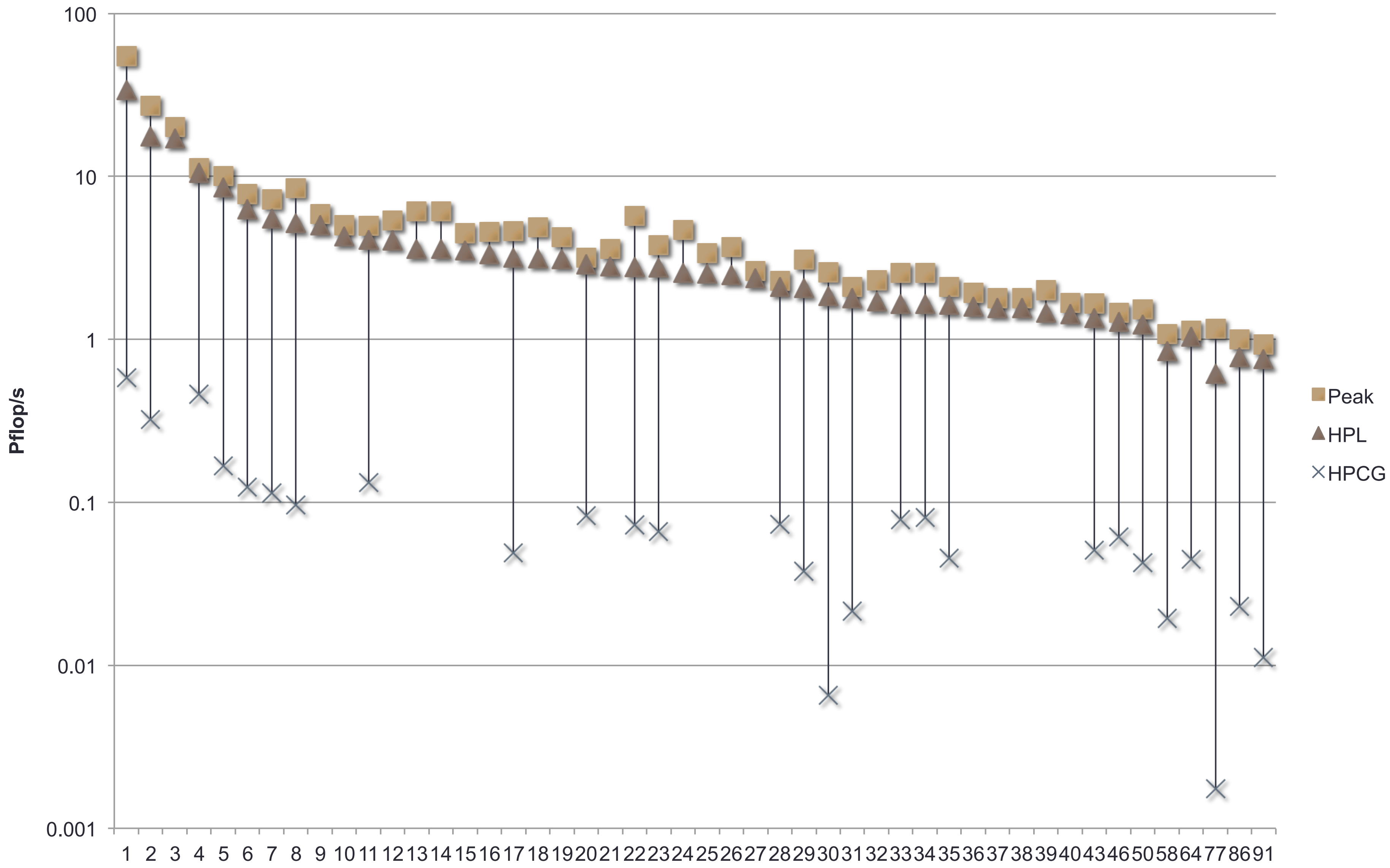
HPCG Results, July 2015

Rank	Site	Computer	Cores	HPL Rmax (Pflops)	HPL Rank	HPCG (Pflops)	HPCG /HPL	% of Peak
11	DOE/OS LBNL	Edison, Cray XC30, Xeon, 12c, 2,4GHz + Custom	133,824	1.66	33	.0786	4.8%	3.1%
12	Plasma Simulator	Fujitsu FX100, Sparc64 Xifx 32C + custom	82,944	2.38	27	.073	3.1%	2.8%
13	GSIC Center TiTech	Tsubame 2.5 Xeon 6C, 2.93GHz + Nvidia K20x + IB	76,032	2.79	22	.0725	2.6%	1.3%
14	HLRS/Universitaet Stuttgart	Hornet Cray XC40, Xeon 2.5GHz + custom	94,656	2.76	23	.066	2.4%	1.7%
15	Max-Planck	iDataPlex Xeon 10C, 2.8GHz + IB	65,320	1.28	46	.061	4.8%	4.2%
16	Earth Simulator	NEC SX-ACE 4C, 1 GHz + custom	8,192	0.487		.058	12%	11%
17	CEA/TGCC-GENCI	Curie thin nodes Bullx B510 Intel Xeon 8C 2.7 GHz + IB	77,184	1.36	43	.051	3.8%	3.1%
18	Exploration and Production Eni S.p.A.	HPC2, Intel Xeon 10C 2.8 GHz + Nvidia Kepler 14C + IB	62,640	3.00	17	.049	1.6%	1.2%
19	Grand Equipement National de Calcul Intensif	Occigen Bullx Xeon 12C 2.6Ghz + IB	50,544	1.63	35	.045	2.8%	2.2%
20	Oakleaf-FX	PIMEHPC FX10, Sparc64 16C, 1.85 GHz + custom	76,800	1.04	64	.0448	4.3%	3.9%

Peak, HPL Pflop/s



Peak, HPL, HPCG Pflop/s



HPCG Tech Reports

Toward a New Metric for Ranking High Performance Computing Systems

- Jack Dongarra and Michael Heroux

HPCG Technical Specification

- Jack Dongarra, Michael Heroux, Piotr Luszczek

SANDIA REPORT
SAND2013- 8752
Unlimited Release
Printed October 2013

HPCG Technical Specification

Michael A. Heroux, Sandia National Laboratories[1]
Jack Dongarra and Piotr Luszczek, University of Tennessee

Prepared by

SANDIA REPORT
SAND2013-4744
Unlimited Release
Printed June 2013

Toward a New Metric for Ranking High Performance Computing Systems

Jack Dongarra, University of Tennessee
Michael A. Heroux, Sandia National Laboratories[1]

Prepared by
Sandia National Laboratories
Albuquerque, New Mexico 87185 and Livermore, California 94550

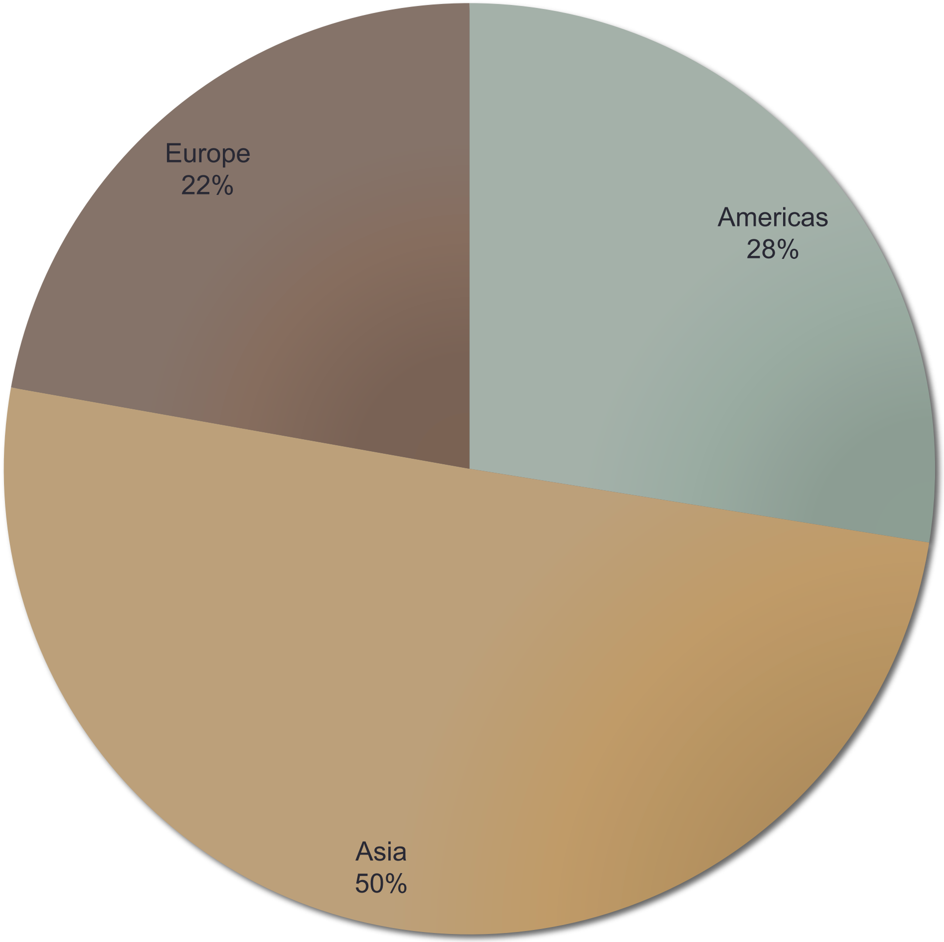
Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

Approved for public release; further dissemination unlimited.

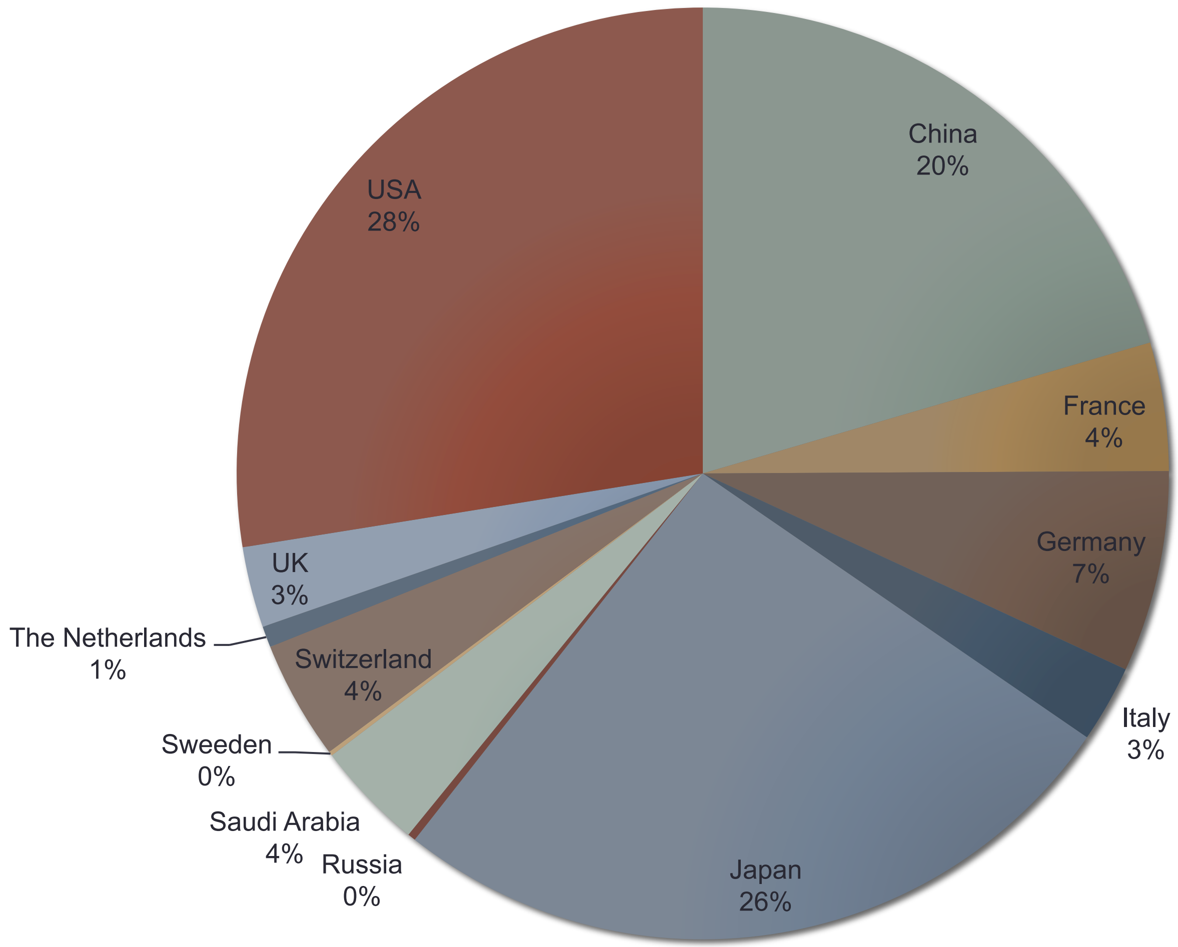
Sandia National Laboratories

[1] Corresponding Author, maherou@sandia.gov

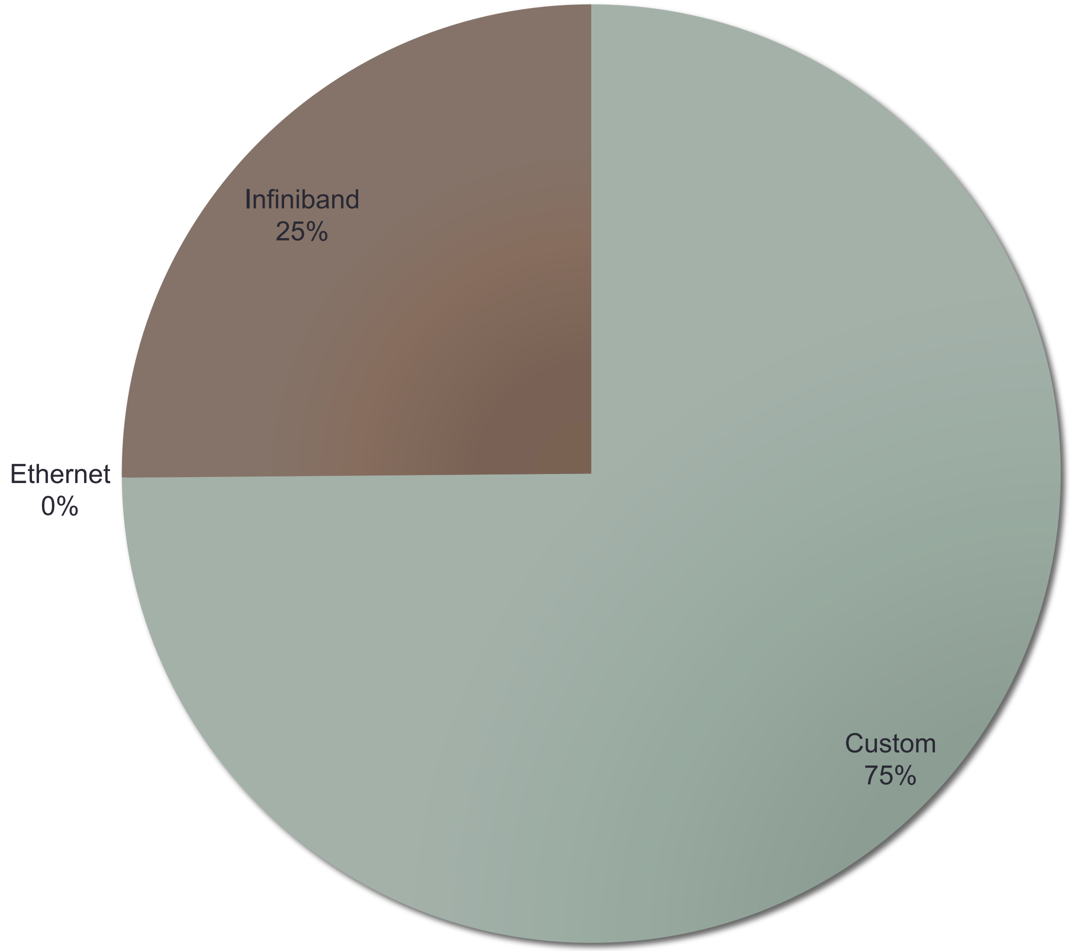
By Region



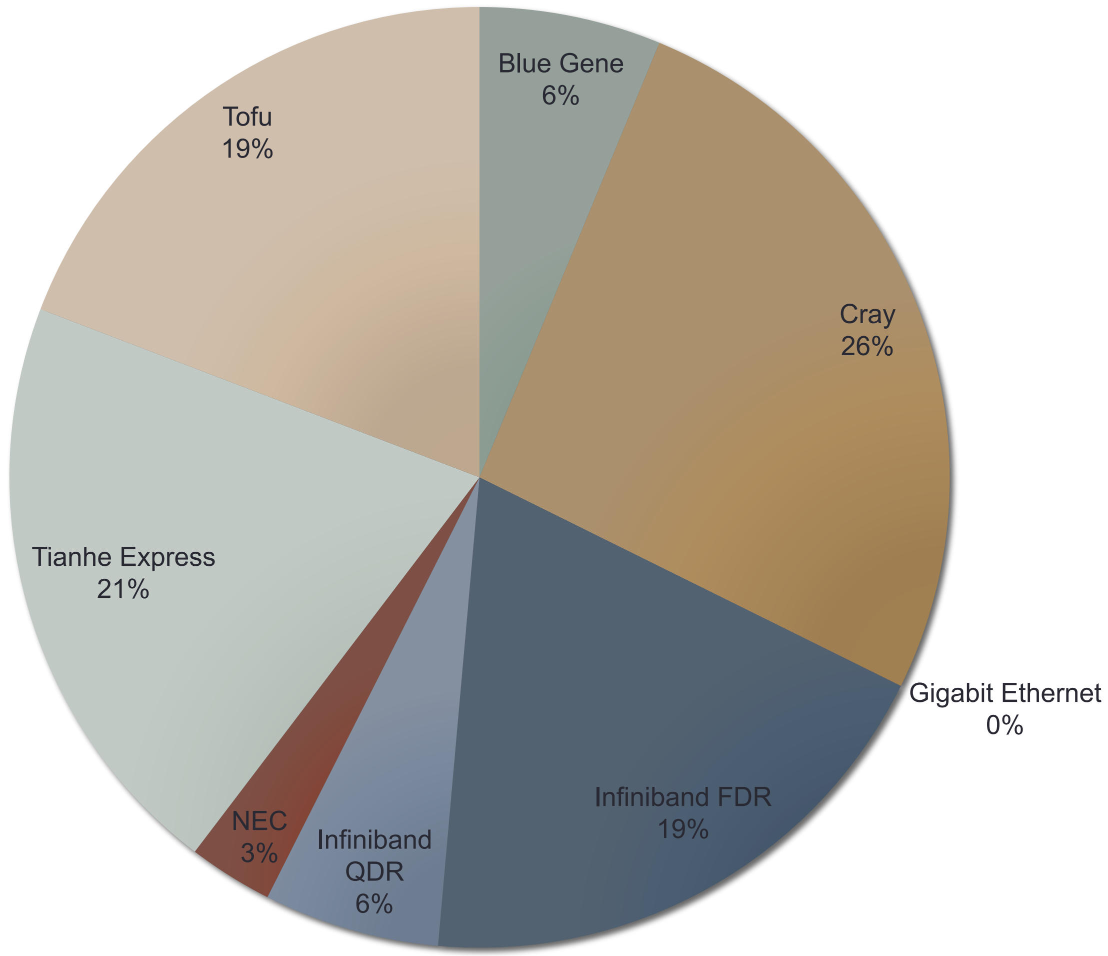
By Country


China
20%
France
4%
Germany
7%
Italy
3%
Japan
26%
Russia
0%
Saudi Arabia
4%
Sweeden
0%
Switzerland
4%
The Netherlands
1%
UK
3%
USA
28%

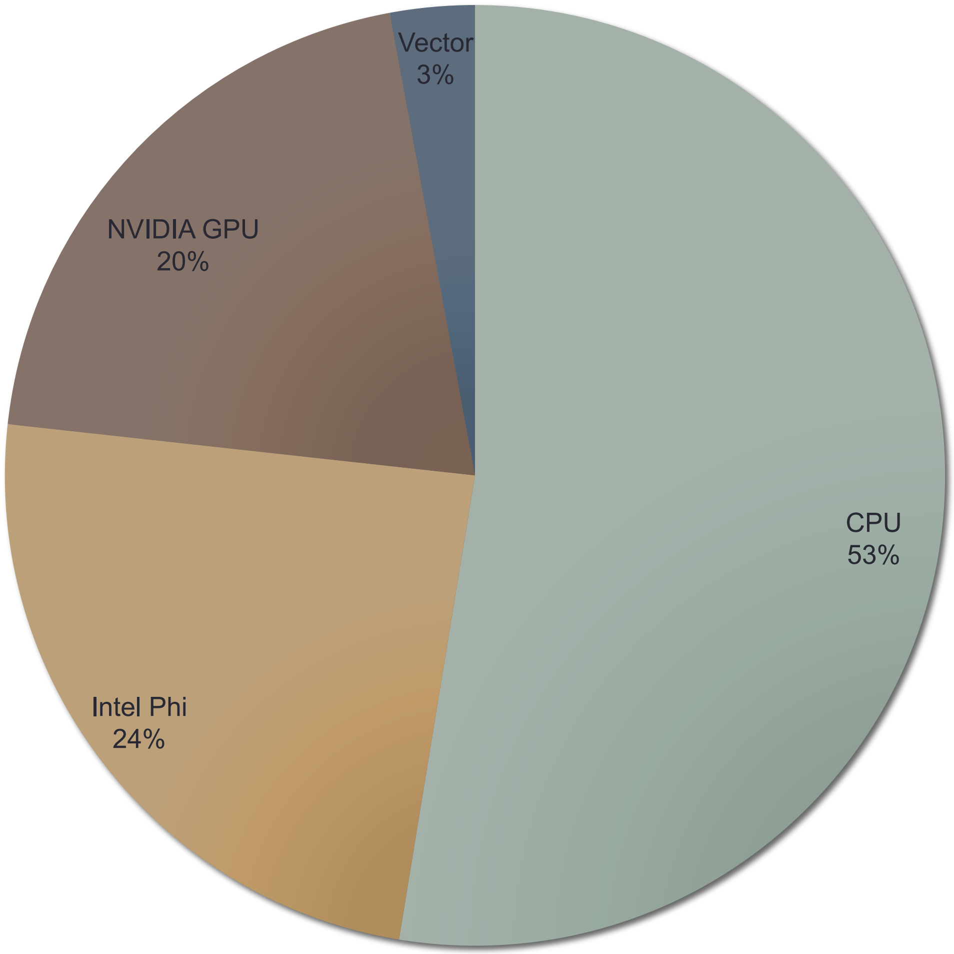
By Network



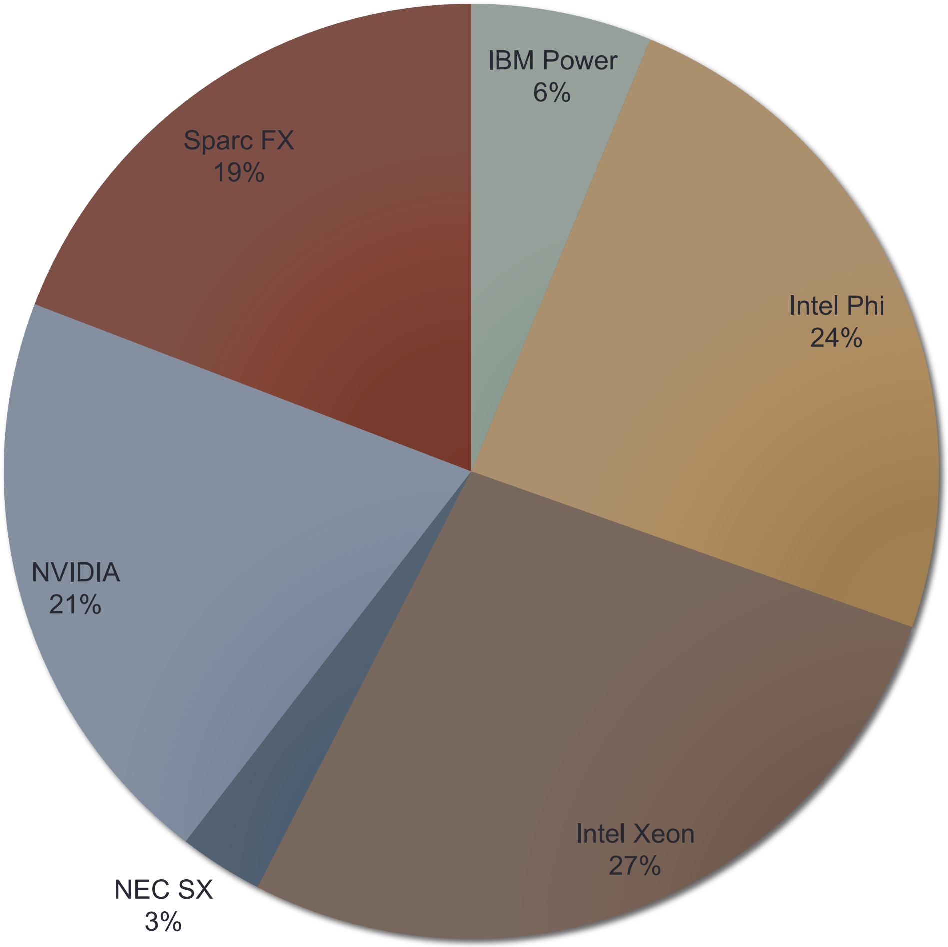
By Network Detail



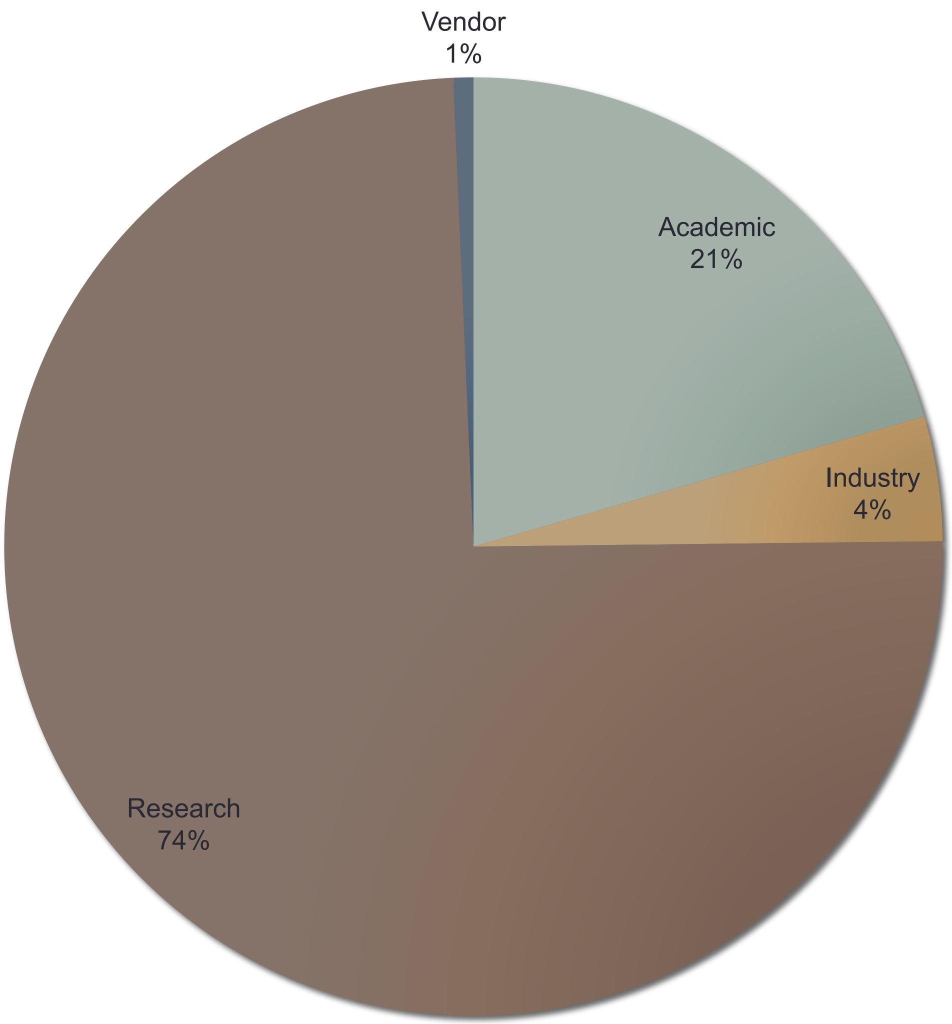
By CPU/Accelerator



By CPU/Accelerator - Details



By Market Segment



By Integrator