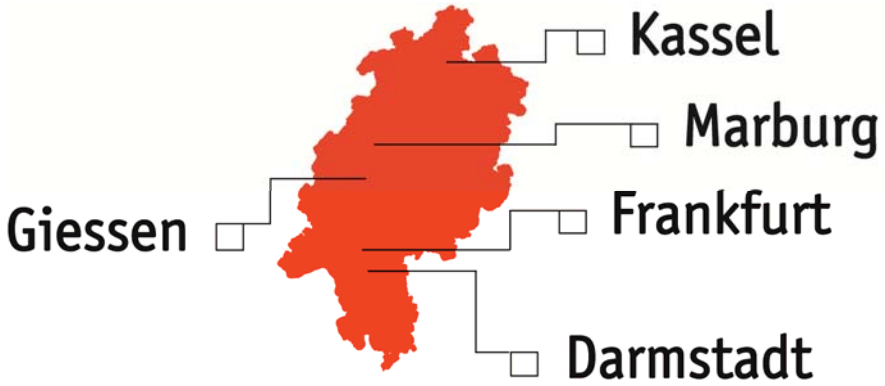
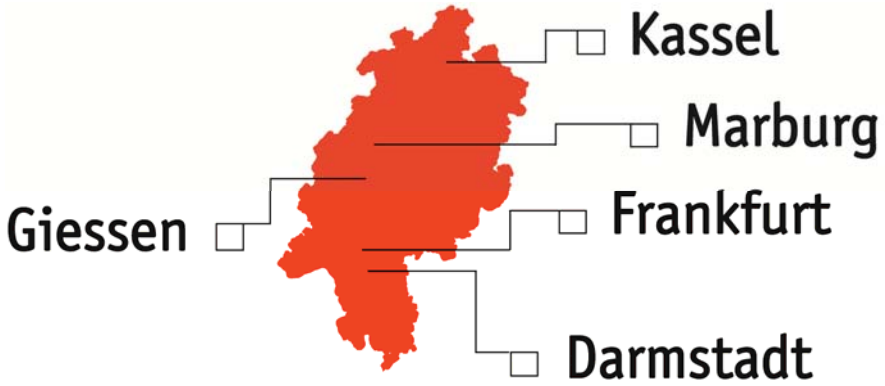
The Hessian Infrastructure for HPC



Hessisches Kompetenzzentrum
für Hochleistungsrechnen

Agenda Introduction HPC-Hessen

- HPC in Germany
- HPC in Hesse
- Brainware for Science
  - Motivation
  - Hessian clusters
  - Benefits for Users
  - User surveys
- Clusters in Darmstadt & Frankfurt



HPC in Germany

Classification

Tier 1: National

- Gauss Center for Supercomputing
  - HLRS, Stuttgart (Höchstleistungsrechenzentrum)
  - JSC, Jülich (Jülich Supercomputing Centre)
  - LRZ, near Munich (Leibniz Rechen Zentrum)

• Tier 2: Regional or Thematic

- Gauss Alliance
  - 8 Members
  - 7 Associates

• Tier 3: Universities & Research Institutions

- ~ 220 Members

More information:

- www.gauss-centre.eu
- www.gauss-allianz.de
- www.hpc-hessen.de

Ref.:→http://www.gauss-allianz.de

HPC in Hesse

Research Institutions using HPC facilities (state aided)

- Universities
- Universities of Applied Science
- GSI
- ...
- → accessible for all Hessian researchers

- Different faculties
- Different computers
- Different (research) interests

- Coordination of interests → Hessischer Beirat
  (Advisory board for HPC, incl. HMWK)

HPC Report is available: http://www.hpc-hessen.de/fileadmin/user_upload/Group_HPC/download/HPC-Report-2015_7MB.pdf

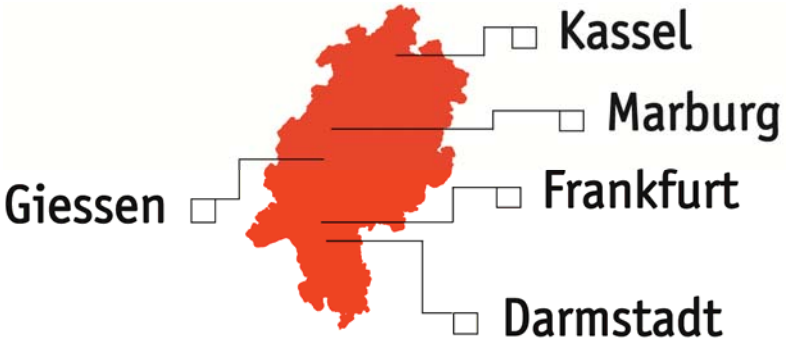
Introducing HPC Hesse (HKHLR)

Management and
Dr. Dörte Sternel, Dr. Alexandra Feith

Directorial Board
Prof. Christian Bischof
Prof. Hans Jürgen Lüdde
Prof. Cristian Heiliger
Prof. Peter Dräxler
Dr. Clemens Helf

Staff Members in:

Darmstadt@hpc-hessen.de
Dr. Sergey Boldyrev
Thomas Opfer
Frankfurt@hpc-hessen.de
Anja Gerbes
David Palao
Giessen@hpc-hessen.de
Dr. Michael Feldmann
Kassel@hpc-hessen.de
Kai Liebel
Sven Siebler
Marburg@hpc-hessen.de
René Sitt
Christian Strack



Brainware for Science

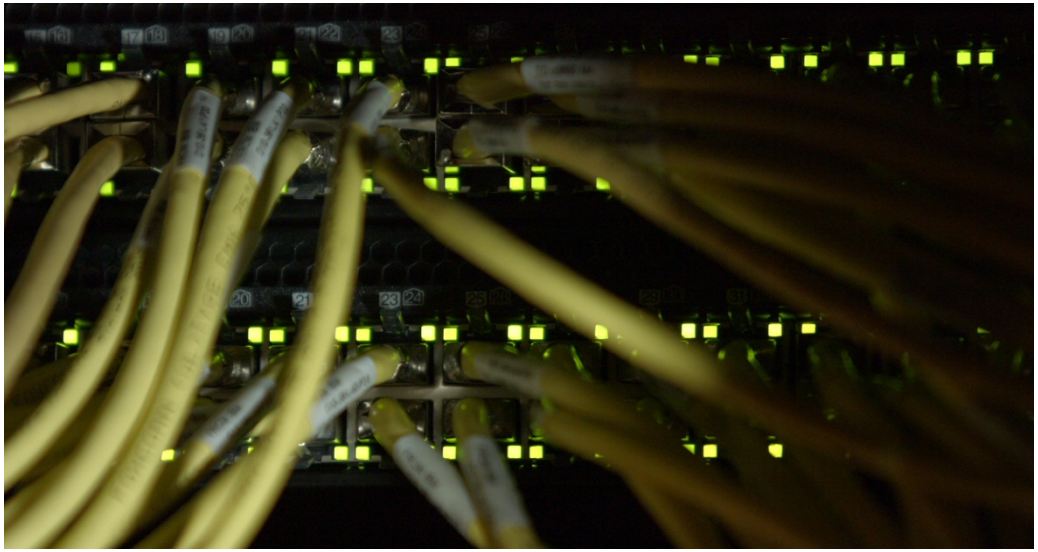
MOTIVATION

Motivation - Clusters

Life-cycle of research software is around 4 x life-cycle of hardware

- several generations of researchers and of computer architectures during the lifetime of a research software

→ **Software maintainability for future generations (researchers and hardware) must be improved**



Motivation - Research

HPC is an important infrastructure for innovative research!

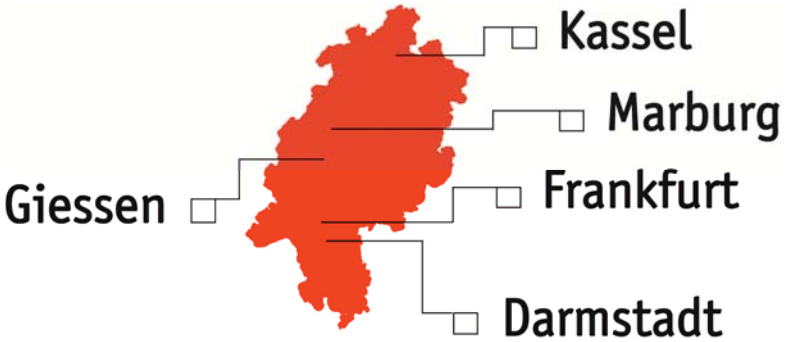
- Most of the HPComputers (world wide) are not efficiently used for relevant problems
  - scientific software often utilizes only a fraction of the reserved compute power while preventing these resources from being used by other jobs

In the last years: Changed view on the **costs of hardware & energy**

Computers have to be used more efficiently!

- in the interest of the universities and funding institutions
- in the interest of research

→ **brainware for science → HPC Hessen**



Brainware for Science

HESSIAN CLUSTERS

Hessian Clusters

	Lichtenberg Cluster TU Darmstadt				LOEWE-CSC Cluster Goethe University Frankfurt				Linux Cluster University of Kassel		MaRC2 Philipps University Marburg		Skylla Cluster JLU Giessen	
Typical Node Parameters	"Phase 1"		"Phase 2"		Nodetype A		Nodetype B							
	Typical	Max	Typical	Max					Typical	Max	Typical	Max	Typical	Max
Cores (sockets x cores/socket)	2x8	8x8	2x12	4x15	2x12		2x10		2x16	2x16	4x16	4x16	2x6	4x8
Memory (RAM)	32 GB	1 TB	64 GB	1 TB	64 GB		128 GB		32 GB	256 GB	256 GB	256 GB	32 GB	64 GB
CPU Type (typical)	Intel Xeon E5-2670		Intel Xeon E5-2680v3		AMD Opteron 6172 (2.1GHz)		Intel Xeon E5-2670v2(2.5GHz)		AMD Opteron (Interlagos, 2.3GHz)		AMD Opteron (Interlagos/Abu Dhabi, 2.3 GHz)		AMD Opteron 2431 (Istanbul, 2.4GHz)	
Flops/Core (DP, theor. peak)	20.8 Gflops		20.0 Gflops		8.6 Gflops		20.8 Gflops		9.2 Gflops		9.2 Gflops		9.6 Gflops	
MPI Communication (pt2pt)	Intranode	Internode	Intranode	Internode	Intranode	Internode	Intranode	Internode	Intranode	Internode	Intranode	Internode	Intranode	Internode
- Bandwidth	4.6 GB/s	4.7 GB/s	4.1 GB/s	6.5 GB/s	-	-	-	-	4.2 GB/s	2.7 GB/s	2.5 GB/s	1.9 GB/s	2.5 GB/s	1.6 GB/s
- Latency (64 bytes)	0.73 µs	1.34 µs	0.71 µs	1.29 µs	-	-	-	-	0.8 µs	2.5 µs	1.0 µs	3.8 µs	0.8 µs	3.8 µs
Memory Bandwidth (triad)	40 GB/s per socket		60 GB/s per socket		38.2 GB/s		48.6 GB/s		33.0 GB/s		71.5 GB/s		15.5 GB/s	
Accelerators	2x NVIDIA K20Xm à 1.3 Tflops, 6 GB (some nodes)		2x NVIDIA K40m à 1.4 Tflops, 12 GB (some nodes)		AMD Radeon HD 5800 à 0.75 TFLOPS, 1GB		2xAMD FirePro S10000 à 1.5 TFLOPS, 12 GB		-		-		-	
Local Temporary Storage	100 GB				1.4 TB				1.8 TB		2 TB		400 GB	
Node Allocation	shared and exclusive				exclusive				shared and exclusive		shared		shared and exclusive	
Global Cluster Parameters														
Processors (CPU, DP, peak)	579 Tflops				226 TFLOPS				40.7 TFLOPS		56.5 TFLOPS		9.3 TFLOPS	
Accelerators (GPU, DP, peak)	174 Tflops				597 TFLOPS				-		-		-	
Compute Cores (CPU)	22,208				18,064				4,452		6,144		992	
Permanent Storage	0.5 PB				1.5 PB				40 TB		33 TB		8 TB	
Scratch Storage	1.5 PB				0.76 PB				5 TB		33 TB		24 TB	
Jobmanager	LSF				Slurm				Slurm		Grid Engine		Grid Engine	
Other Job Constraints	runtime: 24h, max. 14d				runtime: max. 30d				mem: 8 GB/core		runtime: 10d max. 1000 jobs/user		runtime: 5d/unlimited max. 128 (192) cores/user	

Hessian Clusters

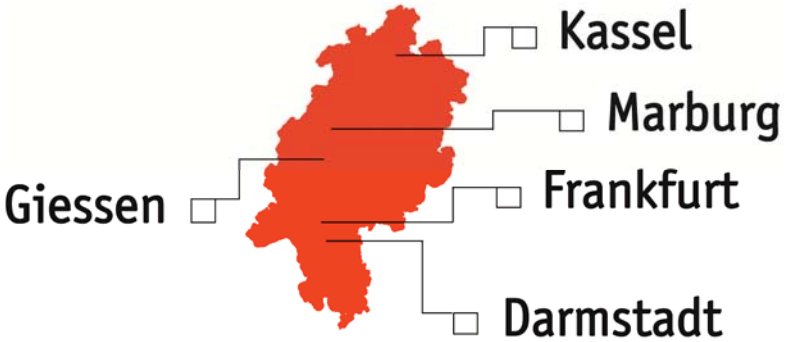
	Lichtenberg Cluster TU Darmstadt				LOEWE-CSC Cluster Goethe University Frankfurt			
Typical Node Parameters	"Phase 1"		"Phase 2"		Nodetype A		Nodetype B	
	Typical	Max	Typical	Max				
Cores (sockets x cores/socket)	2x8	8x8	2x12	4x15	2x12		2x10	
Memory (RAM)	32 GB	1 TB	64 GB	1 TB	64 GB		128 GB	
CPU Type (typical)	Intel Xeon E5-2670		Intel Xeon E5-2680v3		AMD Opteron 6172 (2.1GHz)		Intel Xeon E5-2670v2(2.5GHz)	
Flops/Core (DP, theor. peak)	20.8 Gflops		20.0 Gflops		8.6 Gflops		29.8 Gflops	
MPI Communication (pt2pt)	Intranode	Internode	Intranode	Internode	Intranode	Internode	Intranode	Internode
- Bandwidth	4.6 GB/s	4.7 GB/s	4.1 GB/s	6.5 GB/s	-	-	-	-
- Latency (64 bytes)	0.73 µs	1.34 µs	0.71 µs	1.29 µs	-	-	-	-
Memory Bandwidth (triad)	40 GB/s per socket		60 GB/s per socket		38.2 GB/s		48.6 GB/s	
Accelerators	2x NVIDIA K20Xm à 1.3 Tflops, 6 GB (some nodes)		2x NVIDIA K40m à 1.4 Tflops, 12 GB (some nodes)		AMD Radeon HD 5800 à 0.75 TFLOPS, 1GB		2xAMD FirePro S10000 à 1.5 TFLOPS, 12 GB	
Local Temporary Storage	100 GB				1.4 TB			
Node Allocation	shared and exclusive				exclusive			

Global Cluster Parameters		
Processors (CPU, DP, peak)	579 Tflops	226 TFLOPS
Accelerators (GPU, DP, peak)	174 Tflops	597 TFLOPS
Compute Cores (CPU)	22,208	18,064
Permanent Storage	0.5 PB	1.5 PB
Scratch Storage	1.5 PB	0.76 PB
Jobmanager	LSF	Slurm
Other Job Constraints	runtime: 24h, max. 14d	runtime: max. 30d

Hessian Clusters

	Linux Cluster University of Kassel		MaRC2 Philipps University Marburg		Skylla Cluster JLU Giessen	
Typical Node Parameters						
	Typical	Max	Typical	Max	Typical	Max
Cores (sockets x cores/socket)	2x16	2x16	4x16	4x16	2x6	4x8
Memory (RAM)	32 GB	256 GB	256 GB	256 GB	32 GB	64 GB
CPU Type (typical)	AMD Opteron (Interlagos, 2.3GHz)		AMD Opteron (Interlagos/Abu Dhabi, 2.3 GHz)		AMD Opteron 2431 (Istanbul, 2.4GHz)	
Flops/Core (DP, theor. peak)	9.2 Gflops		9.2 Gflops		9.6 Gflops	
MPI Communication (pt2pt)	Intranode	Internode	Intranode	Internode	Intranode	Internode
- Bandwidth	4.2 GB/s	2.7 GB/s	2.5 GB/s	1.9 GB/s	2.5 GB/s	1.6 GB/s
- Latency (64 bytes)	0.8 µs	2.5 µs	1.0 µs	3.8 µs	0.8 µs	3.8 µs
Memory Bandwidth (triad)	33.0 GB/s		71.5 GB/s		15.5 GB/s	
Accelerators	-		-		-	
Local Temporary Storage	1.8 TB		2 TB		400 GB	
Node Allocation	shared and exclusive		shared		shared and exclusive	

Global Cluster Parameters			
Processors (CPU, DP, peak)	40.7 TFLOPS	56.5 TFLOPS	9.3 TFLOPS
Accelerators (GPU, DP, peak)	-	-	-
Compute Cores (CPU)	4,452	6,144	992
Permanent Storage	40 TB	33 TB	8 TB
Scratch Storage	5 TB	33 TB	24 TB
Jobmanager	Slurm	Grid Engine	Grid Engine
Other Job Constraints	mem: 8 GB/core	runtime: 10d max. 1000 jobs/user	runtime: 5d/unlimited max. 128 (192) cores/user



Brainware for Science

BENEFITS FOR USERS

Benefits for users

- A Hessen-wide concept for user support
- Individual User Support
  - Contact persons at each location
  - Regular introduction courses for new users
  - HiPerCH (High Performance Computing in Hesse) workshop series for advanced users and beginners
  - Further education and support through user meetings

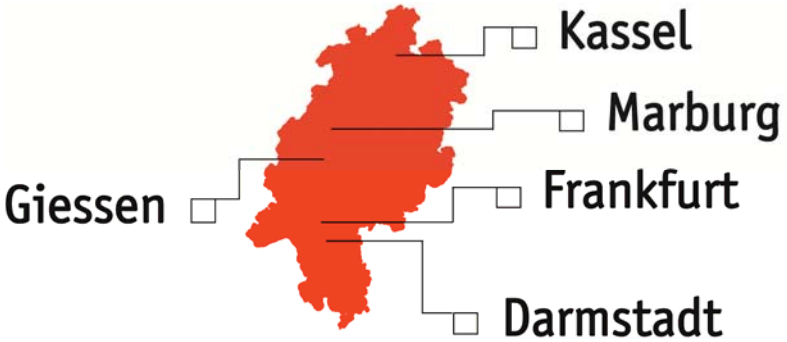
→ Hessen wide user network
→ understand users‘ needs
→ effective usage of all Hessian clusters

User support

- Performance Analysis
  - Vampir licence for all Hessian clusters (funded by HMWK)
  - Together with the user, analysis for example in Darmstadt (FASTEST), Marburg, Giessen (plasmapic)
- Support for porting or installation of for example openFOAM, VASP, Gaussian, COMSOL, different in-house-codes, special libraries
- Support for debugging

Monitoring

- Central monitoring of the computer capacity
  - Discovery of bottlenecks in utilization of cluster resources
- Central monitoring of user support
  - Collection of FAQs, and topics for workshops
- Annual user evaluation
  - Close to users‘ needs, HPC-usage, users‘ competences and demands
- HiPerCH-Workshops: individual course evaluation
  - Anonymous feedback

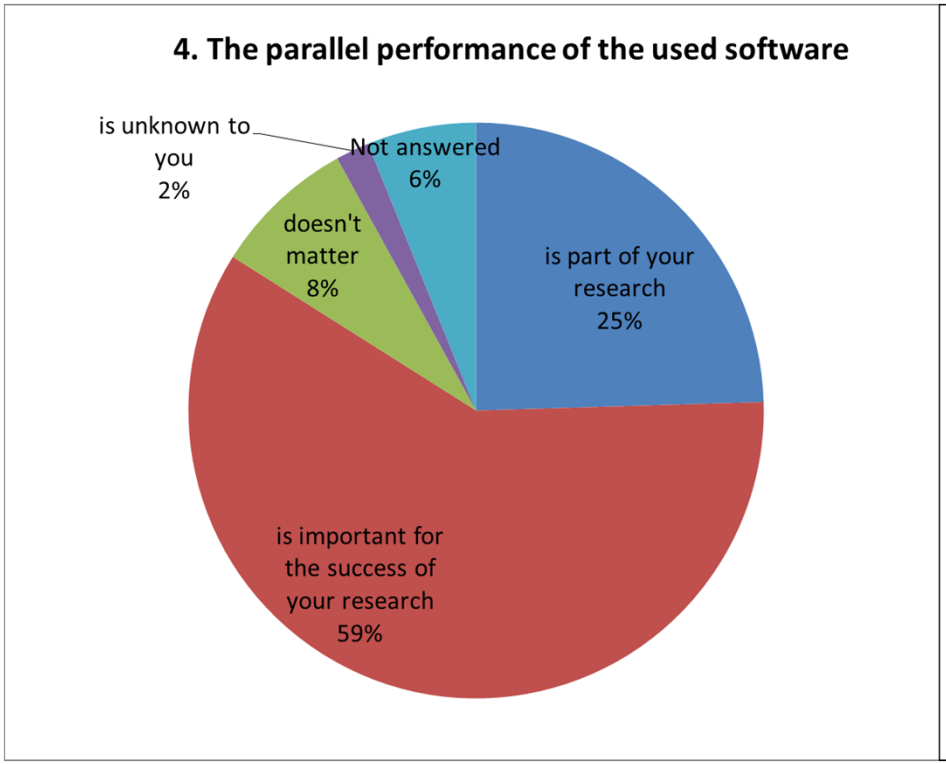


Brainware for Science

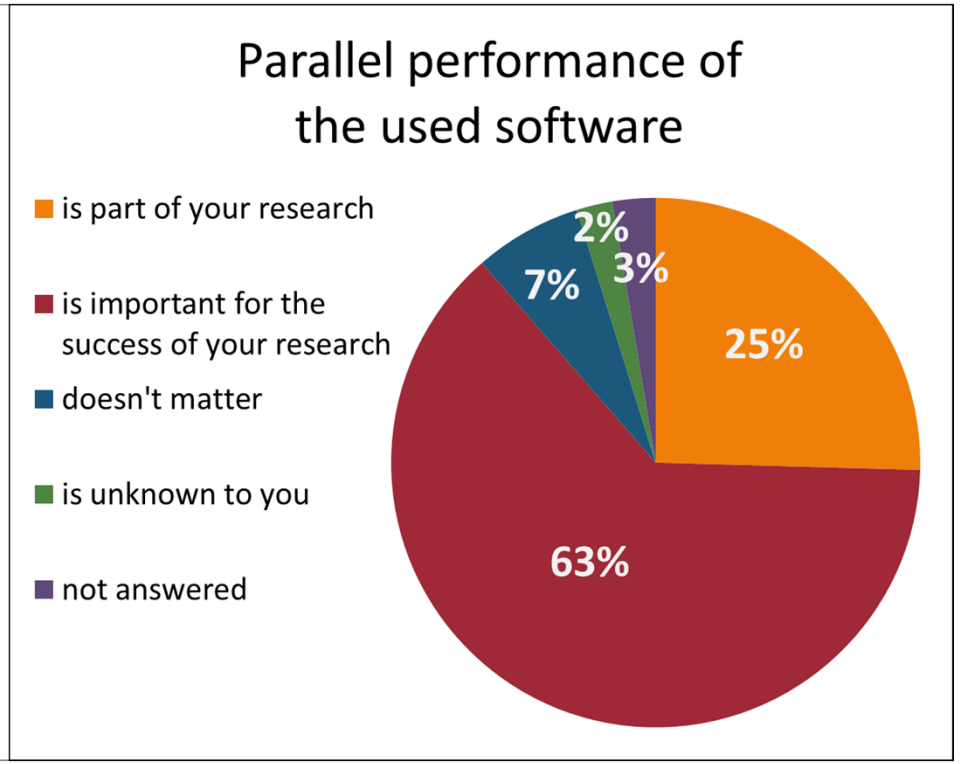
USER SURVEYS

Online Survey, 2014: n=212; 2015: n=228

2014

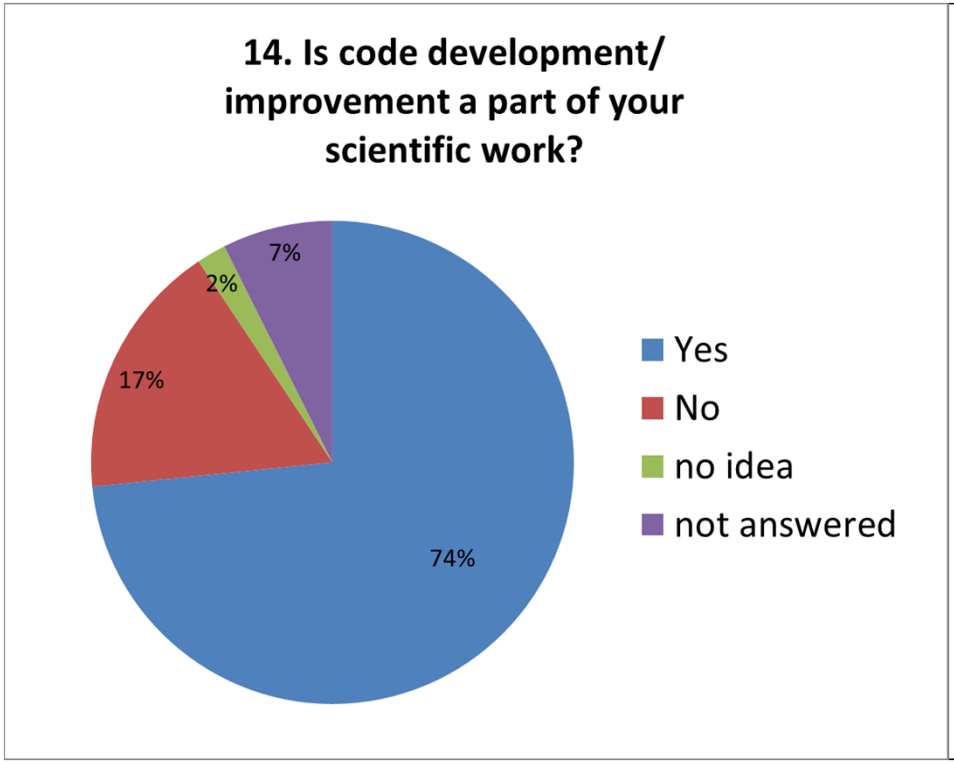


2015

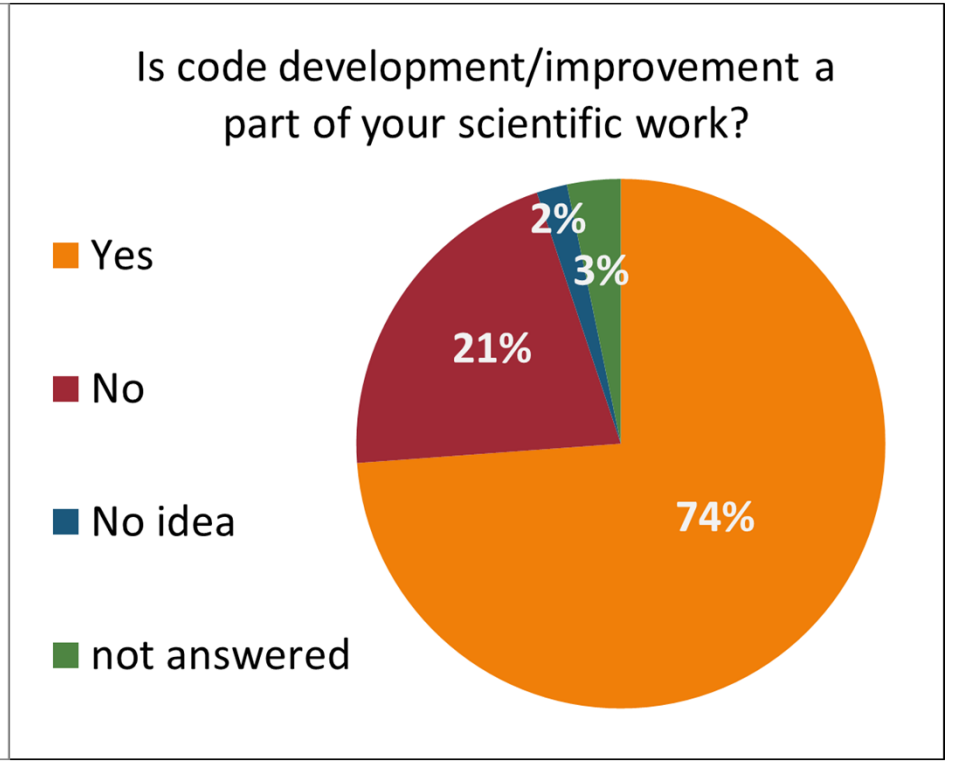


Online Survey

2014

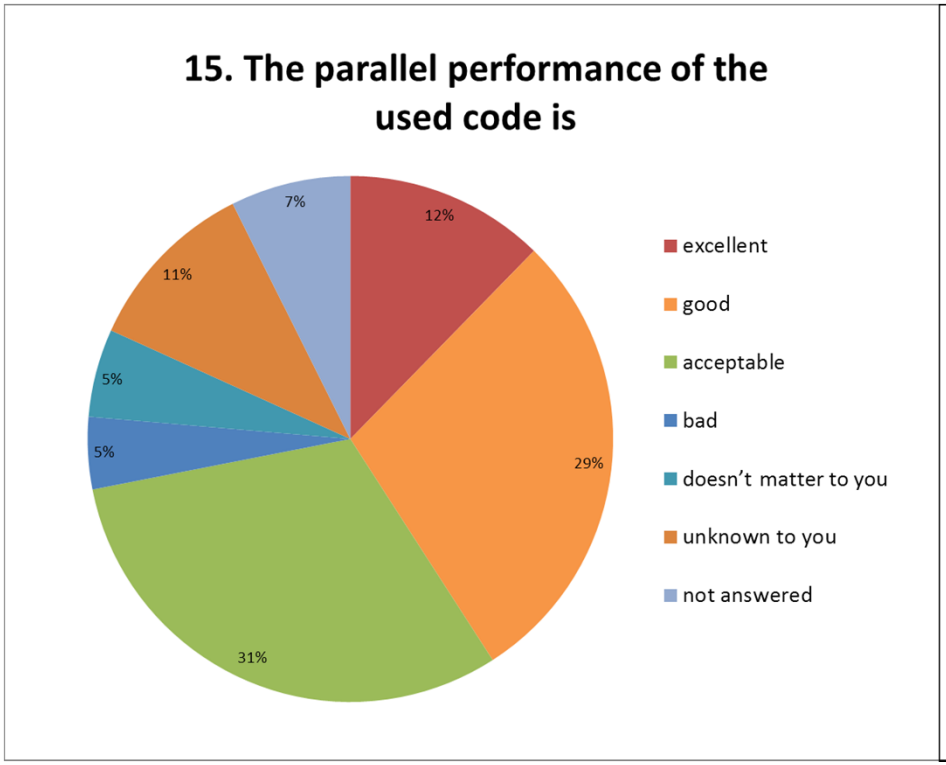


2015

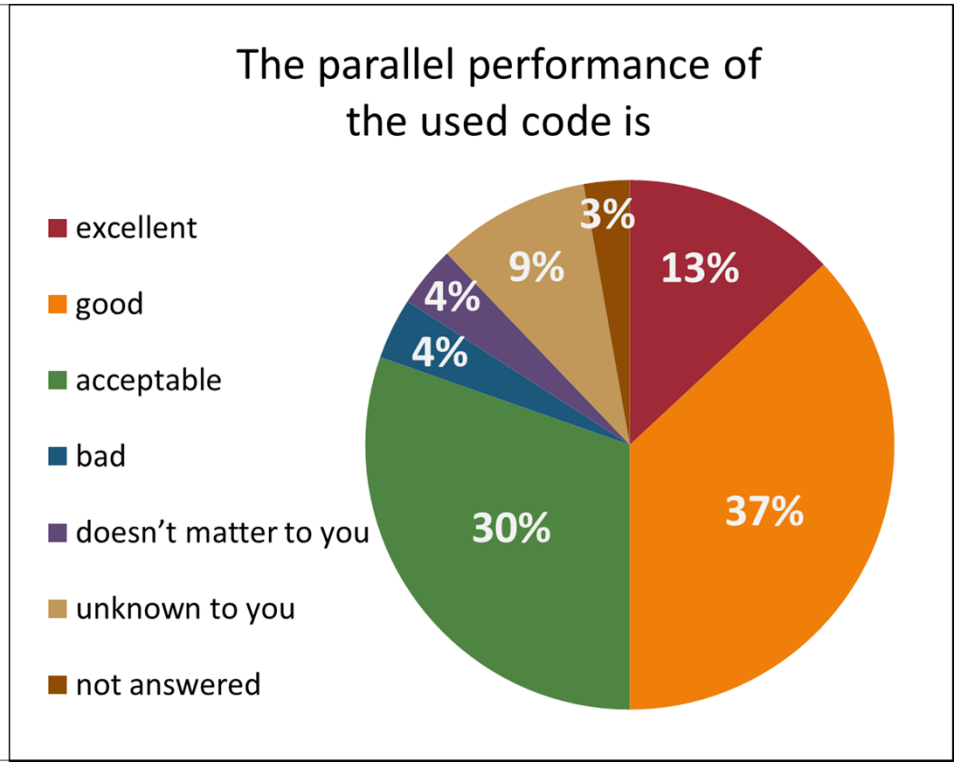


Online Survey

2014

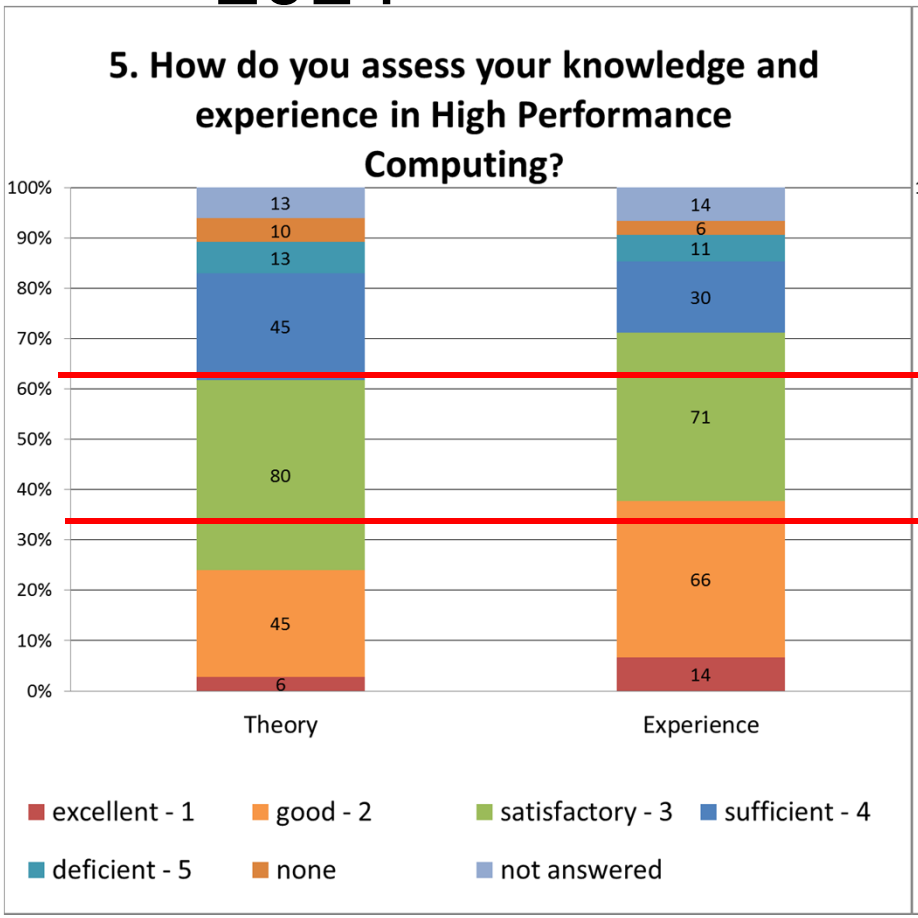


2015

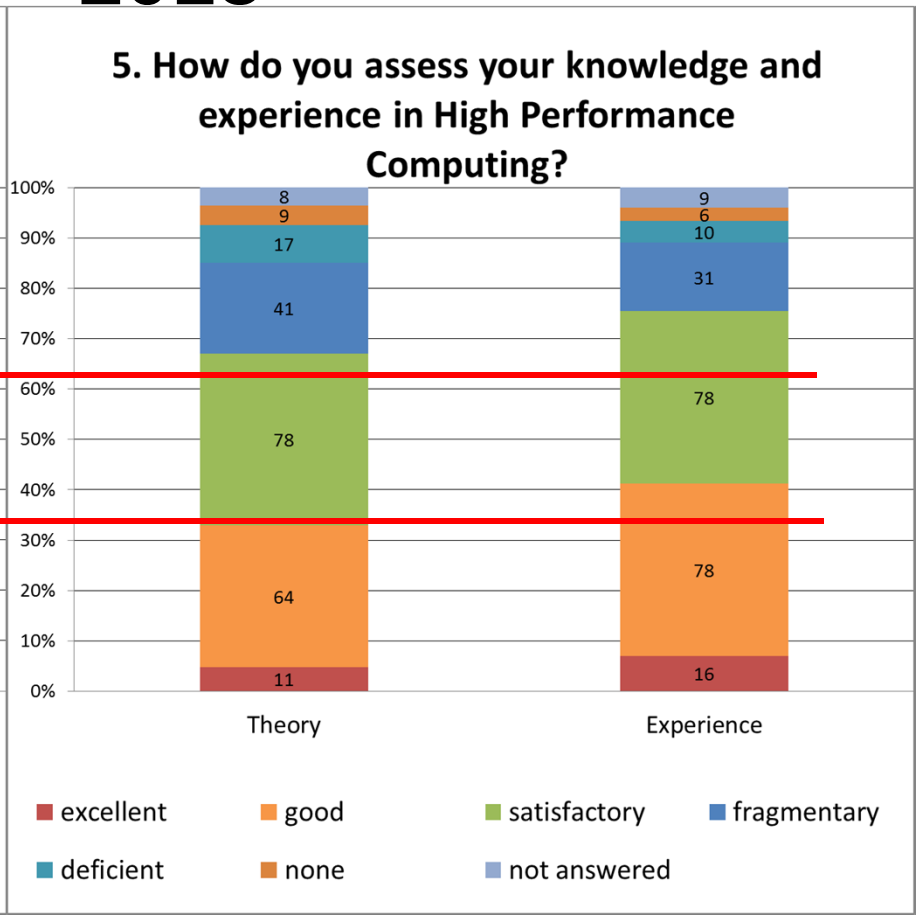


Online Survey

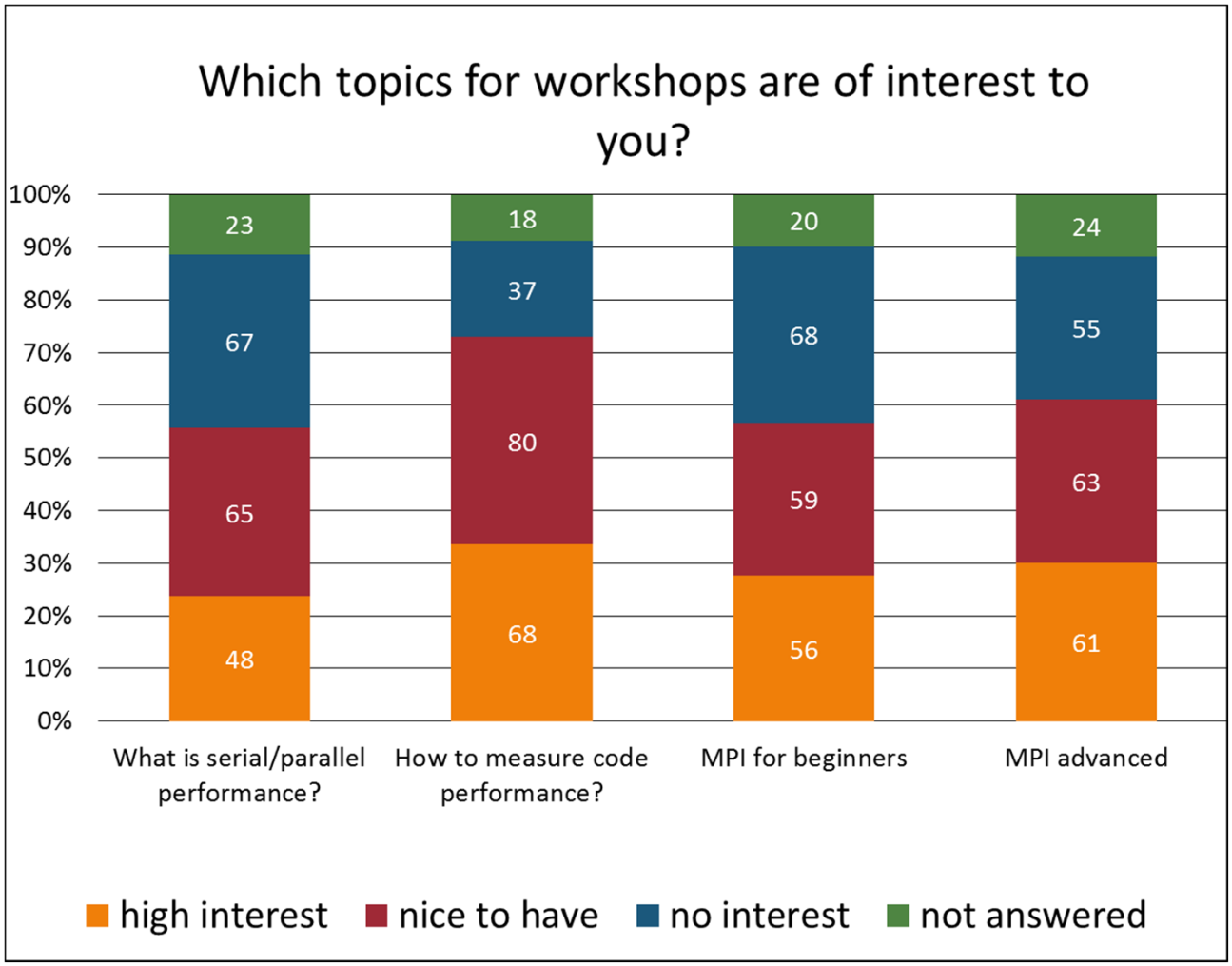
2014



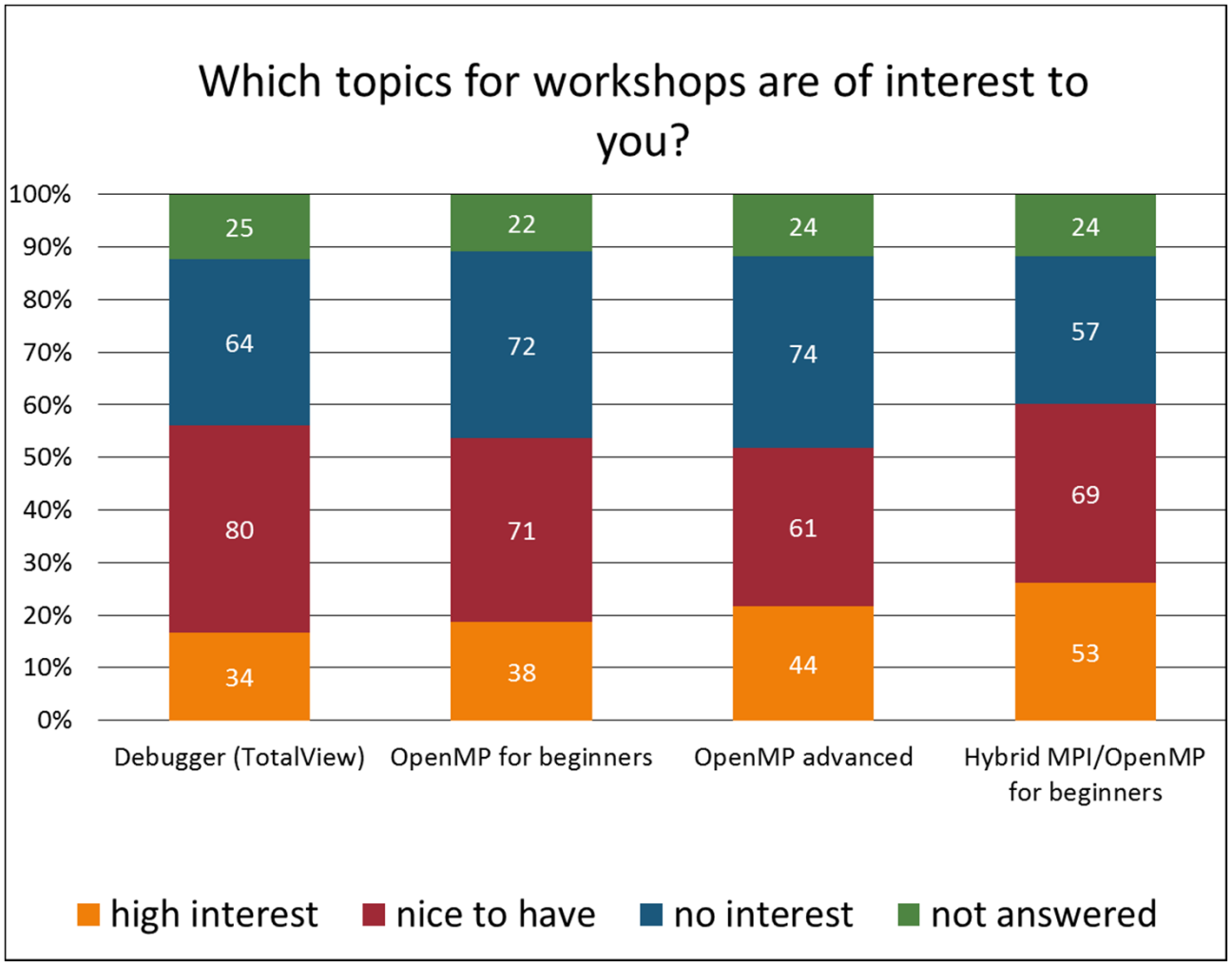
2015



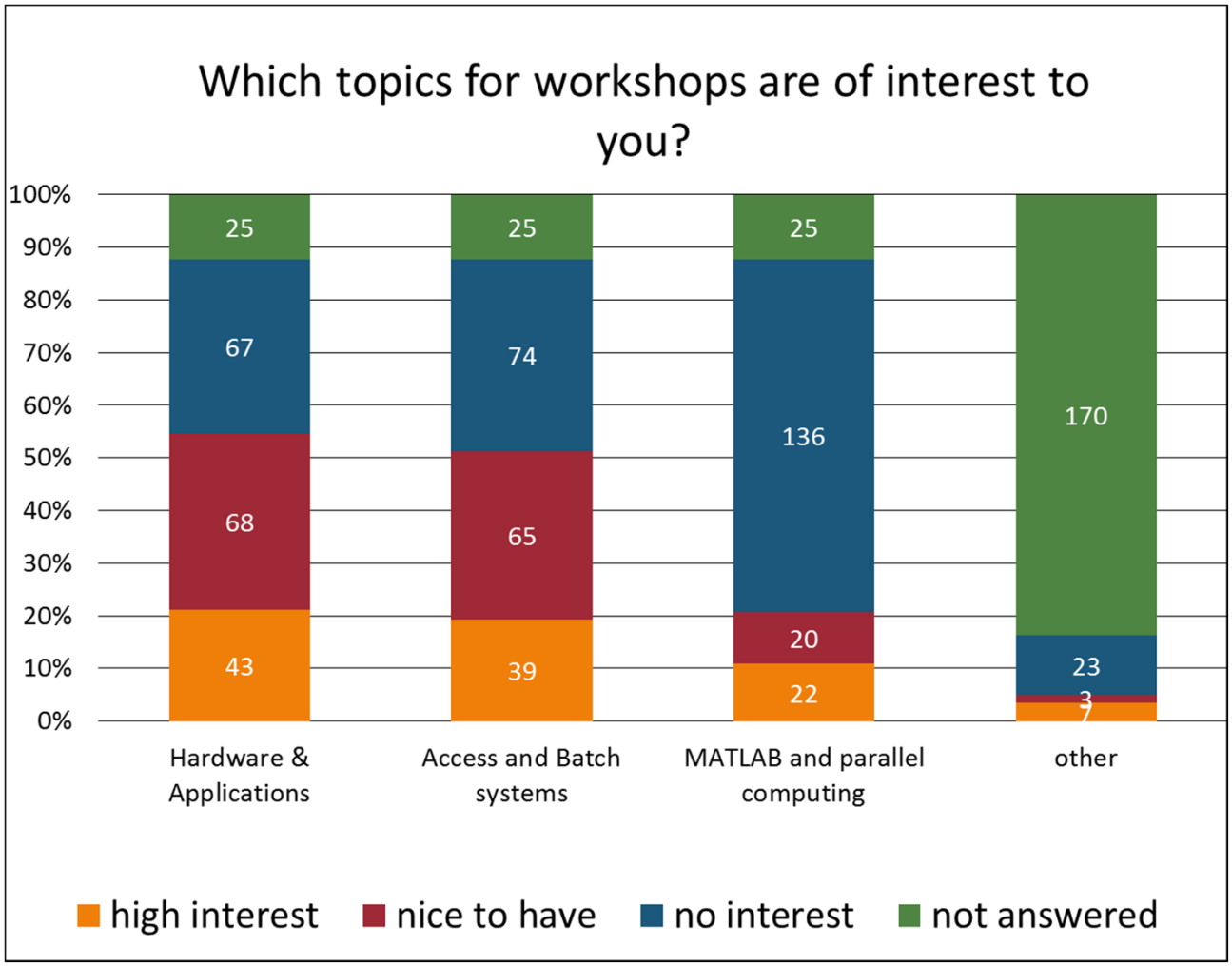
Online Survey 2015



Online Survey 2015



Online Survey 2015



Online Survey 2015