WG 4.1
HPC R&D cartography

H. Huber, Chair

R. Brunino, Vice Chair (Speaker)

WG4.1 Experts

Leif Nordlund	AMD	SE	Business Development Manager
Jean-Pierre Panziera	Bull	FR	Director of Performance Engineering
Wilfried Oed	Cray	DE	HPC Engineer, EMEA
Luigi Brochard	IBM	FR	Distinguished HPC Engineer
Andrey Semin	Intel	RUS	HPC Technology Manager, EMEA
Frank Baetke	HP	DE	Worldwide Director of High Performance Computing
Michael Kagan	Mellanox	IL	CTO of Mellanox Technologies
Dev Tyagi	Supermicro	UK	General Manager at Supermicro, UK
Giampetro Tecchiolli	Eurotech	IT	CEO/CTO of Eurotech
Salvatore Rinaudo	ST Microelectronics	IT	ST Microelectronics
Rüdiger Wolff	SGI	DE	Principal Systems Engineer & Systems Engineer Manager, EMEA
Dmitry Tkachev	T-Platforms	RUS	Chief Architect
Pierre Lagier	Fujitsu	FR	Technical Director
Ulrich Brüning	ExTOLL	DE	Computer Architecture Research Manager
Aad van der Steen	NCF	NL	HPC architectures & benchmarking
Alex Ramirez	BSC	ES	Computer Architecture Research Manager
Dominik Ulmer	CSCS	CH	HPC Center
Jean Gonnord	CEA	FR	Numerical Simulation & Computer Sciences

Questions for the experts

- **What are the major challenges/issues in HPC in the next 5 to 10 years?**
- **What efforts will your institution or company undertake to address these challenges/issues?**
  - **a) If applicable, please mention all R&D on promising future multi-Peta to Exascale hardware technologies such as hybrid many-core CPUs, novel memory technologies (3D-stacked memory, memristor, Spin Torque Transfer Magnetic RAM, graphene based memory, etc.), novel background storage technologies, photonic interconnects, highly cooling efficient system packaging technologies, etc.**
    b) If applicable, please mention all R&D on promising future multi-Peta to Exascale software technologies such as highly scalable programming languages, fault tolerant parallel programming environments, parallel file systems with end-to-end data integrity, highly scalable and energy efficient system management software, hierarchical storage solutions, OS with coherent inter-node scheduling mechanisms, etc.
    c) If applicable, please mention all R&D on promising highly scalable numerical algorithm for multi-Peta to Exascale applications

Preliminary results on international HPC R&D(*) cartography

PROCESSING ELEMENTS																	
System on a Chip			X			X	X	X		X		X	X	X	X	X	10
Heterogeneous many-core			X			X	X	X		X			X	X	X	X	9
Homogeneous many-core			X				X			X	X		X			X	6
In-interconnect processing	X	X		X	X		X			X							6
Company	SGI	Mellanox	Intel	CRAY	BULL	SuperMicro	IBM	STMicroelectronic	Eurotech	AMD	Fujitsu	HP	T-Platforms	Tilera	nVIDIA	ARM	Counts

(*) An “X“ in a field of the table means the vendor is performing active research and development in this field

Preliminary results on international HPC R&D(*) cartography

MEMORY TECHNOLOGY																	
3D Stacking			X				X	X		X	X	X		X	X	X	9
Novel Memory Technologies	X		X				X				X	X					5
Company	SGI	Mellanox	Intel	CRAY	BULL	SuperMicro	IBM	STMicroelectronic	Eurotech	AMD	Fujitsu	HP	T-Platforms	Tilera	nVIDIA	ARM	Counts

(*) An “X“ in a field of the table means the vendor is performing active research and development in this field

Preliminary results on international HPC R&D(*) cartography

POWER DISTRIBUTION AND MANAGEMENT																	
Power distribution	X		X	X	X		X		X		X	X	X				9
Energy-aware Software	X		X		X		X		X		X	X	X				8
Company	SGI	Mellanox	Intel	CRAY	BULL	SuperMicro	IBM	STMicroelectronic	Eurotech	AMD	Fujitsu	HP	T-Platforms	Tilera	nVIDIA	ARM	Counts

(*) An “X“ in a field of the table means the vendor is performing active research and development in this field

Preliminary results on international HPC R&D(*) cartography

RESILIENCY AND FAULT TOLERANCE																	
HW level	X	X	X	X	X	X	X		X	X	X	X	X		X	X	14
SW level	X	X	X	X			X		X			X	X				8
Company	SGI	Mellanox	Intel	CRAY	BULL	SuperMicro	IBM	STMicroelectronic	Eurotech	AMD	Fujitsu	HP	T-Platforms	Tilera	nVIDIA	ARM	Counts

(*) An “X“ in a field of the table means the vendor is performing active research and development in this field

Preliminary results on international HPC R&D(*) cartography

MULTI-PETA TO EXASCALE SYSTEM AND SOFTWARE STACK																	
Runtime Systems with(**)/(without) fault tolerance	X		X	X	X		X		X		X	X	X				9
Compiler, Debugger, Tools			X	X			X			X	X		X		X		7
Programming Languages			X	X			X				X				X		5
Company	SGI	Mellanox	Intel	CRAY	BULL	SuperMicro	IBM	STMicroelectronic	Eurotech	AMD	Fujitsu	HP	T-Platforms	Tilera	nVIDIA	ARM	Counts

(*) An “X“ in a field of the table means the vendor is performing active research and development in this field
(**) A fault tolerant runtime system will be capable to run a user job to completion even in the case of a compute node failure

Preliminary results on international HPC R&D(^) cartography

Company	SGI	Mellanox	Intel	CRAY	BULL	SuperMicro	IBM	STMicroelectronic	Eurotech	AMD	Fujitsu	HP	T-Platforms	Tilera	nVIDIA	ARM	Counts
Power efficiency	X		X	X	X	X	X		X	X		X	X			X	11
Network Interconnect & Routing Algorithms (**)	X	X	X	X	X		X		X	X	X	X					10 +1
Silicon Photonics (**)	X		X			X	X			X	X	X	X				8 +1
Novel system packaging and cooling	X		X	X	X		X		X	X			X				8
Numerical libraries			X	X			X			X	X				X		6
Co-design	X		X	X	X		X			X	X						7
Novel background storage technologies	X		X				X					X					4
Parallel File Systems (*)					X		X				X						3
New Semi Conductor materials (nanocarbon)			X				X				X						3
(*) Wamcloud, Xyratech				(**) ExTOLL													

(^) An "X" in a field of the table means the vendor is performing active research and development in this field

Summary and Conclusions

- Processors
  - ARM
    - Hybrid multi-core
    - SoCs
  - AMD
    - Fusion
    - Opteron
    - Firestream (GPGPU)
  - IBM
    - Power
    - PowerPC
  - Intel
    - Xeon
    - MIC
    - Atom
    - Itanium
  - nVIDIA
    - Tesla (GPGPU accelerated clusters)
    - Tegra (hybrid multi-core SoC)
  - Fujitsu
    - Sparc64 VIIIfx (8 core processor)
    - Sparc64 (16 core processor, …, (8 years development program))
  - Godson (China)
    - Godson-3B (8-core MIPS, 128 GF@1.05 GHz 65nm, 40W)
    - Godson-3C (16-core MIPS, 512 GF@2 GHz, 28 nm, ~40W, ~ 12 GF/W)

Summary and Conclusions

- Memory
  - HP: Memristor
    - SSDs: 2014-2015
    - DRAM & NV caches: 2015-2016
    - SRAM: 2017-2018
  - IBM
    - Phase Change Memory (PCM)

- Interconnect

  CRAY
  ExTOLL
  HP
  IBM
  Intel
  SGI
  T-Platforms

Summary and Conclusions

- There is a clear trend towards more and more processing cores
  - More and more hierarchies will have to be mastered by application codes
    - Memories
      - Inter- and intra-node, various levels of caches and RAM (CPU & GPU)
    - Functional units
      - Big cores, small cores, vector and specialized units
- The Exascale system architecture is still an open debate (processor, memory, interconnect , …)
- MPI + (OpenMP, Threads, PGAS, ..) seems to be the programming model for the present decade
  - But big uncertainty, what programming model should be used for a long term development?
  - Old programs not targeting performance will still be usable in the ExaFlop/s performance decade
    - Running on general purpose CPUs with a low fraction of peak performance!