More About the Future

Computers and Applications

More About the Future

1. FPGA

(Field Programmable Gate Array)

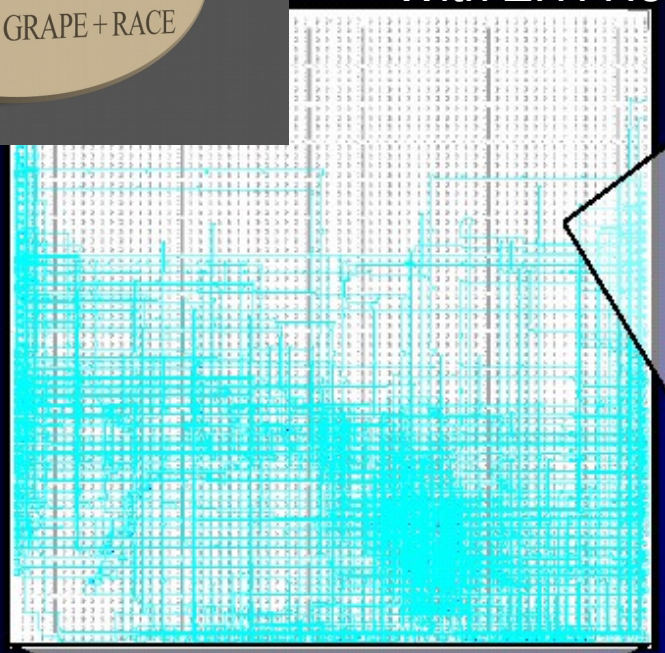
GRACE

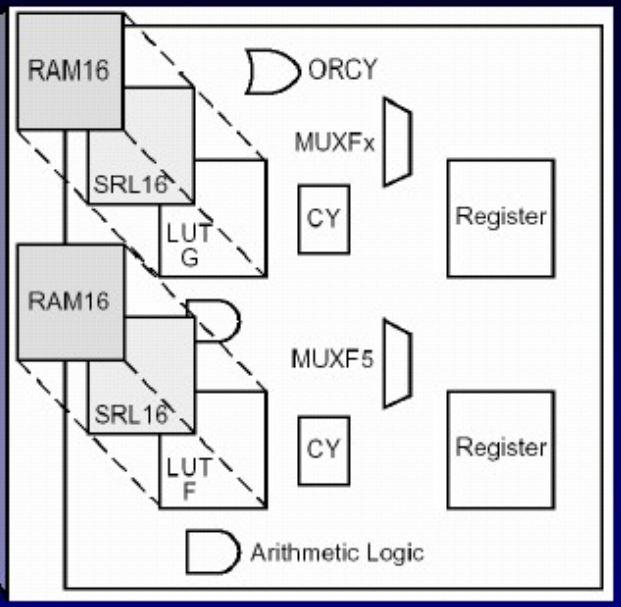
GRACE = GRAPE + RACE

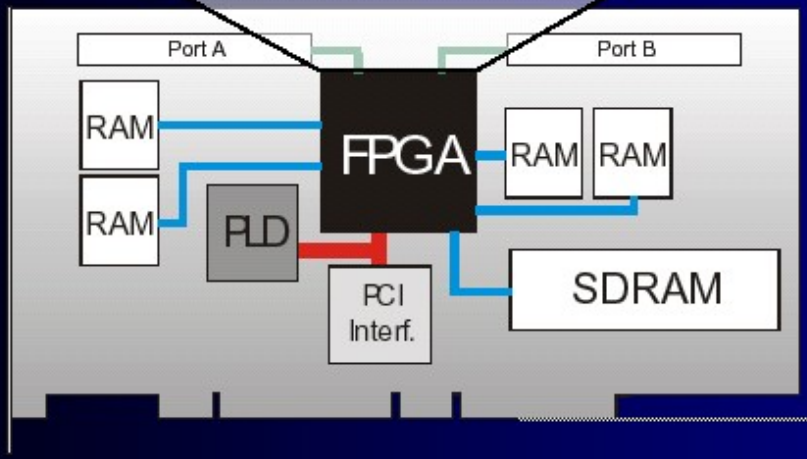
FPGA-Plattform MPRACE

2005 - 2009

With ZITI Heidelberg (former TI Mannheim), Prof. Männer







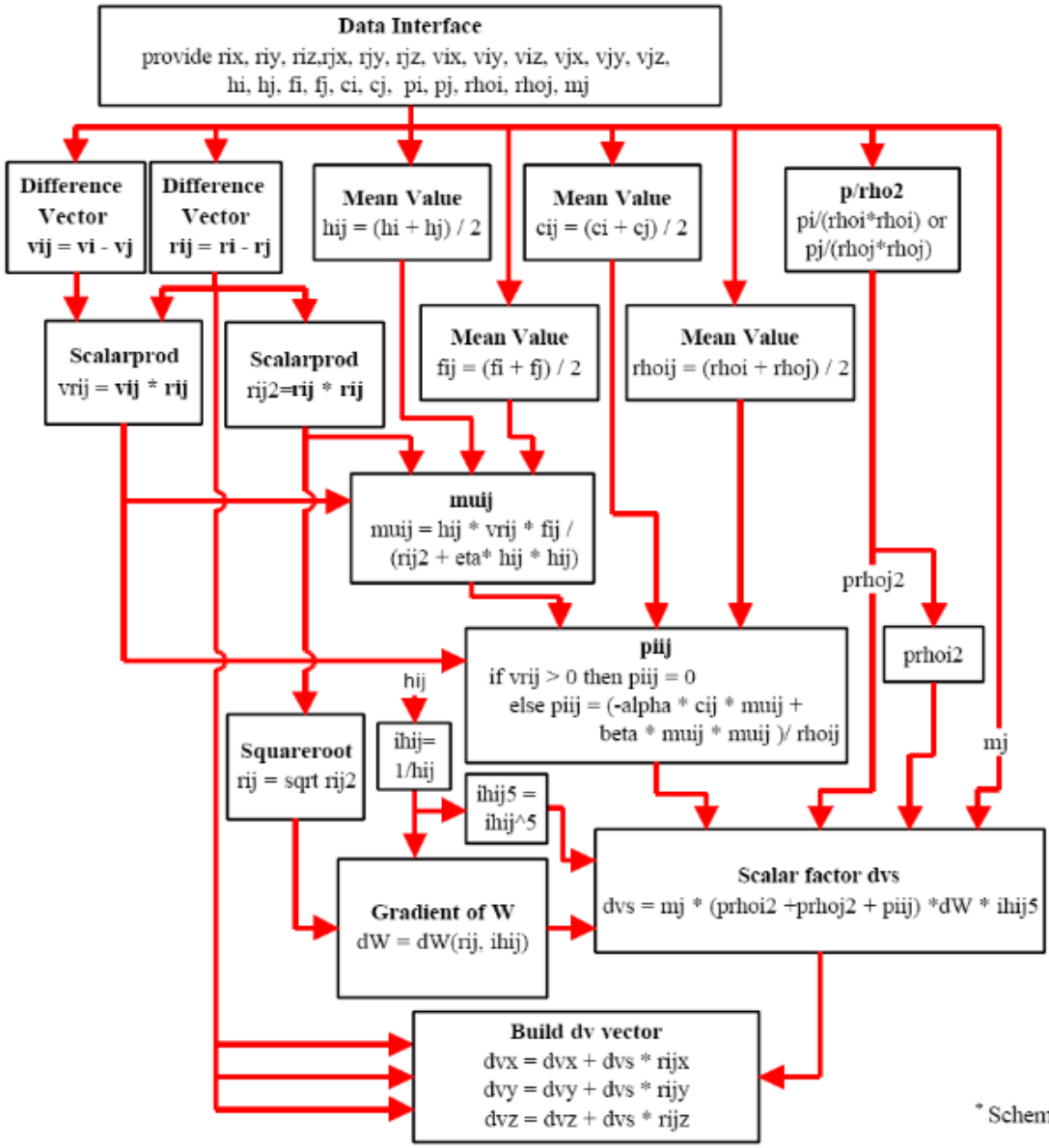
lienhart@ti.uni-mannheim.de

GRACE

GRACE = GRAPE + RACE

FPGA...

Pressure force pipeline:



* Scheme doesn't show energy term

Reconfigurable Computing (FPGA) at ZITI, Heidelberg

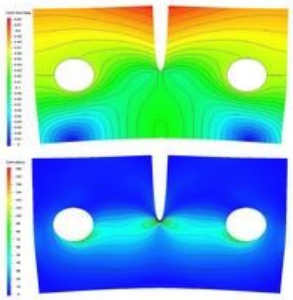
Chair on Application Specific Computing

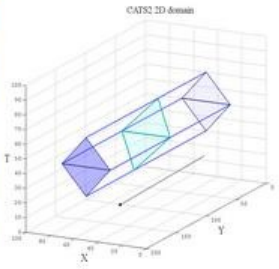
https://asc.ziti.uni-heidelberg.de/en/start

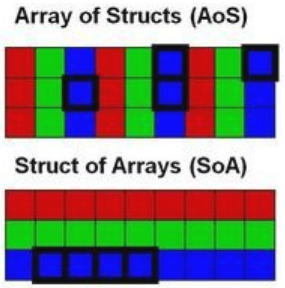
APPLICATION SPECIFIC COMPUTING

ziti

- Home
- How to find us
- How to join us
- Team
- Research
- Teaching







Our research focuses on significant improvements of performance and accuracy in application specific computing through a global optimization across the entire spectrum of numerical methods, algorithm design, software implementation and hardware acceleration.

These layers typically have contradictory requirements and their integration poses many challenges. For example, numerically superior methods expose little parallelism, bandwidth efficient algorithms convolve the processing of space and time into unmanageable software patterns, high level language abstractions create data layout and composition barriers, and high performance on today's hardware poses strict requirements on parallel execution and data access. High performance and accuracy for the entire application can only be achieved by balancing these requirements across all layers.

The following topics are given particular attention:

- Mixed precision methods
- Multigrid methods
- Adaptive data structures
- Data representation
- Bandwidth optimization
- Reconfigurable computing

Heidelberg University
Department of Mathematics and Computer Science
Department of Physics and Astronomy
Institute of Computer Engineering (ZITI)
View PDF

Prof. Robert Strzodka
(successor of retired
Prof. Reinhard Männer)

Computers and Applications

More About the Future

2. Research Centers and Computing Centers

National Astronomical Observatories, CAS

Astron. Rechen-Inst.
ZENTRUM FÜR ASTRONOMIE
Uni Heidelberg

GPU Clusters used:

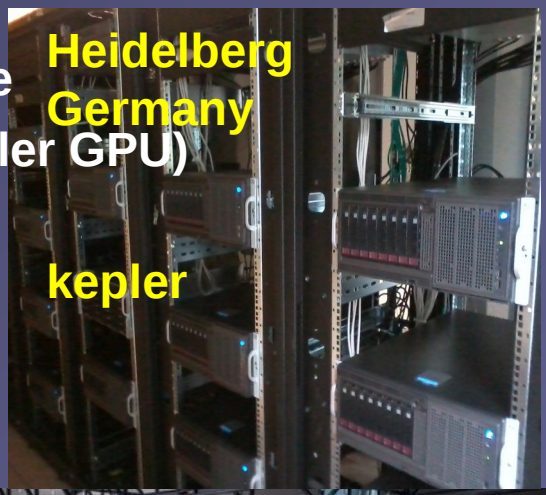
Heidelberg/Beijing *(NAOC/CAS and Silk Road Project)*

GPU servers wn14/hansolo/obiwan/… RTX 2080 Ti

JUWELS Booster (Nodes with 4x Ampere A100 GPU)

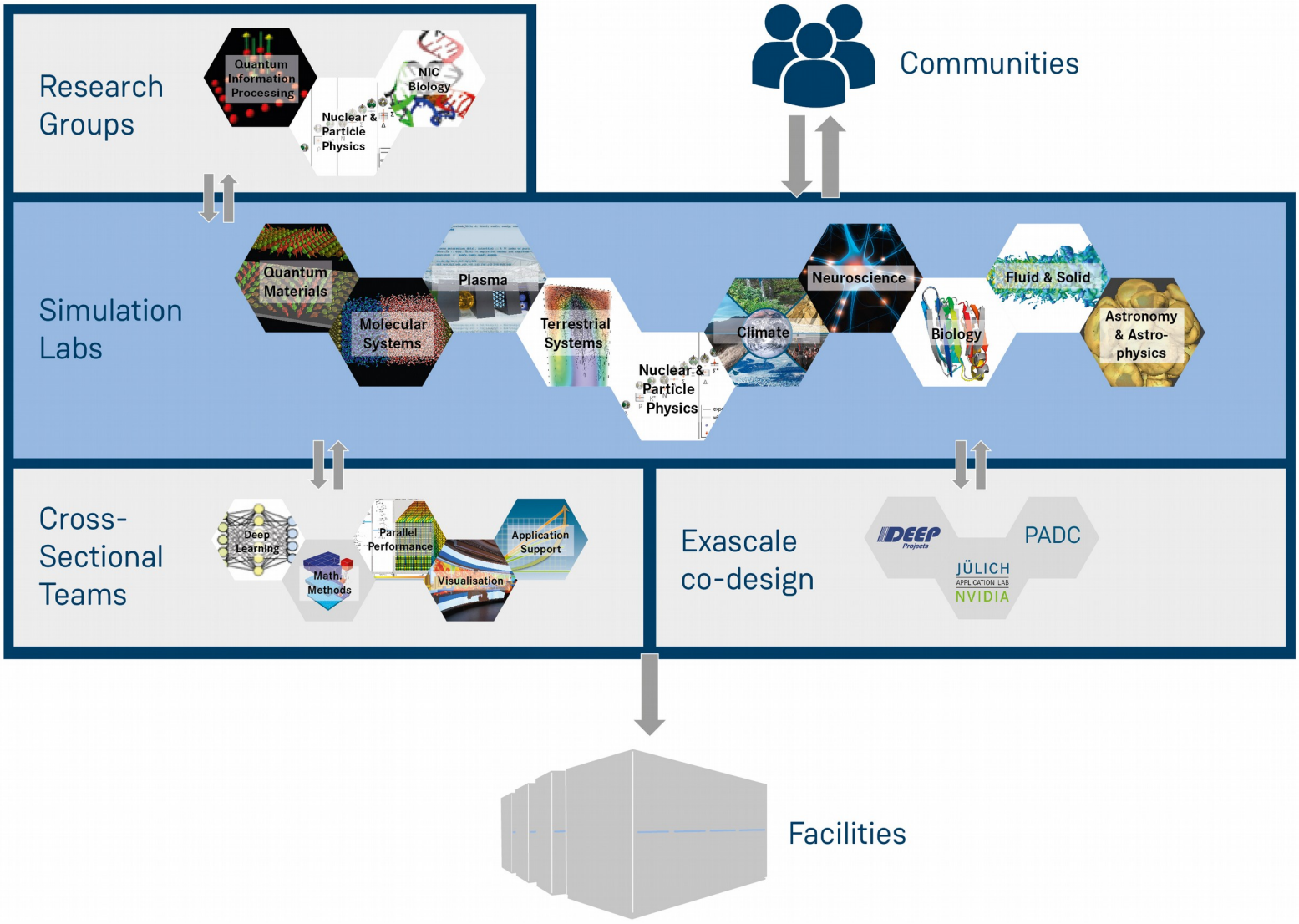
Golowood cluster, Main Astron. Observatory, Kiev, Ukraine

Kepler/bwFor clusters Heidelberg, Germany (12x +18x Kepler GPU)



Jülich Supercomputing Center (JSC)

https://www.fz-juelich.de/ias/jsc/EN/Expertise/SimLab/simlab_node.html



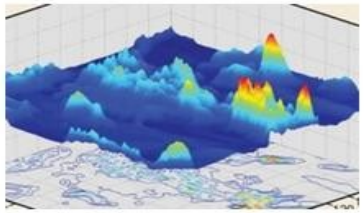
BERKELEY LAB

A-Z INDEX | DIRECTORY |

EXASCALE @ BERKELEY LAB

https://www.exascaleproject.org/ https://exascale.lbl.gov/

HOME SOFTWARE APPLICATIONS CO-DESIGN HARDWARE LEADING THE WAY



EQSIM

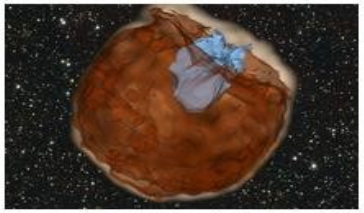
EQSIM: High Performance, Multidisciplinary Simulations for Regional Scale Seismic Hazard and Risk Assessments EQSIM: High Performance, Multidisciplinary Simulations for Regional Scale Seismic Hazard and Risk Assessments is led by David McCallen...

Continue Reading

ExaBiome

ExaBiome: Exascale Solutions for Microbiome Analysis ExaBiome: Exascale Solutions for Microbiome Analysis is led by Associate Lab Director for Computing Sciences Kathy Yelick, with support from Los Alamos National Laboratory and DOE's...

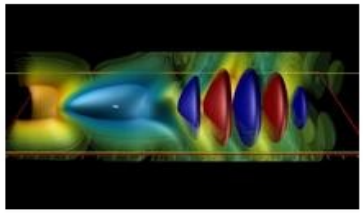
Continue Reading



ExaStar

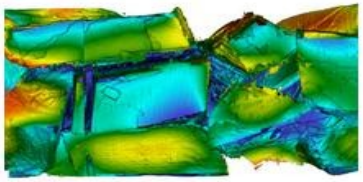
ExaStar: Exascale Models of Stellar Explosions: Quintessential Multi-Physics Simulation ExaStar: Exascale Models of Stellar Explosions: Quintessential Multi-Physics Simulation is led by Daniel Kasen of the Nuclear Science Division with support from...

Continue Reading



WarpX

WarpX: Exascale Modeling of Advanced Particle Accelerators WarpX: Exascale Modeling of Advanced Particle Accelerators is led by Jean-Luc Vay of the Accelerator Technology and Applied



Subsurface

Subsurface: An Exascale Subsurface Simulator of Coupled Flow, Transport, Reactions and Mechanics Subsurface: An Exascale Subsurface Simulator of Coupled Flow, Transport, Reactions

http://english.ipe.cas.cn/

March 04, 2021 Please input keywo Search Home | Contact | Sitemap | 中文 | CAS

Institute of Process Engineering, Chinese Academy of Sciences

Home | News | About Us | Research | People | International Cooperation | Graduate Education | Papers | Join Us

Events

» more

· New System Improves Coking Wastewater Treatment Efficiency

Based on the idea of whole-process pollution control, researchers led by Prof. CAO Hongbin from IPE, lowered the cost of coking wastewater treatment by 20 percent in their new system, and achieved stable and efficient removal of toxic and polluting particles.

· China Focus: New Technology Enables Large-scale Production of Artemisinin for Malaria

Researchers from the Institute of Process Engineering (IPE) of Chinese Academy of Sciences have developed a new technology to produce artemisinin, a top malaria treatment, on a large scale.

Upcoming Events

photosynthesis 06-11

Lecture: Rate Processes in Particle and Powder Technology 06-10

Lecture: Three Approaches to choose an ionic solvent 05-19

Lecture: Magnetic Resonance Imaging and Computational Modeling of Multiphase Granular Flows 05-19

Join Us

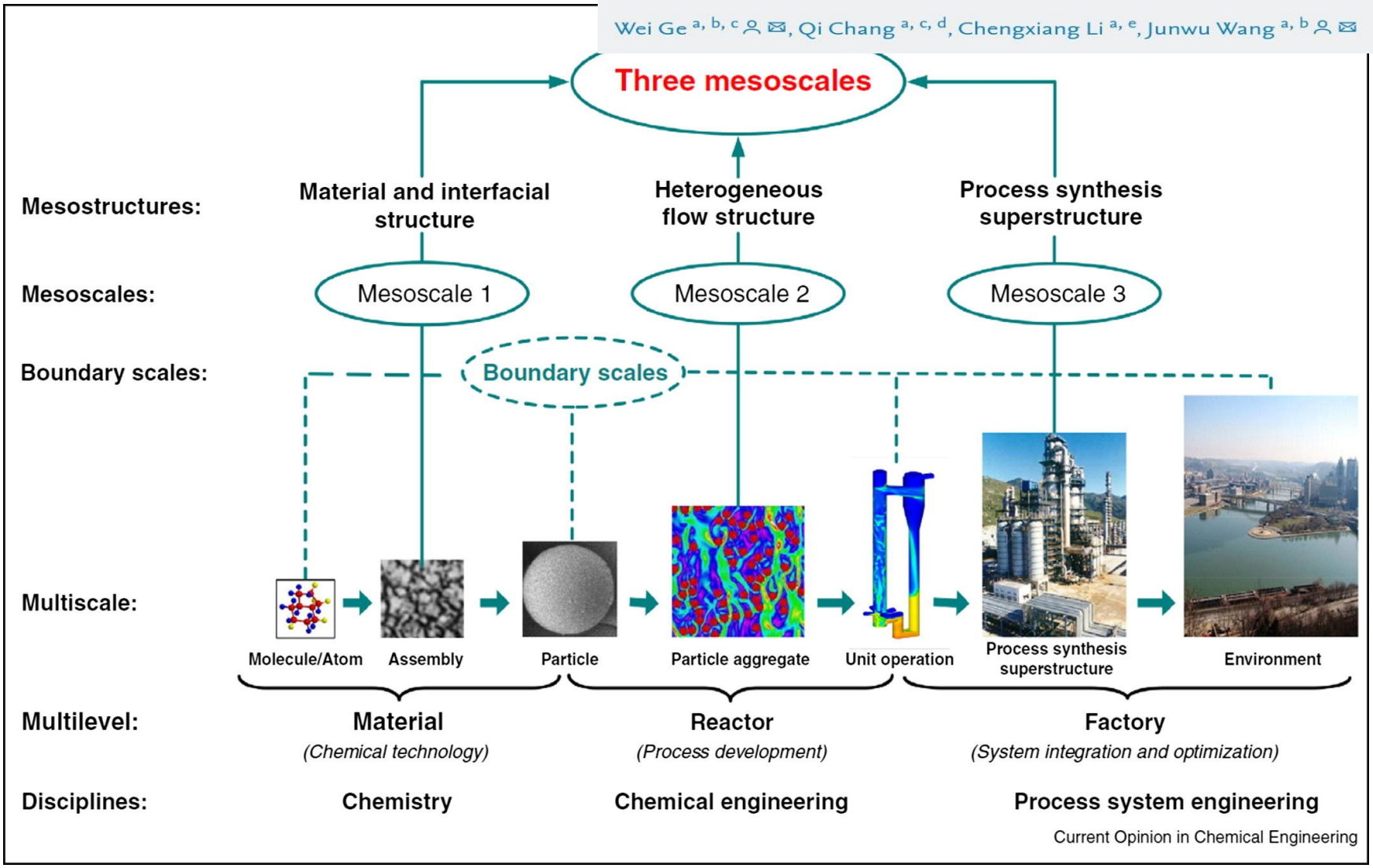
Vacancies at Mesoscience Center

Chemical Engineering Science
Volume 198, 28 April 2019, Pages 198-223

Review

Multiscale structures in particle–fluid systems: Characterization, modeling, and simulation

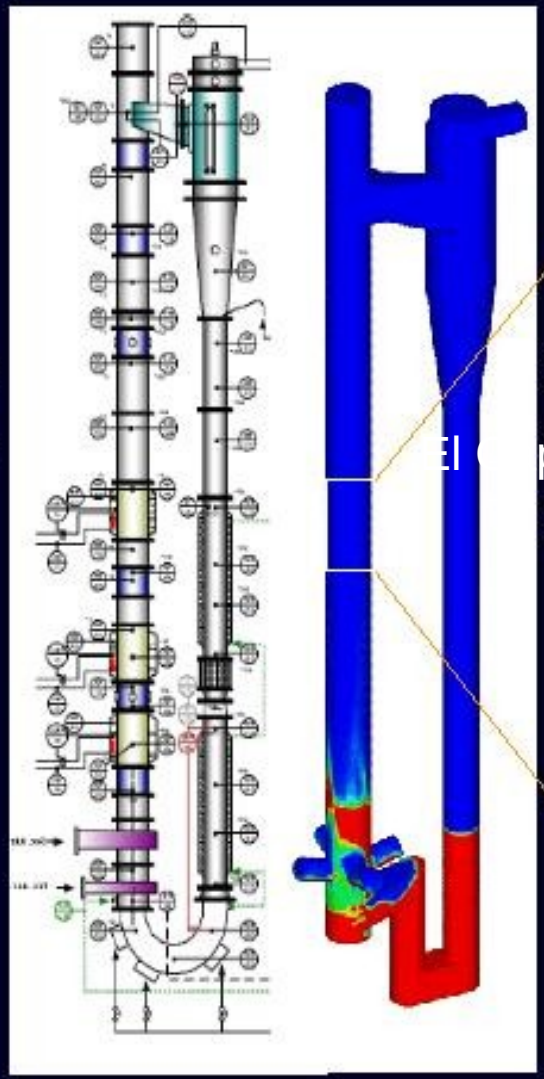
Wei Ge [a, b, c], Qi Chang [a, c, d], Chengxiang Li [a, e], Junwu Wang [a, b]



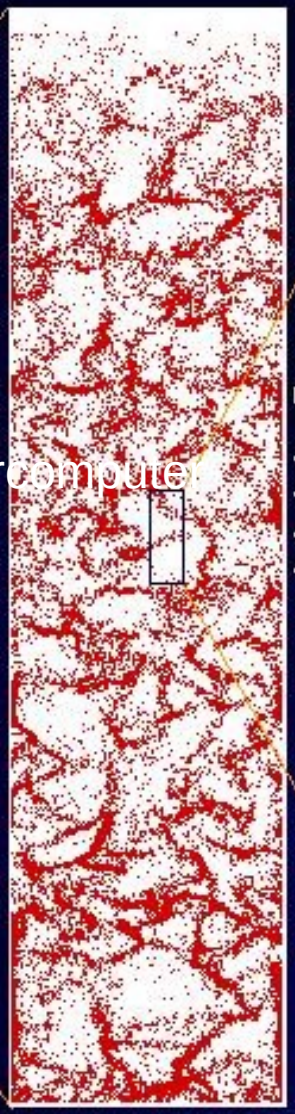
DNS of gas-solid flow : >20x speedup (1C1060/1E5430 core)

120K Particles + 400M pseudo-particles

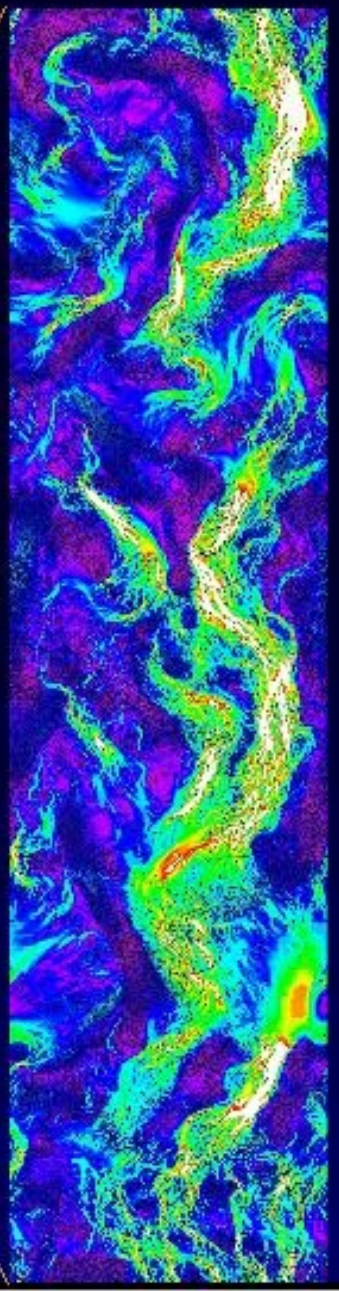
Reactor:
0.4*20m
3D



Section:
0.4*1m
2D



Cell:
2*10cm
2D



Computers and Applications

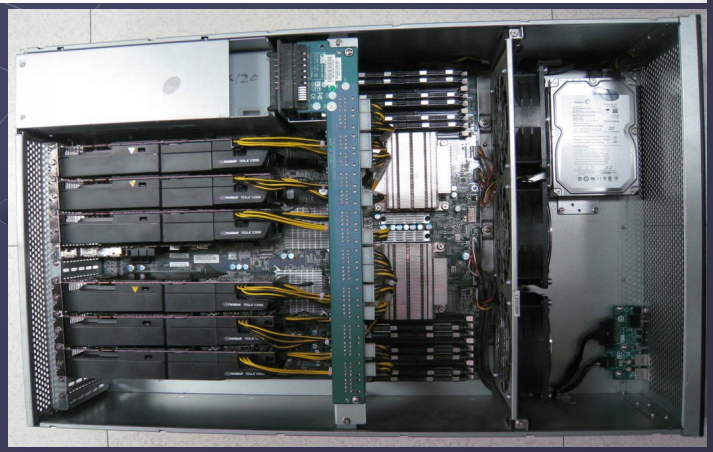
More About the Future

3. Supercomputing Systems

Fermi-based GPU supercomputer IPE

Mole-8.5(2010.04.24)

Rpeak SP：	2Pflops
Rpeak DP：	1Pflops
Linpack:	207.3T (Top500 19th)
Mflops/Watt:	431 (Green500 8th)
Total RAM：	17.2TB
Total VRAM：	6.6TB
Total HD：	360TB
Inst. Comm.：	H3C GE
Data Comm.：	Mellanox QDR IB
Occupied area：	150 sq.m.
Weight：	12.6 tons
Max Power：	600kW(computing) 200kW(cooling)
System：	CentOS 5.4, PBS
Monitor：	Ganglia, GPU monitor
Languages：	C, C++, CUDA 3.1 , OpenCL

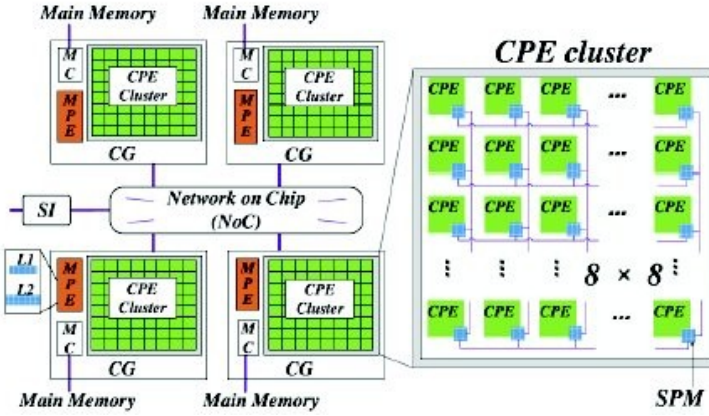


中国科学院过程工程研究所
Institute Of Process Engineering, Chinese Academy Of Sciences

TOP500 #4: Wuxi, Jiangsu Prov., near Shanghai, China

SUNWAY TAIHULIGHT

- SW26010 processor (Chinese design, ISA, & fab)
- 1.45 GHz
- Node = 260 Cores (1 socket)
  - 4 – core groups
  - 32 GB memory
- 40,960 nodes in the system
- 10,649,600 cores total
- 1.31 PB of primary memory (DDR3).
- 125.4 Pflop/s theoretical peak
- 93 Pflop/s HPL, 74% peak
- 15.3 Mwatts water cooled
- 3 of the 6 finalists for Gordon Bell Award@SC16



Chinese Processor Architecture
260 cores on socket.

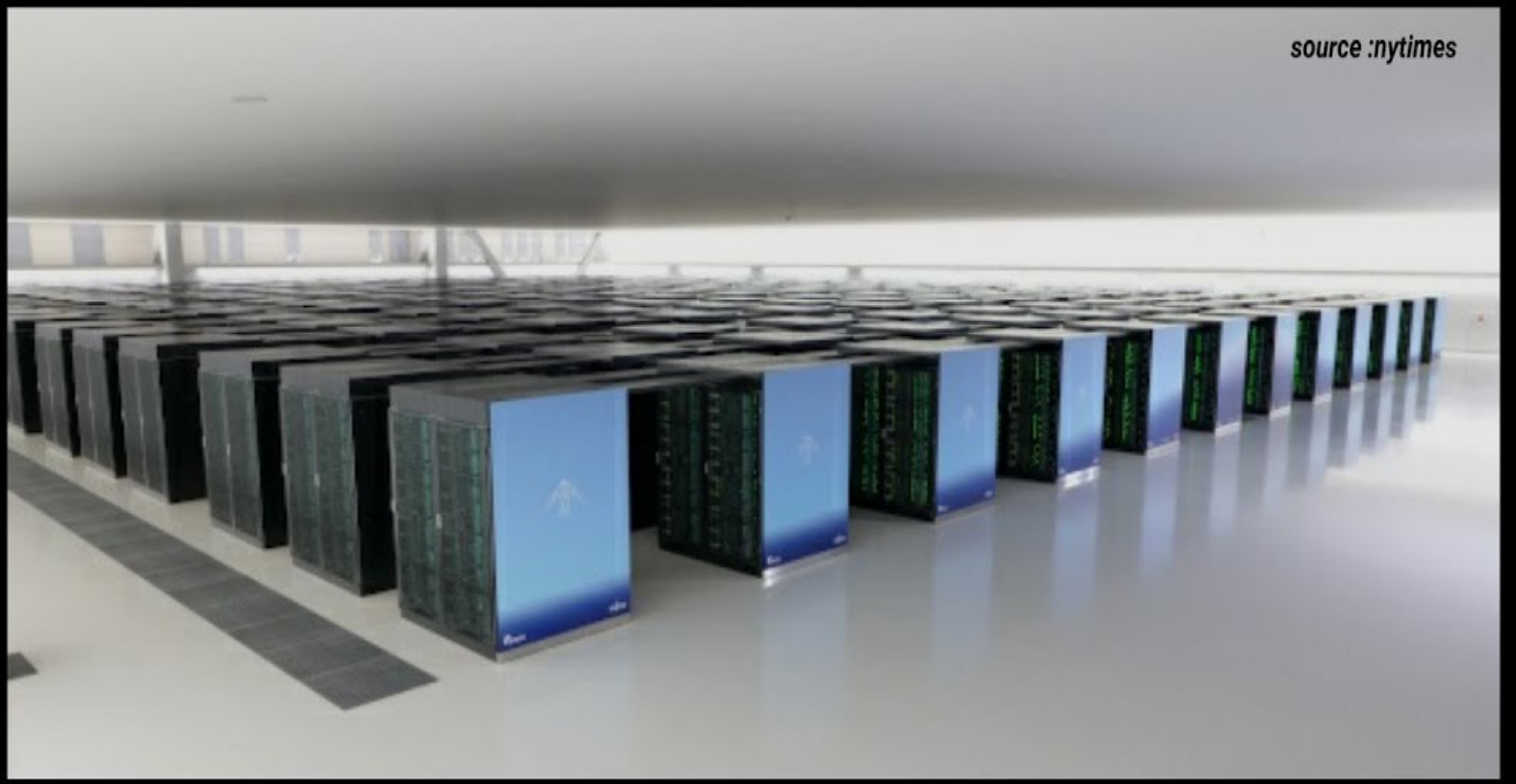
FUGAKU

Nature's Secrets

富岳

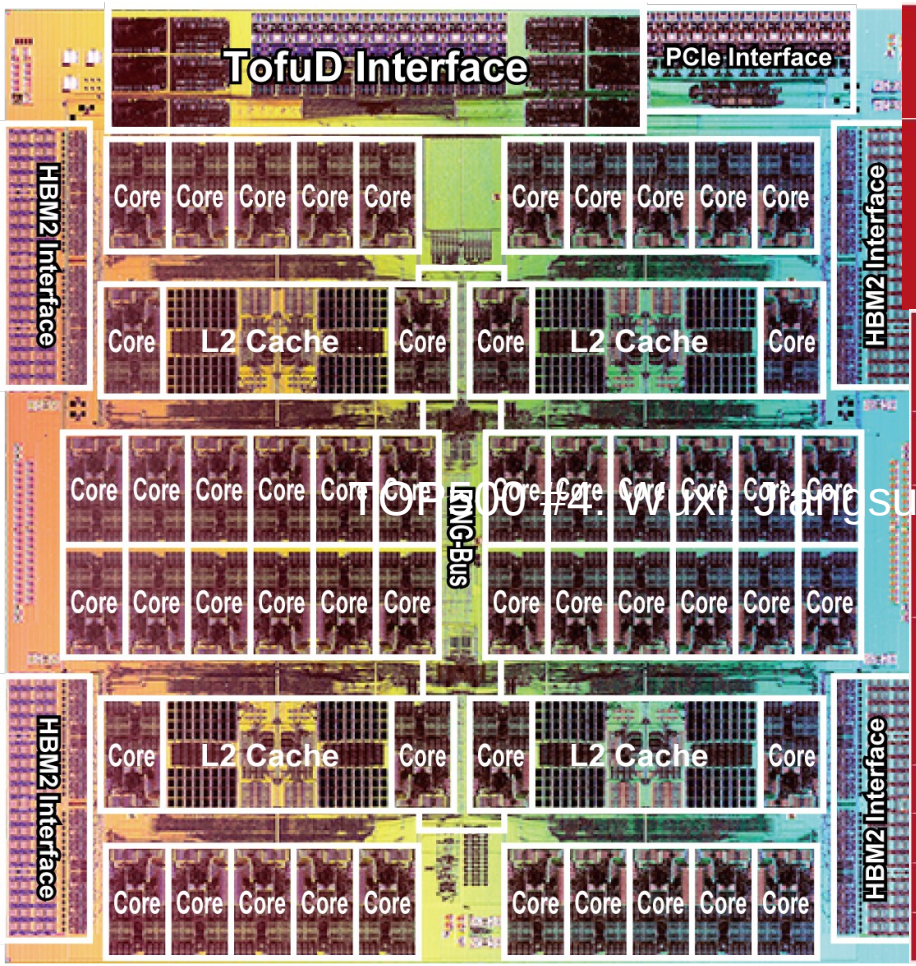
Mt. Fuji

The world's fastest Super Computer 2020



Currently fastest supercomputer of the world: Fugaku @ RIKEN Center, Kobe, Japan

理化学研究所 RIKEN



Number of Nodes		158,976 nodes 384 nodes x 396 racks = 152,064 192 nodes x 36 racks = 6,912
Peak	Normal Mode: 2.0 GHz	• Double Precision (64 bit) 488 Petaflops • Single Precision (32 bit) 977 Petaflops • Half Precision (16 bit) 1.95 Exaflops • Integer (8 bit) 3.90 Exa Exaops
Architecture		Armv8.2-A SVE 512bit With the following Fujitsu's extensions: Hardware barrier, Sector cache, and Prefetch
Core		48 cores for compute and 2 or 4 cores for OS activities 4 CMGs (NUMA nodes)
Performance	Normal Mode: 2.0 GHz	DP: 3.072 TF, SP: 6.144 TF, HP: 12.288 TF
	Boost Mode: 2.2 GHz	DP: 3.3792 TF, SP: 6.7584 TF, HP: 13.5168 TF
Cache*1 *2		L1D/core: 64 KiB, 4way, 256 GB/s (load), 128 GB/s (store) L2/CMG: 8 MiB, 16way L2/node: 4 TB/s (load), 2 TB/s (store) L2/core: 128 GB/s (load), 64 GB/s (store)
Memory		HBM2 32 GiB, 1024 GB/s
Interconnect		Tofu Interconnect D (28 Gbps x 2 lane x 10 port)
I/O		PCIe Gen3 x16
Technology		7nm FinFET

TOP500 #4: Wuxi, Jiangsu Prov., near Shanghai, China

https://www.riken.jp/en/

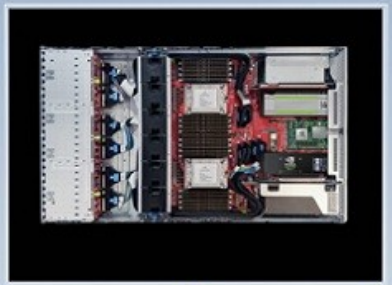
https://www.r-ccs.riken.jp/en/fugaku/project/outline

Computers and Applications

More About the Future

4. On the Software (CUDA?)

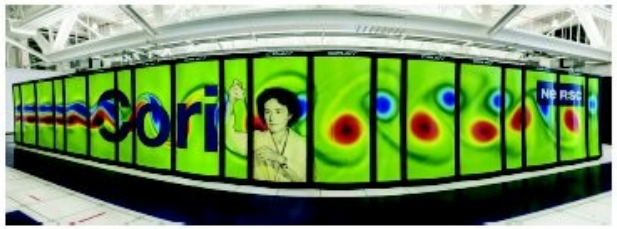
NVIDIA EXPANDS SUPPORT FOR ARM

HPC	CLOUD / DATA CENTER	EDGE AI / ROBOTICS	PC
			
Fujitsu A64FX Mellanox InfiniBand	Ampere Altra Marvell ThunderX2 NVIDIA BlueField DPU	NVIDIA BlueField DPU NVIDIA Isaac	NVIDIA GPU RTX Linux
Arm v8.2 + SVE	Arm Neoverse V/N	Arm Neoverse E	Arm v8.2

NVIDIA AI	NVIDIA RAPIDS	NVIDIA HPC	CUDA-X	NVIDIA Magnum IO	NVIDIA DeepStream	Nsight Tools Suite

SDKs and TOOLS

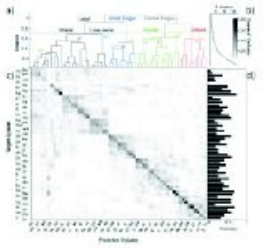
Deep Learning in Science



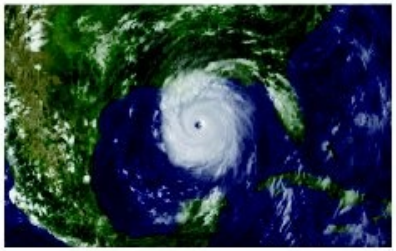
Cray XC40 system at NERSC

Modeling galaxy shapes

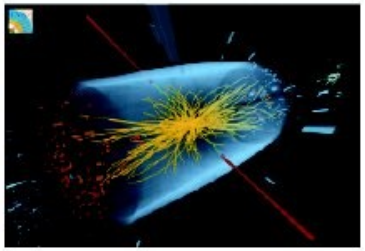
Clustering Daya Bay events



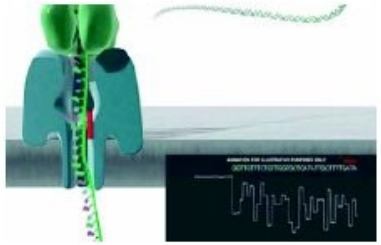
Decoding speech from ECoG



Detecting extreme weather



Classifying LHC events



Oxford Nanopore sequencing

NERSC

BERKELEY LAB
LAWRENCE BERKELEY NATIONAL LABORATORY

Opportunities to apply DL widely in support of classic HPC simulation and modelling

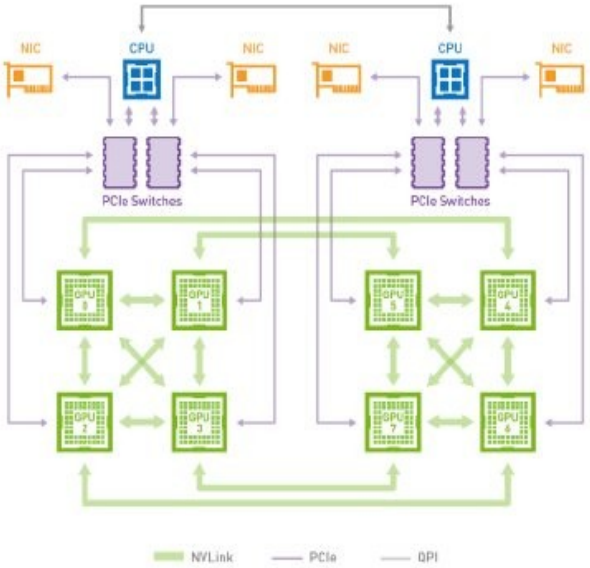
SYSTEMS APPROACHES TO EXASCALE

More GPUs, Fewer CPUs:

Titan: 1GPU/CPU

Summit: 3 GPUs/CPU

Exascale: ?



Faster Serial Processing (~~MANY CORE~~):

Run 8x Fewer Cores @ 2x Speed

Denser Packaging:

Move Networking to Faster Local Networks: NVLINK

EXASCALE: "50X FASTER THAN TITAN"

Per-GPU -hardware- speedups will be less than 50x

	2013 Kepler	2016 Pascal	2017 Volta	2021*	Speedup
FP64 Tflop/s	1.5	4.5	7	7-21	5-15
Memory GB/s	288	720	900	900-4000	3-14
I/O BW GB/s	7	80	150	150-500	20-70
Deep Learning FP16 Tflop/s	3	20	112	112-500	37-166
Deep Learning BW GB/s	576	2880	3600	3600-16000	6-27

*Extremely Fuzzy Public Projections for 2021

El Capitan Supercomputer Detailed: AMD CPUs & GPUs To Drive 2 Exaflops of Compute

55 Comments

+ Add A Comment

by Ryan Smith on March 4, 2020 1:00 PM EST

Posted in CPUs AMD HPC GPUs Cray El Capitan

https://www.anandtech.com/show/15581/el-capitan-supercomputer-detailed-amd-cpus-gpus-2-exaflops

Software? See comments….

Back in August, the United States Department of Energy and Cray announced plans for a third United States exascale supercomputer, El Capitan. Scheduled to be installed in Lawrence Livermore National Laboratory (LLNL) in early 2023, the system is intended primarily (but not exclusively) for use by the National Nuclear Security Administration (NNSA), who uses supercomputers in their ongoing nuclear weapons modeling. At the time the system was announced, The DOE and LLNL confirmed that they would be buying a Shasta system from Cray (now part of HPE), however the announcement at the time didn't go into any detail about what hardware would actually be filling one of Cray's very flexible supercomputers.

Computers and Applications

More About the Future

End of Presentation

(now matmul etc.)

Important Note:

If you do some NBODY research in the future, please contact us (tutors or lecturer); do not use the course code for research it is outdated.

Remember for certificate of course:

* Output files of small experiments on your lecture account (0_hello, 1_add, … , 7-matmul, 8-histo)
* Return two plots, one data file, and a few comments to your tutors Deadline agreement with tutors! (Group 1: Mar 29, 2,3: Mar 19)
* Notice: Student Queues will close Sunday, Mar 7, 23:59. You can run later, but contact me please spurzem@ari.uni-heidelberg.de

Additional deeper material:

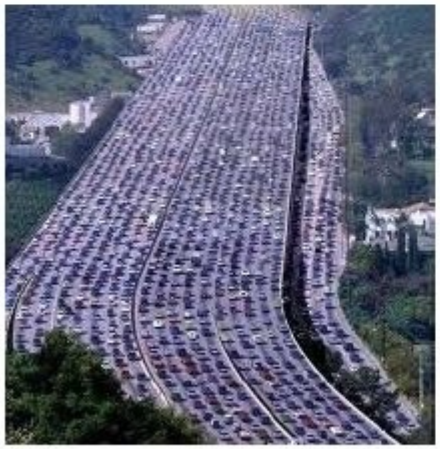
Lectures by Prof. Wen-Mei Hwu Chicago in Berkeley 2012 and Beijing 2013, see http://iccs.lbl.gov/workshops/tutorials.html
(down on page links to all lecture files, also available on request from spurzem@nao.cas.cn)

Lecture1: Computational thinking
Lecture2: Parallelism Scalability
Lecture3: Blocking Tiling
Lecture4: Coarsening Tiling
Lecture5: Data Optimization
Lecture6: Input Binning
Lecture7: Input Compaction
Lecture8: Privatization

See also:
http://freevideolectures.com/Course/2880/Advanced-algorithmic-techniques-for-GPUs/1

Massive Parallelism - Regularity




Main Hurdles to Overcome

- Serialization due to conflicting use of critical resources
- Over subscription of Global Memory bandwidth
- Load imbalance among parallel threads

Computational Thinking Skills

- The ability to translate/formulate domain problems into computational models that can be solved efficiently by available computing resources
  - Understanding the relationship between the domain problem and the computational models
  - **Understanding the strength and limitations of the computing devices**
  - **Defining problems and models to enable efficient computational solutions**

DATA ACCESS CONFLICTS

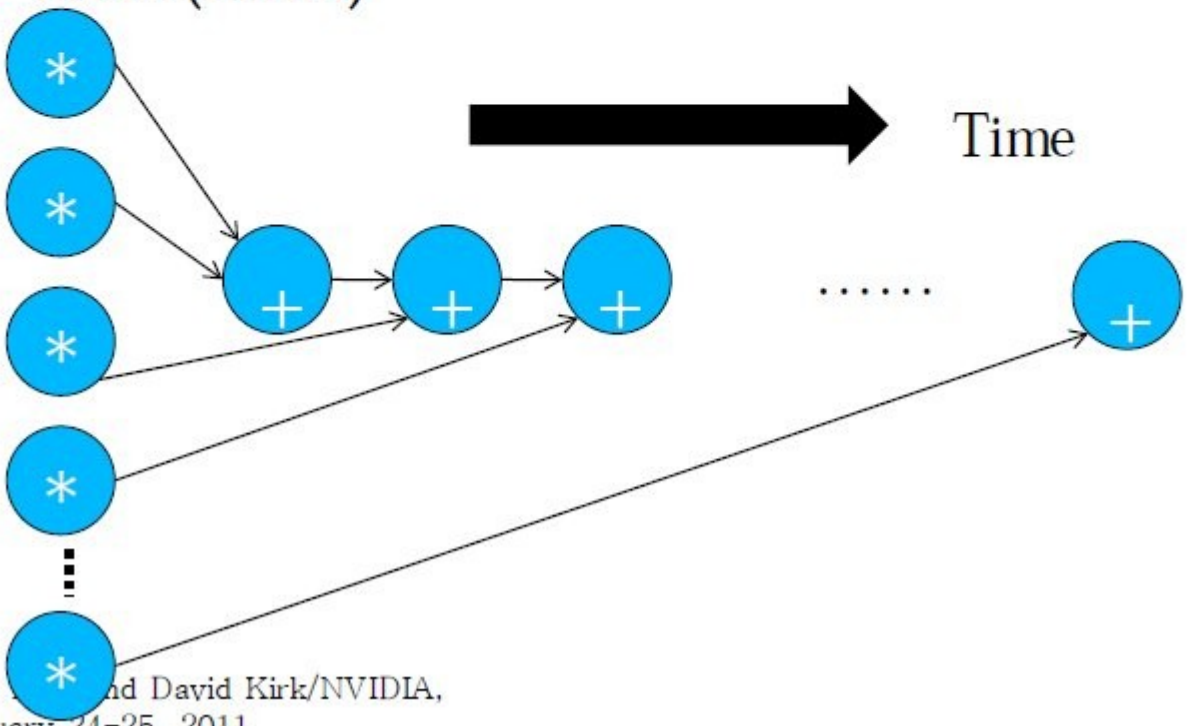
Conflicting Data Accesses Cause Serialization and Delays

- Massively parallel execution cannot afford serialization

- Contentions in accessing critical data causes serialization

A Simple Example

- A naïve inner product algorithm of two vectors of one million elements each
  - All multiplications can be done in time unit (parallel)
  - Additions to a single accumulator in one million time units (serial)



How much can conflicts hurt?

- Amdahl's Law
  - If fraction X of a computation is serialized, the speedup can not be more than 1/(1-X)
- In the previous example, X = 50%
  - Half the calculations are serialized
  - No more than 2X speedup, no matter how many computing cores are used

GLOBAL MEMORY BANDWIDTH

Global Memory Bandwidth

Ideal

Reality

Global Memory Bandwidth

- Many-core processors have limited off-chip memory access bandwidth compared to peak compute throughput
- Fermi
  - 1 TFLOPS SPFP peak throughput
  - 0.5 TFLOPS DPFP peak throughput
  - 144 GB/s peak off-chip memory access bandwidth
    - 36 G SPFP operands per second
    - 18 G DPFP operands per second
  - To achieve peak throughput, a program must perform 1,000/36 = ~28 SPFP (14 DPFP) arithmetic operations for each operand value fetched from off-chip memory

LOAD BALANCE

Load Balance

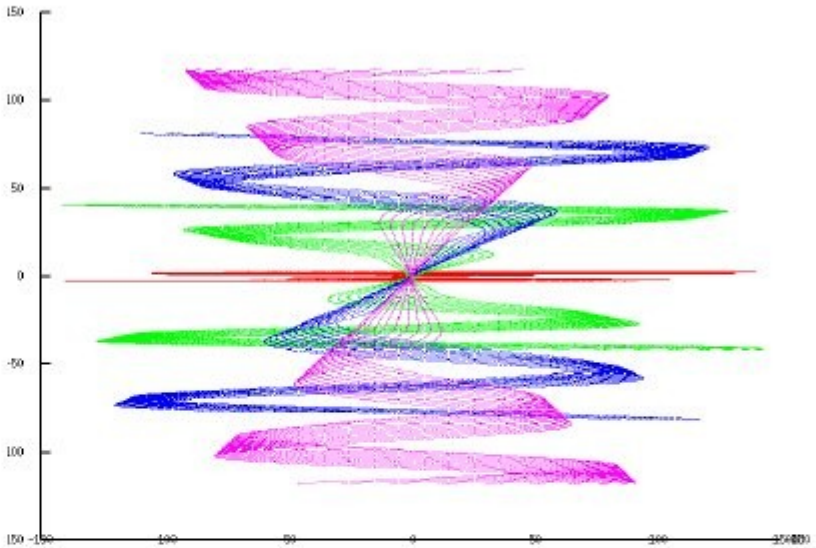
- The total amount of time to complete a parallel job is limited by the thread that takes the longest to finish

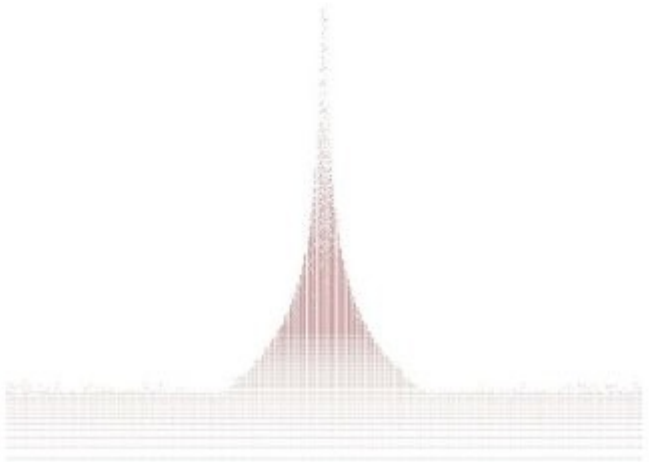
How bad can it be?

- Assume that a job takes 100 units of time for one person to finish
  - If we break up the job into 10 parts of 10 units each and have fo10 people to do it in parallel, we can get a 10X speedup
  - If we break up the job into 50, 10, 5, 5, 5, 5, 5, 5, 5, 5 units, the same 10 people will take 50 units to finish, with 9 of them idling for most of the time. We will get no more than 2X speedup.

How does imbalance come about?

- Non-uniform data distributions
  - Highly concentrated spatial data areas
  - Astronomy, medical imaging, computer vision, rendering, …
- If each thread processes the input data of a given spatial volume unit, some will do a lot more work than others





Eight Algorithmic Techniques (so far)

Technique	Contention	Bandwidth	Locality	Efficiency	Load Imbalance	CPU Leveraging
Tiling		X	X			
Privatization	X		X			
Regularization				X	X	X
Compaction		X				
Binning		X	X	X		X
Data Layout Transformation	X		X			
Thread Coarsening	X	X	X	X		
Scatter to Gather Conversion	X					

http://courses.engr.illinois.edu/ece598/hk/

You can do it.

- Computational thinking is not as hard as you may think it is.
  - Most techniques have been explained, if at all, at the level of computer experts.
  - The purpose of the course is to make them accessible to domain scientists and engineers.

ANY MORE QUESTIONS?