DE LA RECHERCHE À L'INDUSTRIE

ACCELERATED APPLICATIONS AND CONVERGENCE BETWEEN HPC, HPDA AND AI AT CEA

7th ADAC Workshop – ORNL | Christophe CALVIN CEA/DRF | christophe.calvin@cea.fr

OUTLINE

European and French HPC landscape

Big Data – Open data & Open science

Efficient GPU programming

From optimal performances to performances portability

Conclusion

DE LA RECHERCHE À L'INDUSTRIE

www.cea.fr

EUROPEAN AND FRENCH HPC LANDSCAPE

EUROPEAN LANDSCAPE - PRACE

Hazel Hen, GCS@HLRS, Germany
Cray XC40 – 7.4 PFlops
185k cores – Intel Haswell

JOLIOT Curie, GENCI@CEA France
BULL SEQUANA – 9.4 PFlops
79.5k cores – Intel SKL – 6.9 PFlops
56k cores – Intel KNL – 2.5 PFlops

JUWELS, GCS@FZJ, Germany
BULL SEQUANA – 12 PFlops
122.5k cores – Intel SKL – 10.4 PFlops
192 NVIDIA V100 – 1.6 PFlops

MARCONI, CINECA, Italy
Lenovo NeXtScale – 19 PFlops
110k cores – Intel SKL – 8 PFlops
245 k cores – Intel KNL – 11 PFlops

MareNostrum, BSC, Spain
166k cores – Intel SKL – 11 PFlops
Power9+Volta – 1.5 PFlops
Intel KNH – 0.5 PFlops
ARM-v8 – 0.5 PFlops

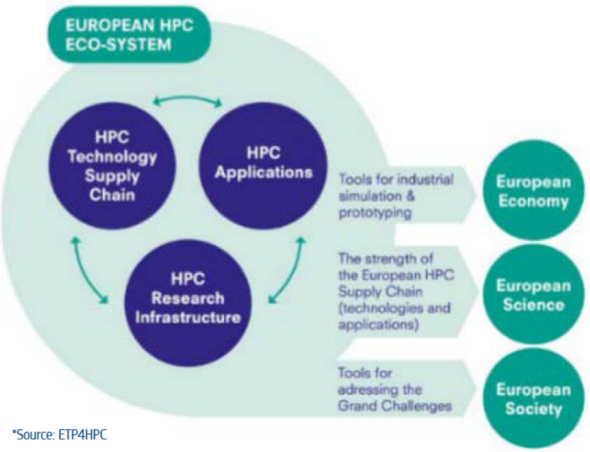
Piz Daint, ETH Zurich/CSCS, Switzerland
Cray XC40/XC50
5230 nodes Intel Xeon E5 + P100 – 25 PFlops
1431 nodes Intel Xeon E5 – 1.7 PFlops

SuperMUC-NG, GCS@LRZ, Germany
311 k cores – Intel SKL – 27 PFlops

Over 110 Pflops
- Mainly x86 Intel (>70% Peak Perf)
- GPU (NVIDIA)
- Manycore
- Power
- ARM

The European HPC strategy

The overall strategic goal is to develop a thriving European HPC ecosystem :



*Source: ETP4HPC

Infrastructure: Capacity of acquiring leadership-class computers

Technology: Securing our own independent HPC system supply

Applications: Excellence in HPC applications, and widen HPC use

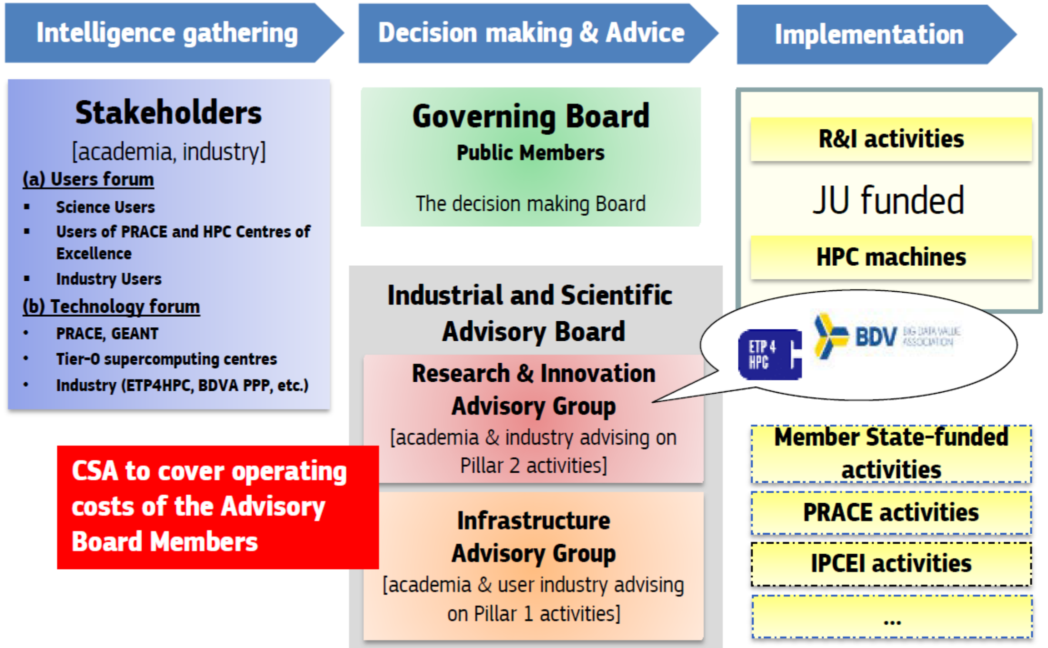
EuroHPC JU - Objectives

A world-class European HPC, Big Data and Cloud Ecosystem

Co-invest on a leading HPC and data infrastructure

for our scientists, industry and the public sector and support the development of technologies and applications across a wide range of fields

Governing Structure



EUROPEAN LANDSCAPE - EUROHPC

EuroHPC – Activities

HPC Ecosystem: Infrastructure & Operations	HPC Ecosystem: R&I, Application & Skills	JU Admin/Running costs		
~270	**min 180**	**10**	**486 m€**	EC
~290	~186	10	**486 m€**	**Participating States**
560	392	20	972 m€	**Total**
0	~420 (in kind)	2	422 m€	**Private Members**

- **Infrastructure + Operations**
  - ***Acquisition of 2 pre-exascale machines and several (tbd) mid-range machines***
- **Applications & Skills + R&I**
  - ***R&I,*** exascale technologies and systems (incl. low-power processor); applications
- **JU Admin/running costs**
  - **JU Operation: 2019 to 2026**

Acquistion of 2 exascale systems: 2022/2023

The Infrastructure Pillar: Supercomputers

Precursors to exascale

At least 2 Precursors to exascale

EU contribution:
≤50% of CAPEX and ≤50% of OPEX
MAX EU budget: 250 M€

Operations: end 2020

Petascale

At least 2 Petascale

EU contribution:
≤35% of CAPEX
MAX EU budget: 30 M€

Operations: mid 2020

FRENCH LANDSCAPE - GENCI

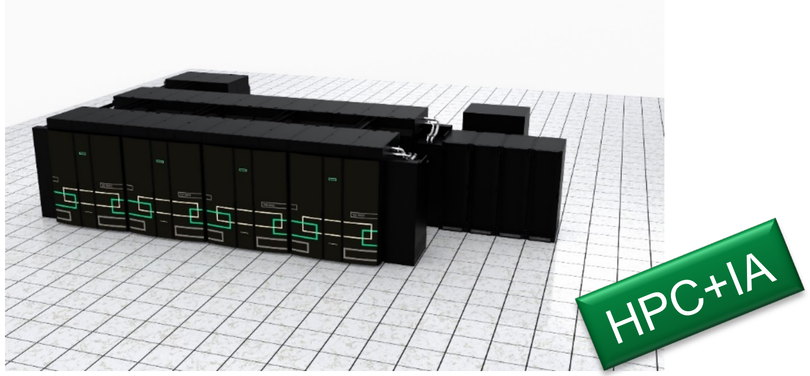
TGCC

JOLIOT Curie, GENCI@CEA France
BULL SEQUANA – 9.4 PFlops
79.5k cores – Intel SKL – 6.9 PFlops
56k cores – Intel KNL – 2.5 PFlops

CINES

OCCIGEN
BULL BullX – 3.5 PFlops
85k cores – Intel Haswell

IDRIS



JEAN ZAY
HPE – 14 PFlops
60k cores – Intel CascadeLake – 5 Pflops
262 converged nodes - 10k cores + 1k GPUs - 9 PFlops

1 new supercomputer in 2018: Joliot-Curie – SKL + KNL
1 new supercomputer in 2019: Jean Zay – CascadeLake+V100
Extension of Joliot-Curie in 2019: Yet another architecture?

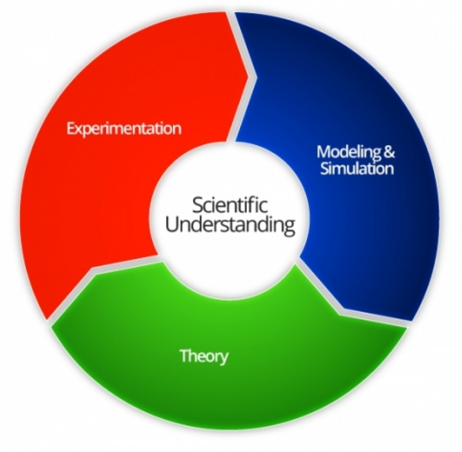
French researchers may have access up to 120 Pflops
But a quite large heterogeneity of architecture
No clear vision of future machines (pre-exa and exa)

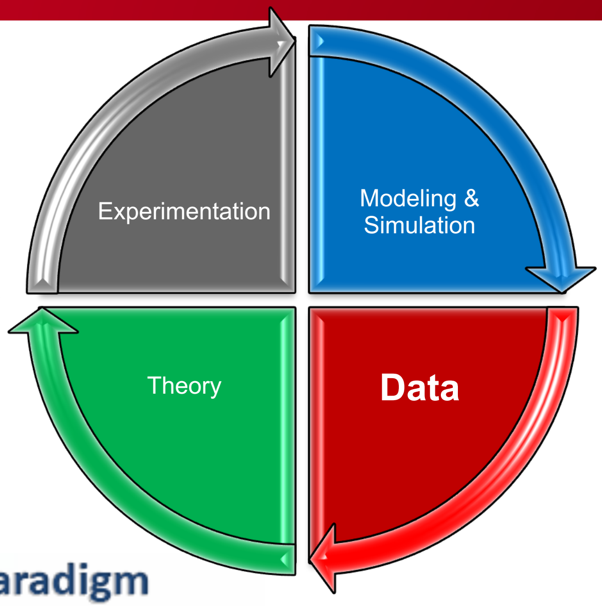
DE LA RECHERCHE À L'INDUSTRIE

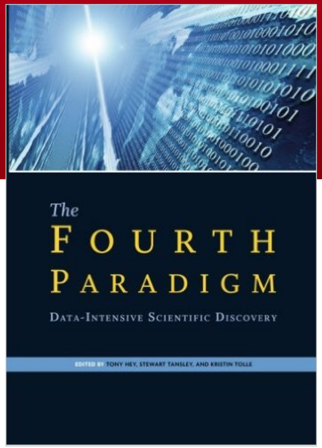
BIG DATA – OPEN DATA & OPEN SCIENCE

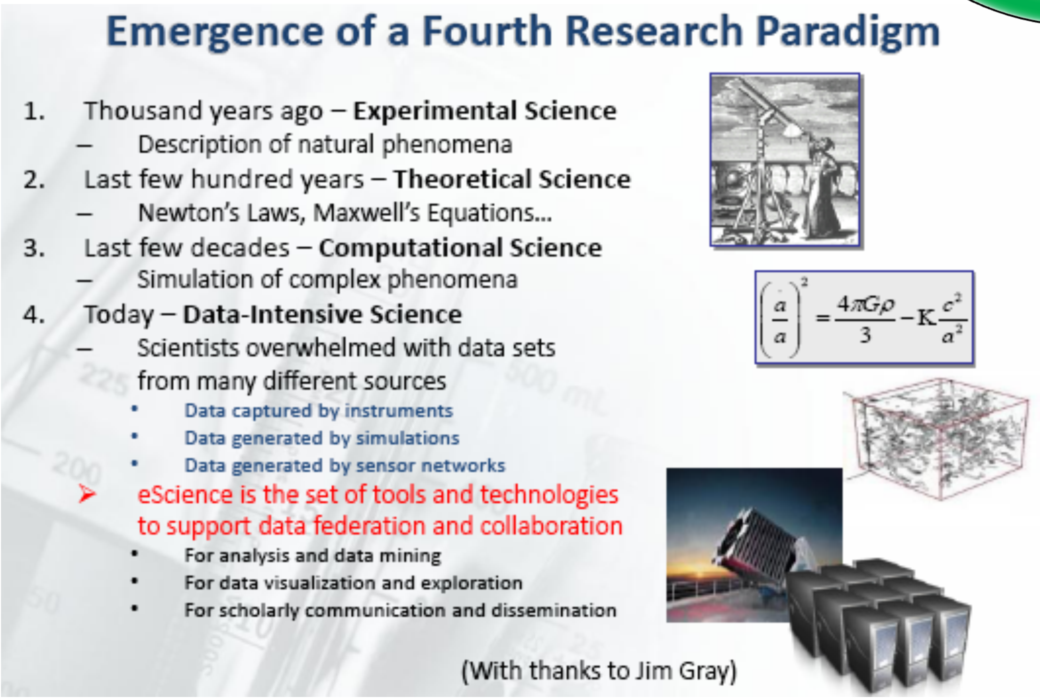
www.cea.fr

DATA: 4TH PILAR OF SCIENCE









Emergence of a Fourth Research Paradigm

1. Thousand years ago – **Experimental Science**
   – Description of natural phenomena
2. Last few hundred years – **Theoretical Science**
   – Newton's Laws, Maxwell's Equations...
3. Last few decades – **Computational Science**
   – Simulation of complex phenomena
4. Today – **Data-Intensive Science**
   – Scientists overwhelmed with data sets from many different sources
     - Data captured by instruments
     - Data generated by simulations
     - Data generated by sensor networks
   ➤ eScience is the set of tools and technologies to support data federation and collaboration
     - For analysis and data mining
     - For data visualization and exploration
     - For scholarly communication and dissemination

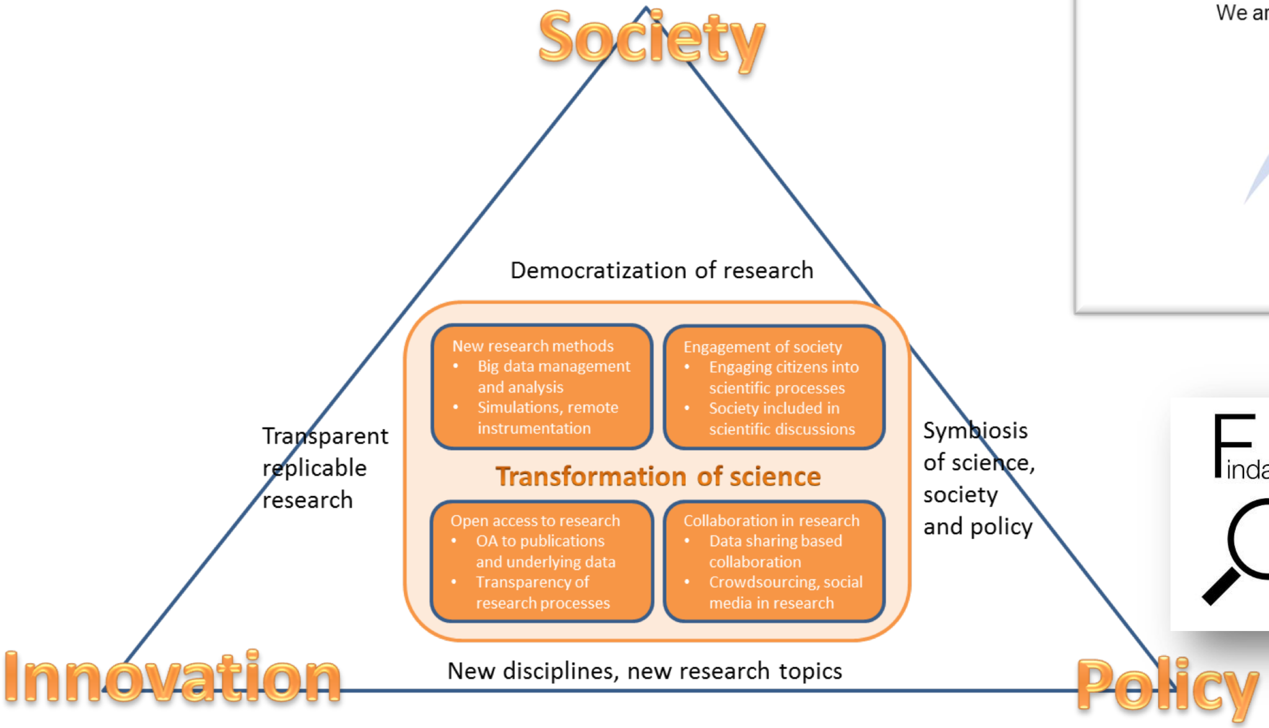
$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{4\pi G\rho}{3} - K\frac{c^2}{a^2}$$

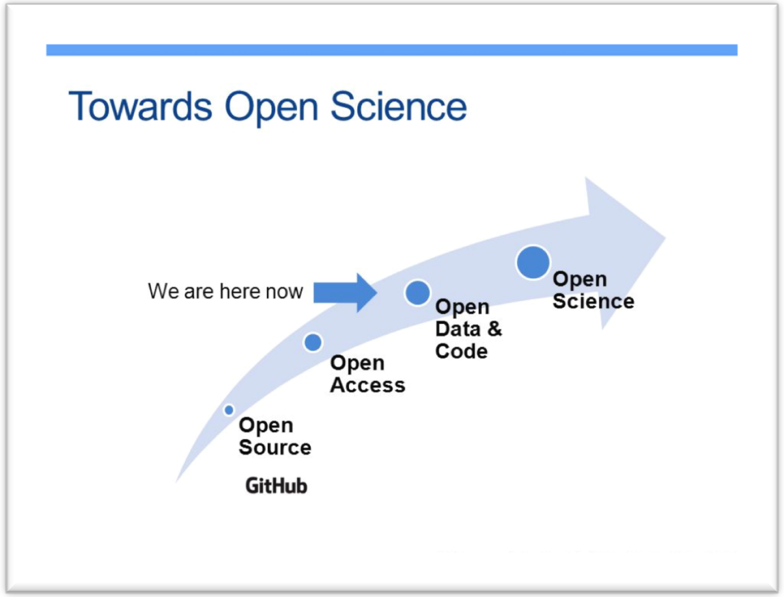
(With thanks to Jim Gray)

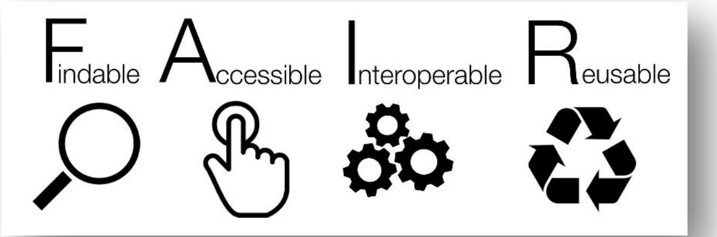
Data analysis
- Is a science → Data Science
- Contributes to scientific knowledge

OPEN DATA & OPEN SCIENCE

Open Science aims at transforming science through ICT tools, networks and media, to make research more open, global, collaborative, creative and closer to society.







But implies non-trivial constraints!

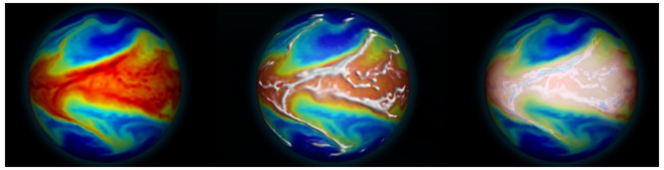
Data is becoming another challenge in today science

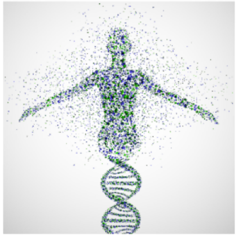
- Data is coming from many different sources:
  - Observations
  - Numerical simulations
  - Instruments and experiments

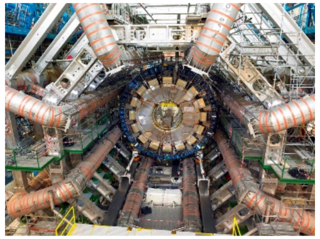
- Volume of data is increasing dramatically due to:
  - Needs for finer modeling
  - Progress in HPC usage
  - Instruments giving more and more information

- Interpretation and analysis of these data:
  - Use of HPC → HPDA
  - Mixing of many different sources of data
  - Mixing exp./obs data and simulation data

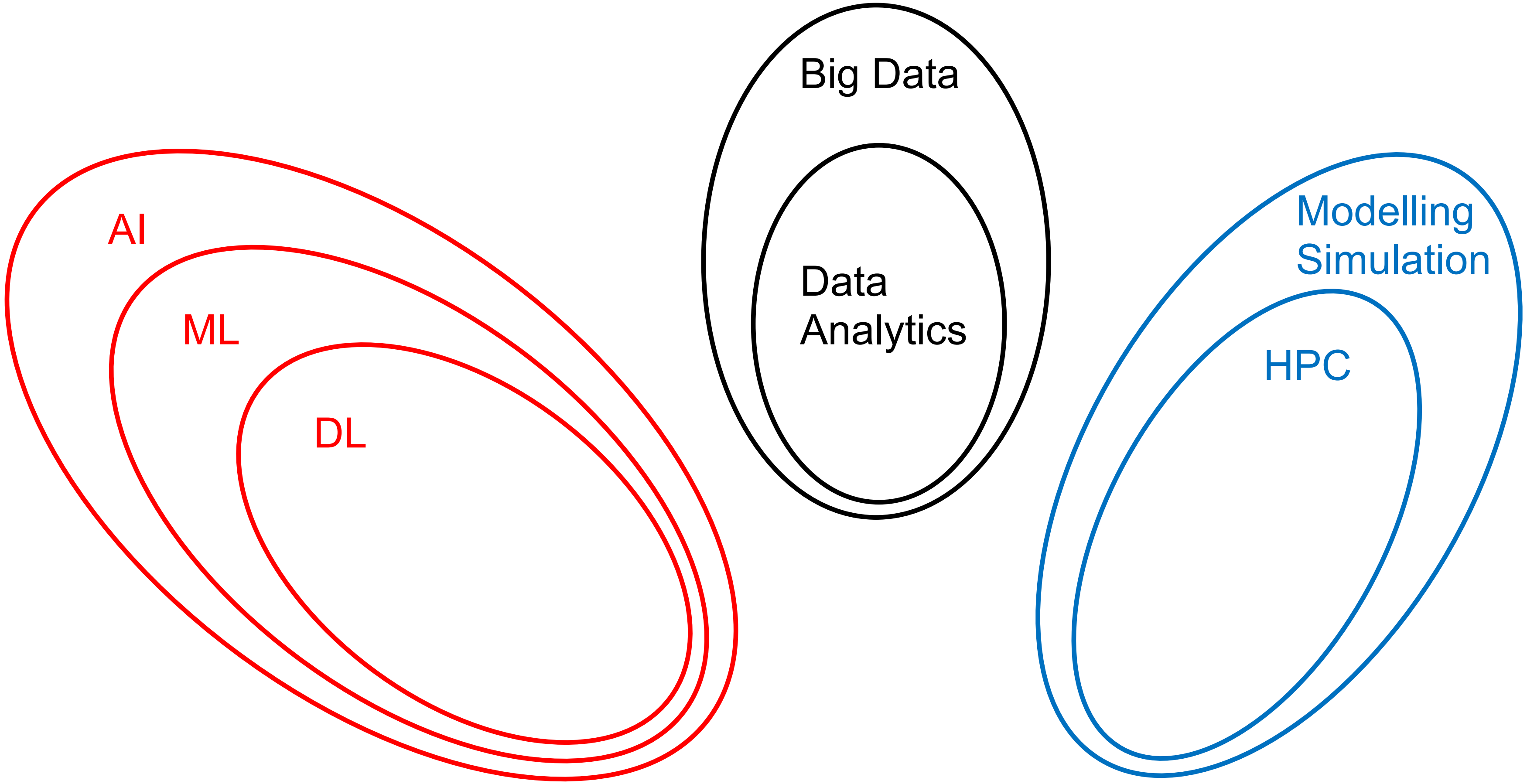
- Data needs to be open and FAIR
  - Complex workflow and dataflow
  - Data have to be accessible + services to explore



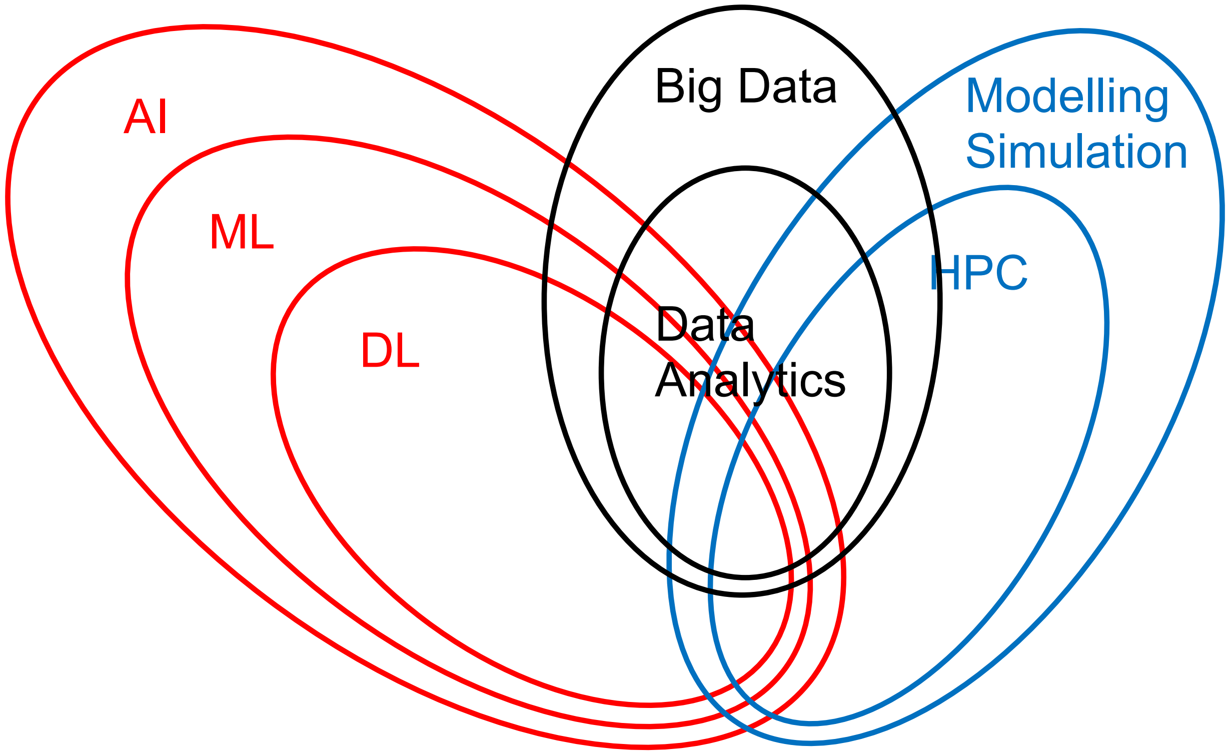




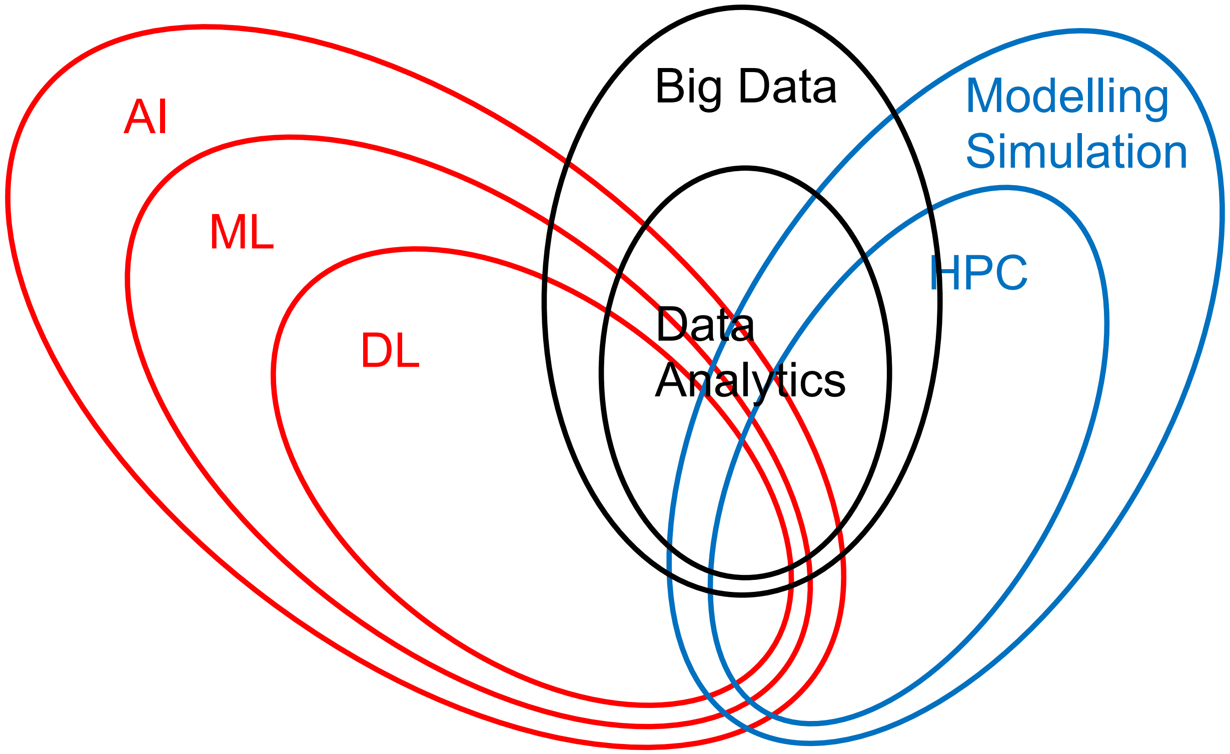
CONVERGENCE BETWEEN HPC / BIG DATA / AI



CONVERGENCE BETWEEN HPC / BIG DATA / AI



CONVERGENCE BETWEEN HPC / BIG DATA / AI

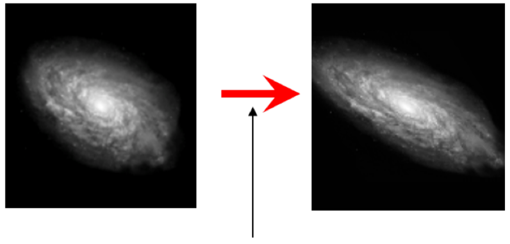


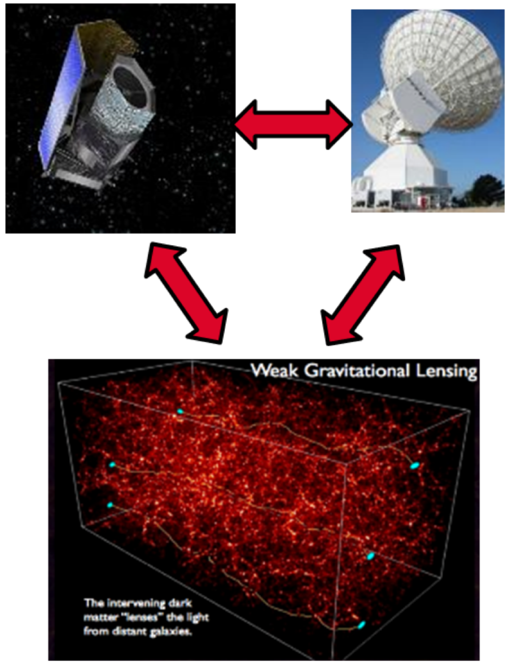
HPC for AI   AI for HPC   AI for data analytics

EUCLID: SEARCHES THE SIGNATURES OF THE DARK ENERGY AND MATTER

The dark matter and energy can be studied by looking at:

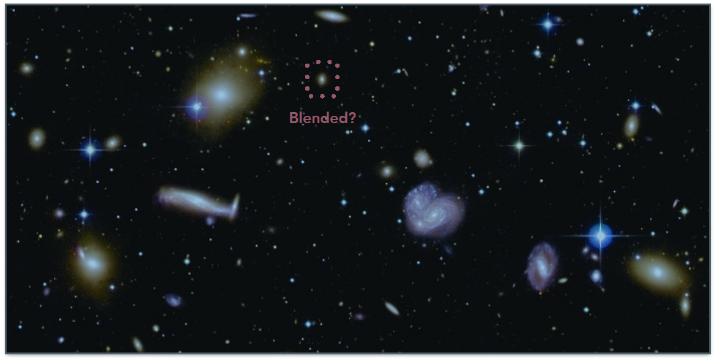
- The geometry of the universe
  - Measure of position of galaxies as a function of redshift
- Growth of density perturbations
  - Evolution of structure as a function of cosmic time, growth rate
- Galaxy image distortion caused by dark matter bending light



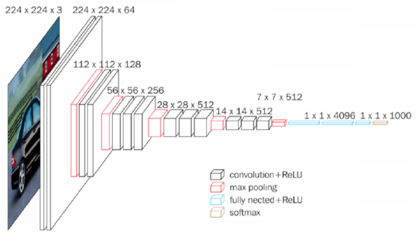


How to measure the dark matter:

- EUCLID: measure image distortion
- Cross observations from the ground: measure distances between galaxies
- Correlate these interpreted observational data with numerical simulations
- Use of DL to analyze the images



VGG-16

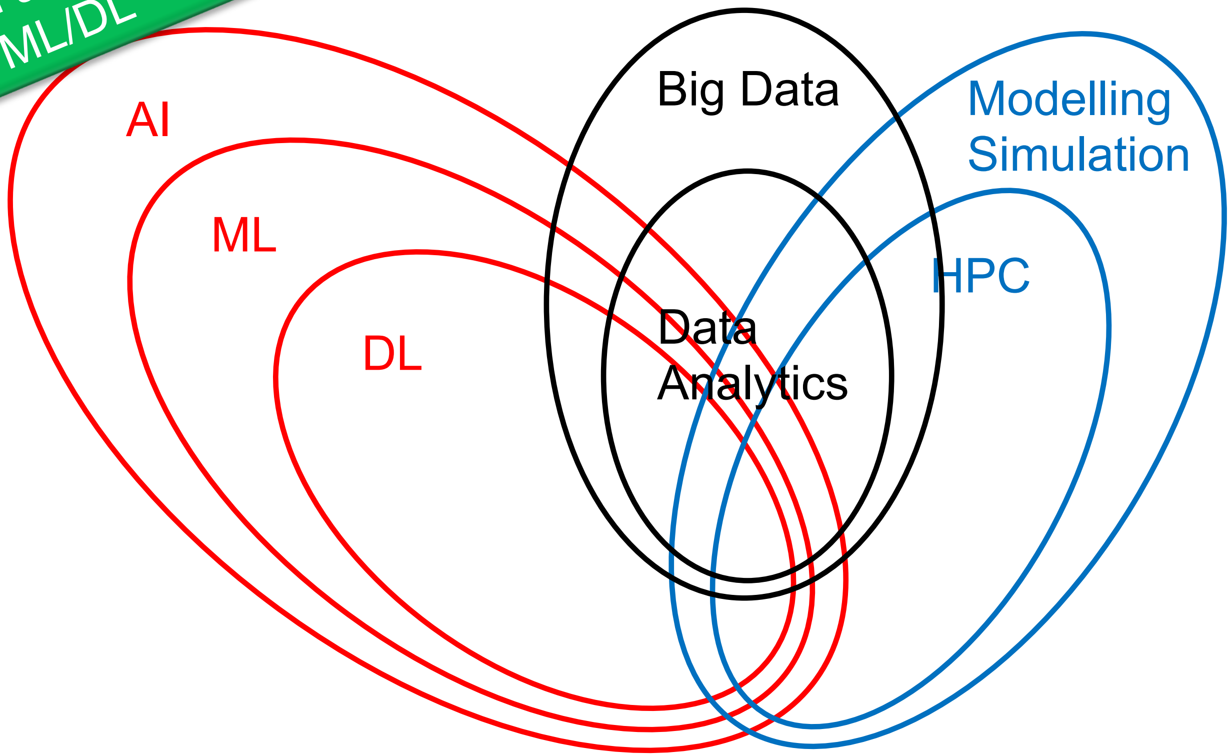


Simonyan & Zisserman (2018)

- Pre-trained network (e.g. using IMAGENET)
- Simple architecture.

CONVERGENCE BETWEEN HPC / BIG DATA / AI

GPU is, at now, viewed as the ideal architecture for ML/DL



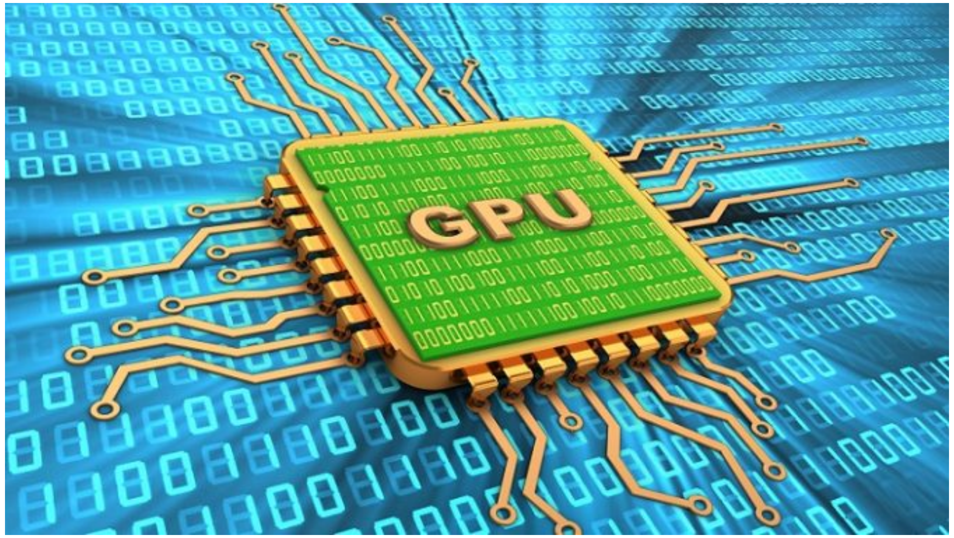
HPC for AI

AI for HPC

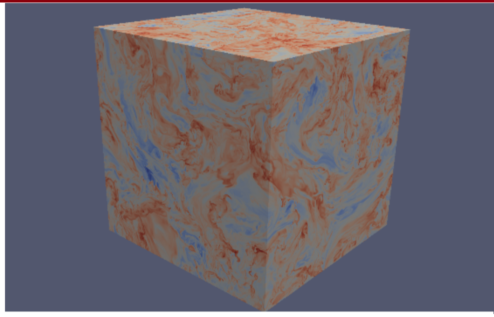
AI for data analytics

DE LA RECHERCHE À L'INDUSTRIE

EFFICIENT GPU PROGRAMMING



www.cea.fr

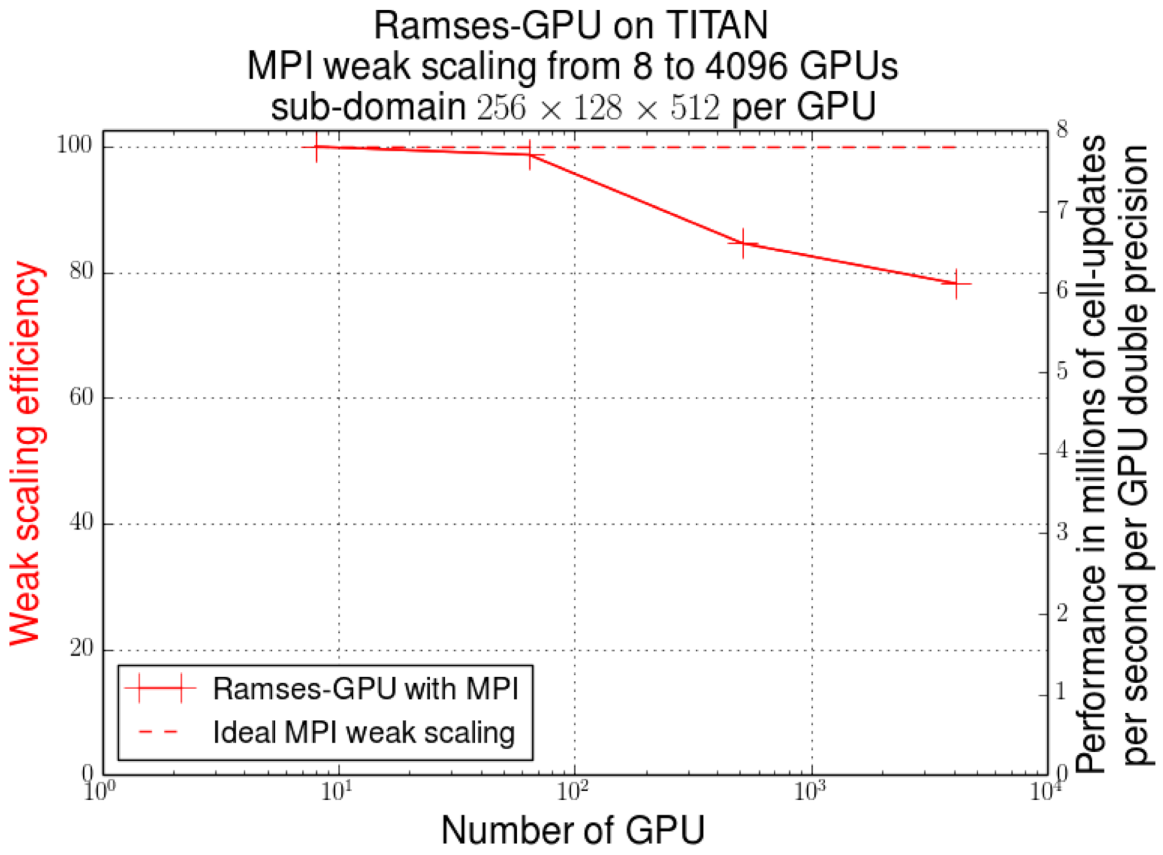


RAMSES

- Developped in Saclay to study large scale structure and galaxy formation.
- Flexible package to be used for general purpose simulations in self-gravitating fluid dynamics.

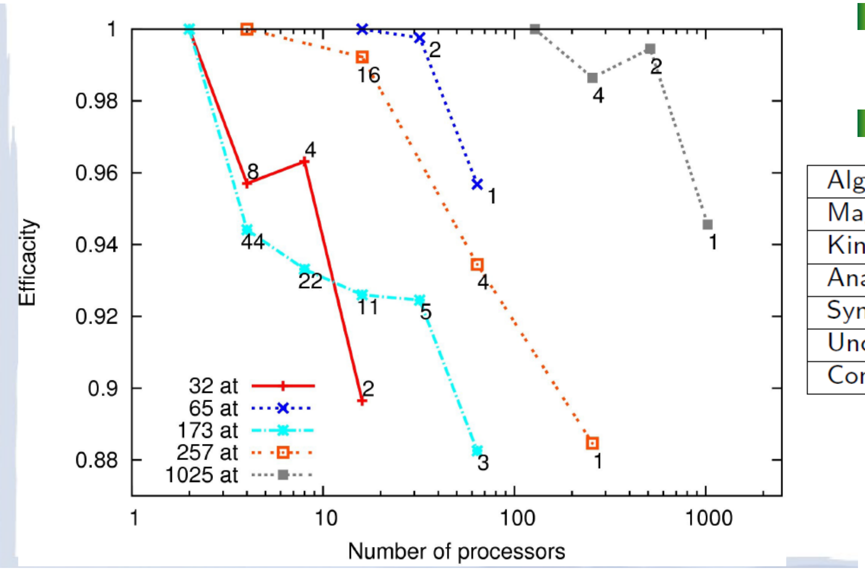
RAMSES-GPU (**MPI** and **GPU (CUDA)** parallelisation)

- Finite volume methods for compressible flows (**HYDRO** and **MHD** 3D cartesian grids
- **Directional splitting Godunov-based, 2nd order** for Hydrodynamics
- **Unsplit Godunov-based, 2nd order** for Hydrodynamics; Rieman solvers are *almost exact* (involves Newtow-Raphson iterations), *HLL*, *HLLC*
- **Unsplit Godunov-based, 2nd order** for Hydrodynamics + **Constraint Transport** for magnetic field update;
- Riemann solvers for **MHD**are *HLL*, *LLF*, *HLLD*
- **Dissipative terms** (viscosity, resistivity)
- **Self-gravity using a FFT-based Poisson solver**



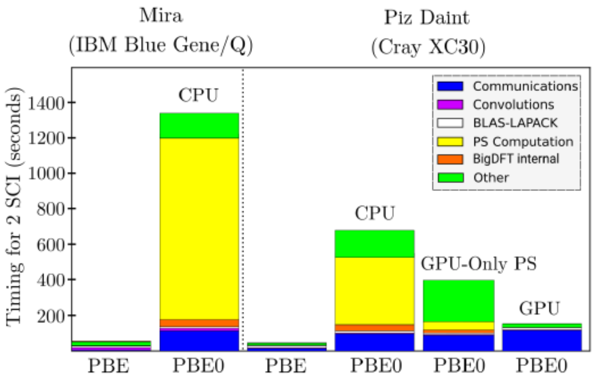
BigDFT

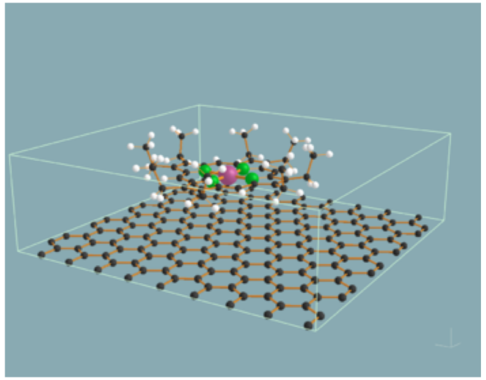
- DFT massively parallel electronic structure code (GPL license) using a wavelet basis set

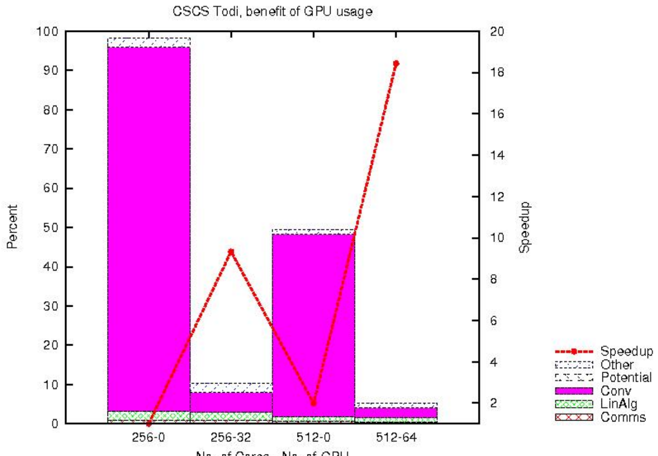


- First port on GPU NVIDIA using CUDA in 2009
- New version using OpenCL in 2010

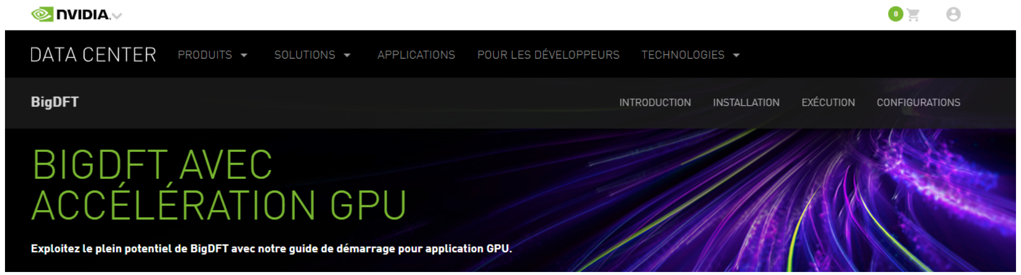
Algorithm	CPU	CUDA	OpenCL	CPU/OpenCL	CUDA/OpenCL
Magic filter	35.7	2.2	1.9	19	1.2
Kinetic	70.2	5.2	3.2	22	1.6
Analysis	33.7	2.3	1.5	22	1.5
Synthesis	43.8	2.3	1.6	28	1.5
Uncompress	195	20.3	10.5	19	2.1
Compress	100	13.6	7.6	13	1.8



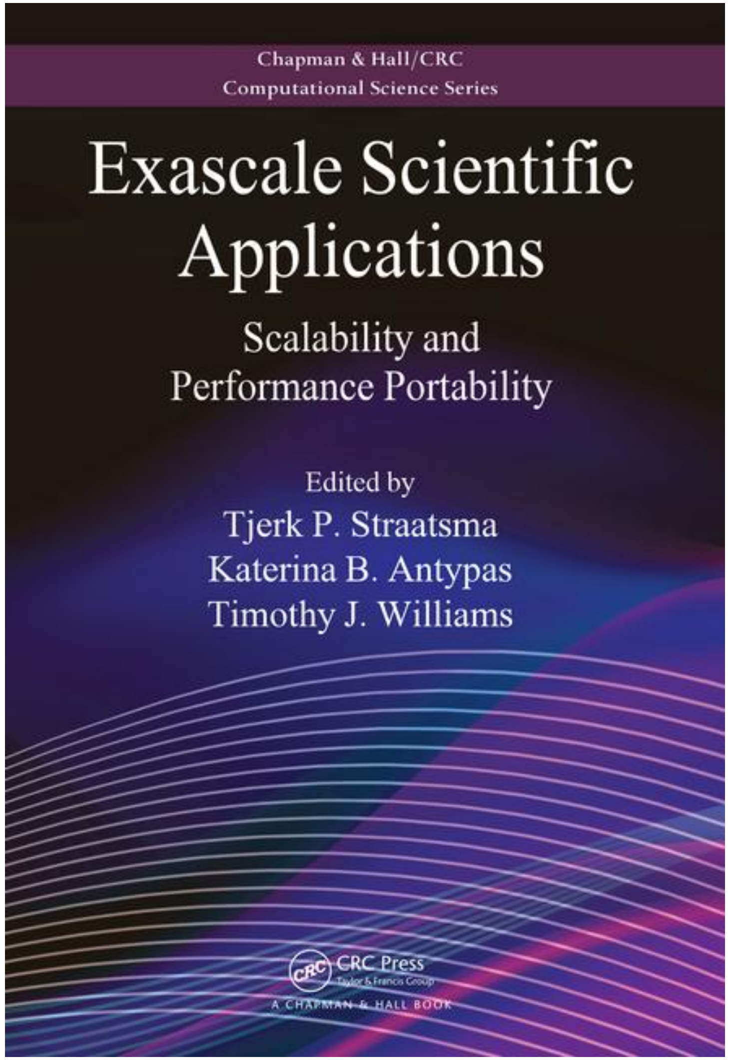




- Included in a docker-container
- Could be run on a Bare Metal server or on NVIDIA GPU Cloud



DE LA RECHERCHE À L'INDUSTRIE



FROM OPTIMAL PERFORMANCES TO PERFORMANCES PORTABILITY

www.cea.fr

Multi-architecture implementation of Spectral Difference Methods (SDM) for compressible flows using performance portable programing tools (c++ kokkos)

Pierre Kestener

CEA Saclay, DRF, Maison de la Simulation, FRANCE

Seminar CERFACS, Toulouse, February 15th, 2019

PERFORMANCE PORTABILITY

- **Developing / maintaining** a **separate** implementation of an application for each **new hardware platform**(Intel KNL, Nvidia GPU, ARMv8, ...) is **less and less realistic**

- **Identical code** will never perform **optimally** on all platforms
- Is it possible to have a **single set of source codes** that can be compiled for different hardware targets ?
- Performance portability should be understood as a single source code base with
  - **good** performance on different architecture
  - a relatively **small amount of effort** required to tune app performance from one architecture to another.

source http://www.nersc.gov/research-and-development/application-readiness-across-doe-labs

- **Low-level native language:** OpenCL, CUDA, ...
- **Directive approach (code annotations)** for multicore/GPU (OpenMP 4.5, OpenACC 2.5)
- **Other high-level library-based approaches** (mostly c++-based, à la TBB):
  - Some provide STL-like algorithmics patterns (Thrust, lift, arrayFire)
  - Kokkos, RAJA, ...
  - Cross-platform frameworks (Chamm++, hpx)
- **Use an embedded Domain Specific Language (DSL) (**Halide, NABLA)

RAMSESGPU WITH KOKKOS

Directionally splitted Hydro - 3D performances

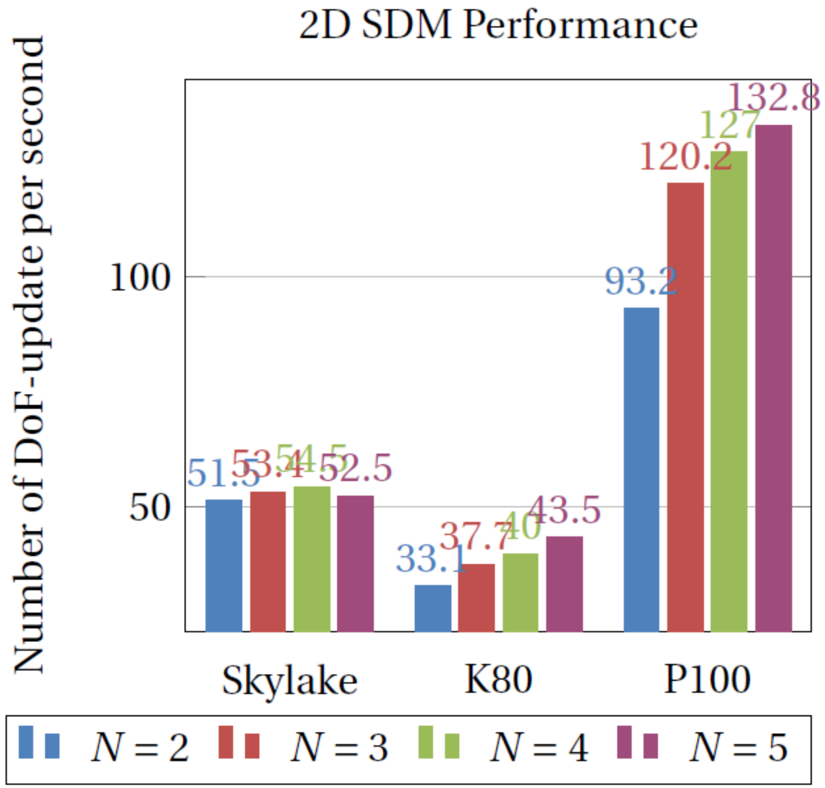
Number of 10^6 cell updates / second versus domain size

taille	M2090 SP/fast	SP	DP	K20 SP/fast	SP	DP
32x32x32	18.8			26.6 (+41%)		
64x64x64	83.4			95.2 (+14%)		
96x96x96	100.7			175.1 (+73%)		
128x128x128	114.7	32.1	9.2	178.7 (+55%)	72.4	34.9
192x192x192	133.0			226.7 (+70%)		
225x225x225	137.2	39.5	11.1	210.5 (+53%)	97.4	40.6

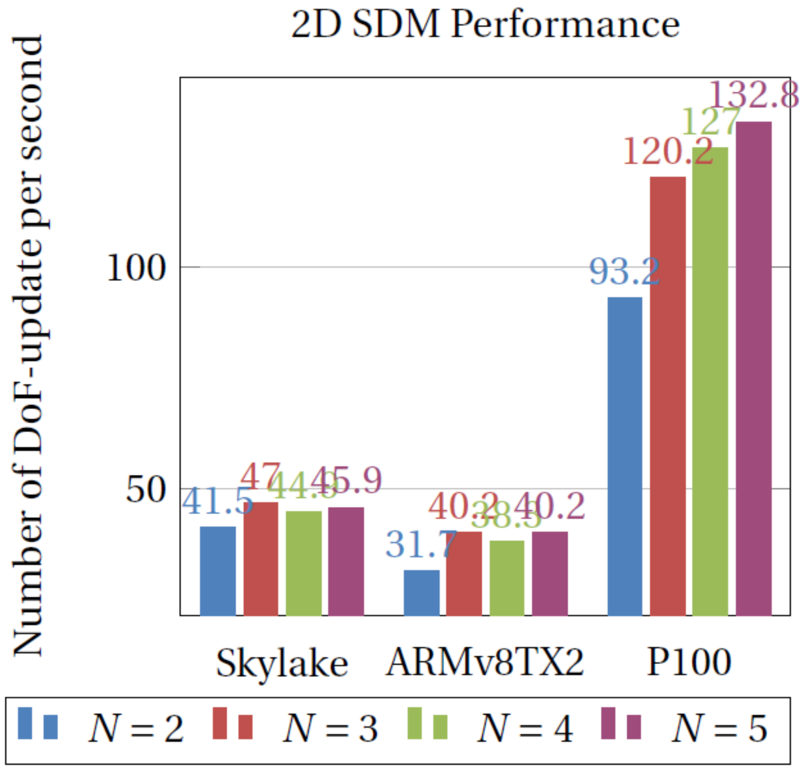
- Architecture Kepler K20 *versus* Fermi M2090
- Rebuild application with CUDA 5.5 toolchain for architecture 3.5 and tune flags
- Tune max register count for Kepler, and care about *read-only data cache* ==> **No more register spilling, DP perf is optimal !**
- in DP : Kepler is ~ 3.5× **faster** than than M2090

- With Kokkos, from Kepler K80 ⇒ Pascal P100, **performance scaling almost perfect** (~ ×3.0); no tuning required ; 360 Mcell-update/s
- With hand-written CUDA, tuning is required to recover this perf scaling
- Number of lines of codes divided by 2-3
- Get for free an efficient OpenMP implementation

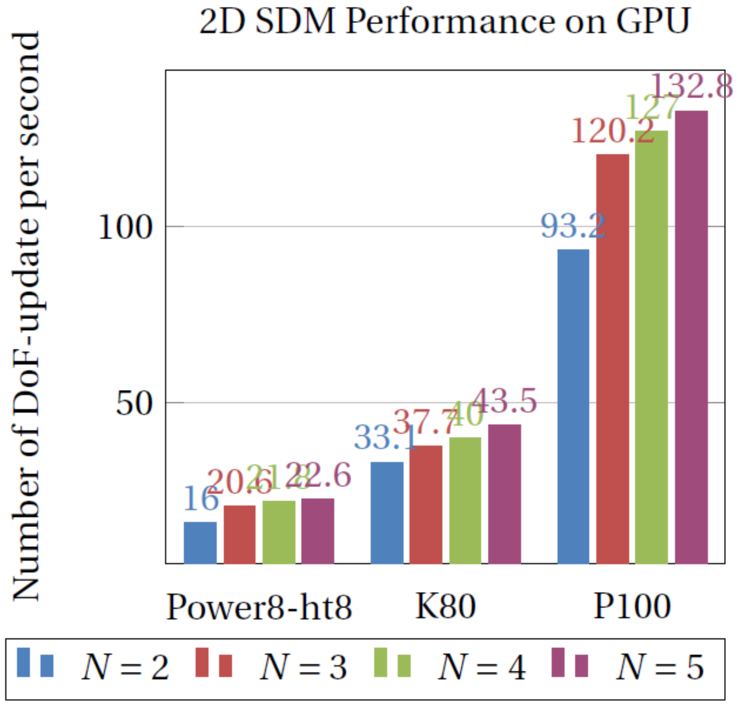
SDM PERFORMANCES ON VARIOUS ARCHITECTURE USING KOKKOS



SKL vs GPU



SKL vs ARM vs GPU



Power8 vs GPU

Illustration on MC neutron transport code (PATMOS – CEA/DEN)

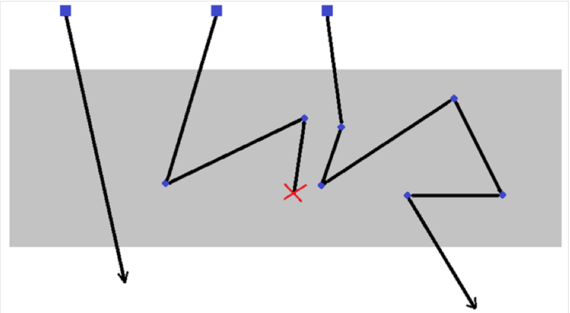
- Main kernel: XS calculation
- 2 benchmarks: S1Bench & PATMOS
- 2 versions: History Based (HB) & Pseudo Event Based (PEB)

Model	Compiler	S1Bench		PATMOS	
		HB	PEB	HB	PEB
MPI+OMPth	GCC,PGI,XLC,Clang	√	√	√	√
MPI+OMPth+CUDA	GCC,PGI,XLC,Clang	√	√	√	√
MPI+OMPth+OACC	PGI	√	√	√	√
MPI+OMPth+OMP	XLC	√	√	√	√
MPI+OMPth+Kokkos	GCC	√	√	√	√
MPI+OCCA	GCC	√	√		

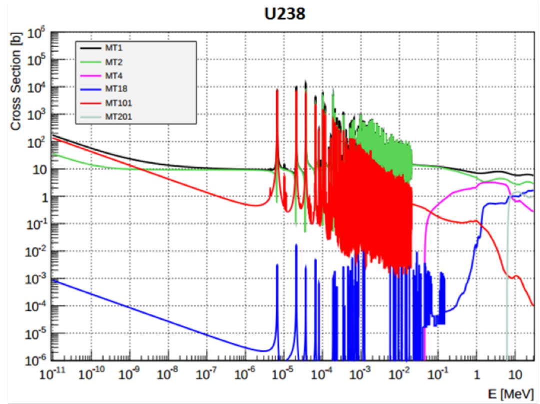
Table: Programming model

The principle of MC transport for Neutrons :

- **Follow each particle** from birth to absorption or leakage (history).
- Governed by **interaction probabilities** described by microscopic **cross sections**



a)No collision
b)Absorption
c)diffusion



Cross sections depending on energy

T. Chang, E. Brun, F. Malvagi, C. Calvin

PORTABITLITY METRICS

Metric

- $p = \frac{L_{TDA}}{L_T}$, where
  - L_{TDA}: Total lines used across different architectures
  - L_T: Total lines

Metric

$$\boldsymbol{P}(H) = \frac{|H|}{\sum_{i \in H} \frac{1}{e_i}}, \text{ where} \begin{cases} H & \text{a set of platforms} \\ e_i & \text{performance efficiency on } H_i \end{cases}$$

{H} = {CPU, GPU}			
Version	L_{TDA}	L_T	Portability (%)
OMPth	0	247	0
CUDA	0	237	0
OpenACC	265	265	100
OpenMP	263	263	100
Kokkos	10	275	96
OCCA	2	270	99

Difficulty

- Hard to measure performance efficiency e_i:
  - Peak performance, such as GFLOPS
  - It is difficult to measure performances of S1Bench and PATMOS in terms of GFLOPs

DE LA RECHERCHE À L'INDUSTRIE

CONCLUSION

CONCLUSION

- It is not a novelty: architecture are more and more complex
- Which is new: workflow are more and more complex, mixing HPC, HPDA, AI …
- Data is becoming the corner stone of the system (and system of systems – edge computing, fog computing …)
- Convergence between HPC, HPDA and AI is coming from usages, not from HW or SW solutions

→ **from a user point-of-view the landscape is becoming more and more complex**

- From the HPC point-of-view (which has been partially addressed in this talk), one major concerns is performance portability: **provide to developers framework, tools, methodologies to deal with all these complexities in a constantly evolving environment**.

DE LA RECHERCHE À L'INDUSTRIE

THANK YOU !