Some (Post-)Exascale Challenges
The NumPEx Program and InPEx Project Contributions

KeyNote 2 - Long-term Computing Vision

Feb 13 2024, 2024 ETP4HPC Conference

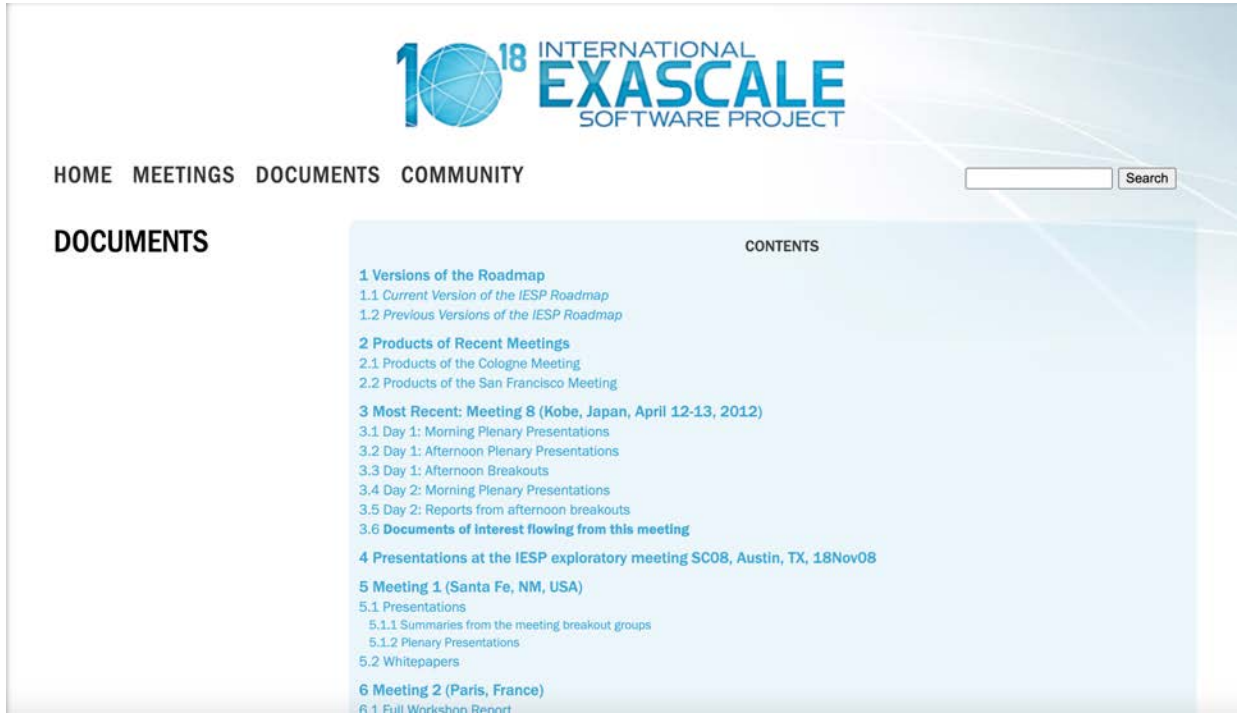
G. Antoniu (INRIA), JY. Berthou (INRIA), J. Bobin (CEA),

Contributors : M. Krajecki (CNRS), C. Prudhomme (U. de Strasbourg), Hélène Barucq (Inria), R. Namyst (Inria/U. de Bordeaux), A. Buttari (IRIT), J. Bigot (CEA), F. Bodin (U. de Rennes), Mark Asch (U. Picardie), T. Deutsch (CEA), J-P. Vilotte (CNRS), V. Brenner (CEA), B. Raffin (Inria)

15 years ago, Simulating The Future'08
One million core programming workshop

Some international Exascale efforts …

IESP 2008, …2012

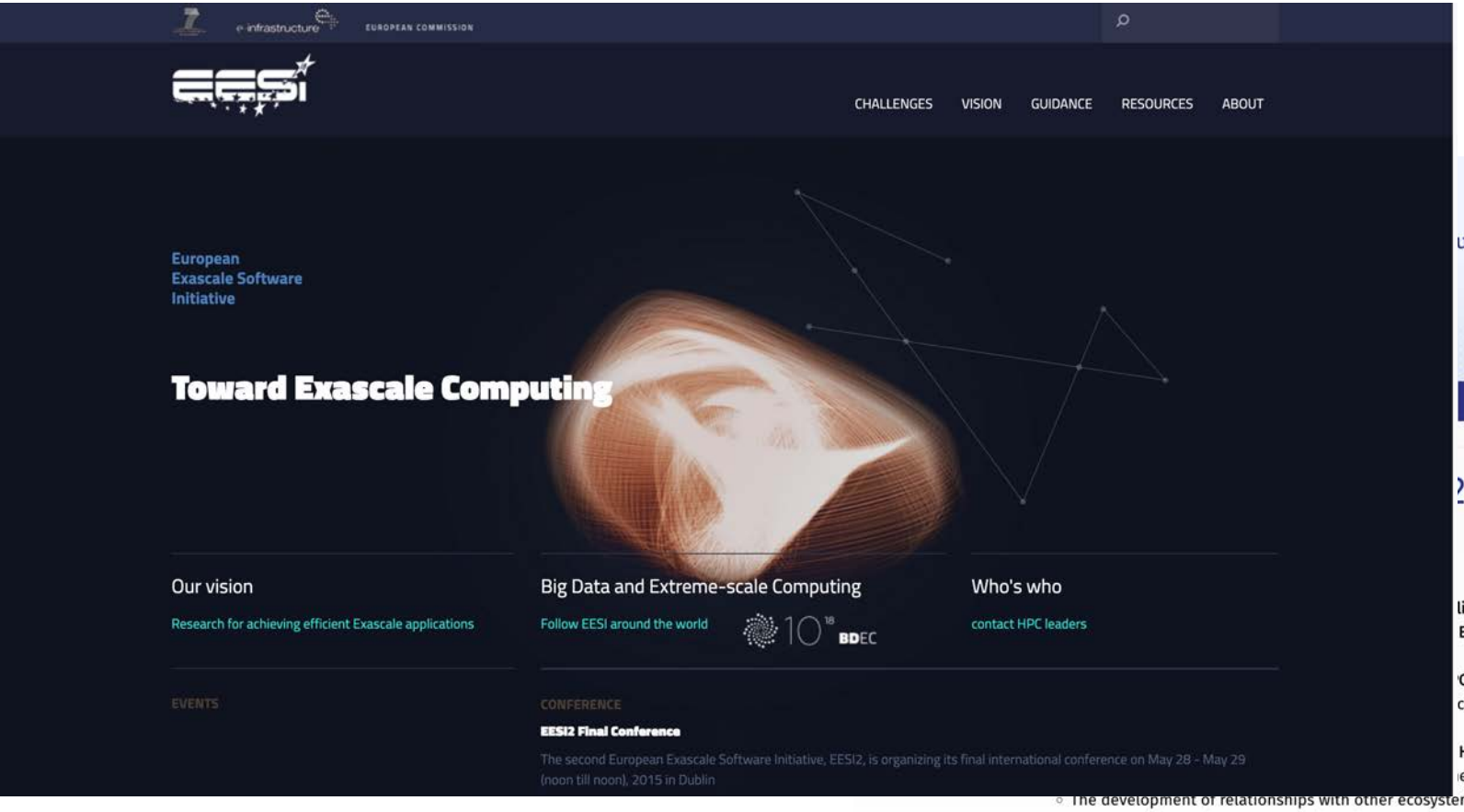


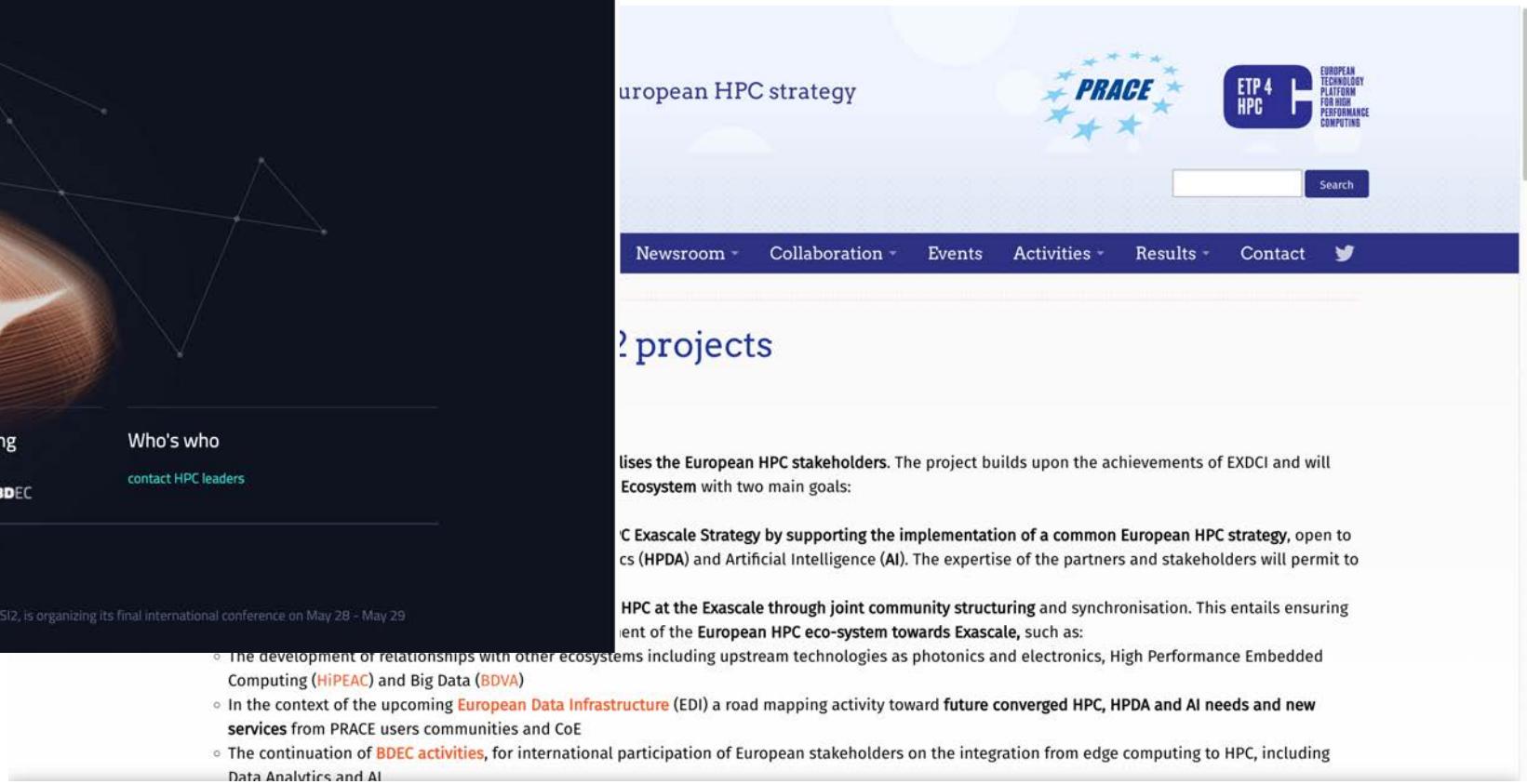
BDEC 2013, 2014 …2020

Some international Exascale efforts …

With coordinated European contribution (CSA)

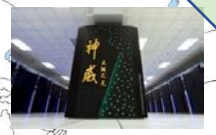
EESI 1&2, EXDCI 1&2 2010, …2020

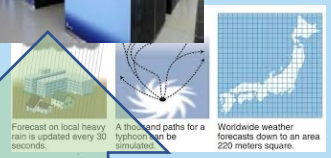




The (post)-Exascale race, where are we?

China initiatives:
- development of applications in preparation for the arrival of the Tianhe3 machine.





ECP : DOE funded, NSF support (end dec. 2023)
+ Creation of 6 co-design centers
Still a challenge: Exascale ready app, sustainable software stack
New perspectives: Gen IA for Science, Trillion Parameter Consortium, 20B€

Japan initiatives:
- FugakuNext prod. 2029 (xB$): co-design HW/SW/Apps
- Next Gen AI (x100M$), Quantum-HPC (140M$)
- (post)-Exascale as a Service (AWS/Rikean): from Fugaku to Virtual Fugaku

A strong effort in both hardware, software and applications/co-design

The (post)-Exascale race, where are we?

The International Post-Exascale (InPEx) Project

What to share and coordinate…

- AI for Science, HPC for AI - Datasets for AI training
- Software Exascale stack production, management and sustainability :
  - Packaging, documentation, builds, results, catalogs, continuous integration, Linux builds (E4S, etc.)
- Digital continuum and data management
- Co-design (HW&SW&Apps), benchmarks and evaluation
- Math and algorithm for Exascale
- Energy and environmental impact and sustainability
- Future and disruptive technologies and usage (investment
- Training


Pre-workshop InPEx, October 2023, Reims, F

The International Post-Exascale (InPEx) Project

InPEx expected outcomes

- Landmark documents largely exploited, worldwide, for supporting future post-exascale science
- Contribute to the implementation of an international, shared, high-quality computing environment based on the principles and practices of co-design
- Formation of a solid network of exascale computing leaders, all around the globe

Actions:

- Dedicated international working groups
- International Post-Exascale (InPEx) workshop series

Participants:

Researchers, engineers, industry, funding bodies

Why InPEx, June 5, 2023, Paris

The International Post-Exascale (InPEx) Project

Inpex.science

Date	(10/2023)	11/2023	06/2024	06/2025
Location	Preparatory phase EU (France)	SC'23 - BOF	Workshop1 EU/BSC	Workshop2 Japan
Date	03/2026	09/2026	06/2027	09/2027
Location	Workshop3 US	Workshop4 EU	Workshop5 Japan	Workshop6 US

How to contribute? Send 2 pages white paper to : inpex@inpex.science

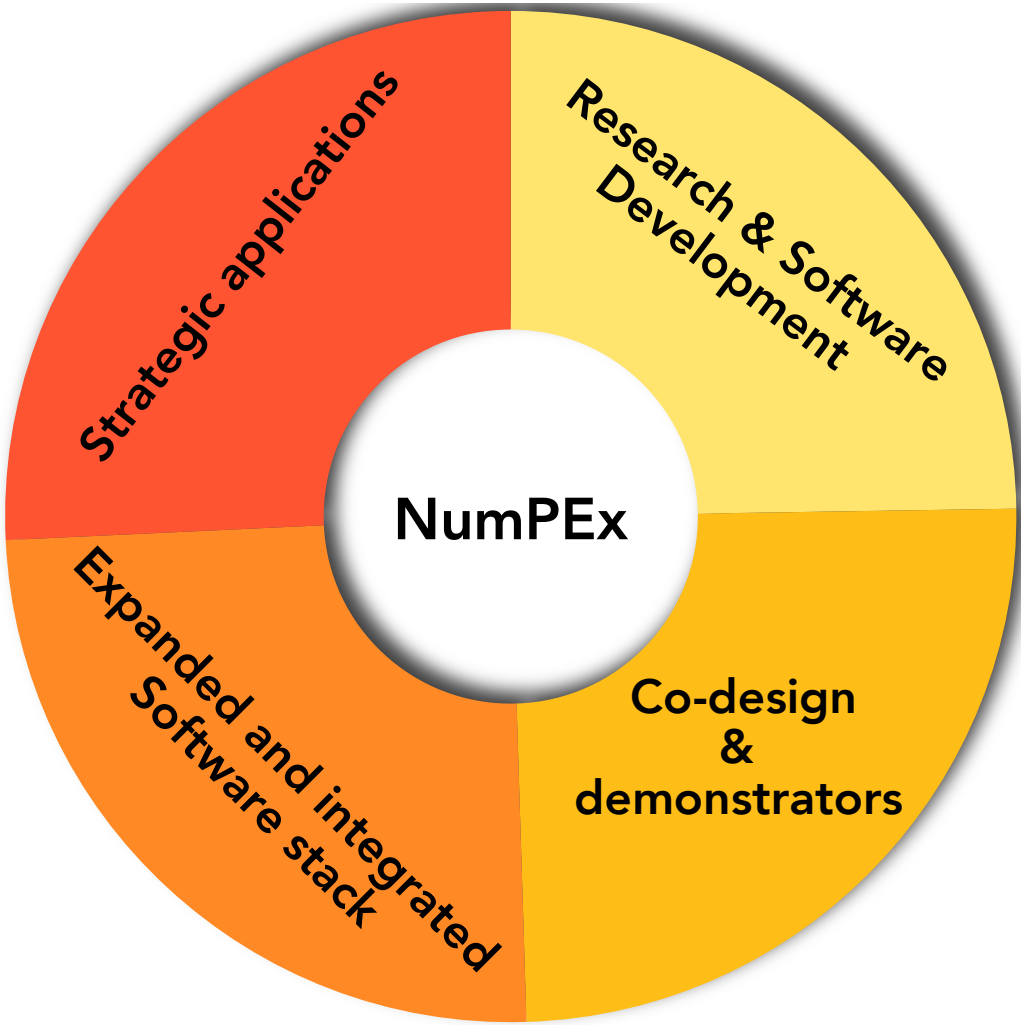
The French NumPEx Program

Consolidating and accelerating the construction of a sovereign European **exascale software stack** and **strategic applications exascale capability** in a **coherent and multi-annuel framework**

Integrate and validate **co-designed** innovative methods, libraries and software stack with demonstrators of strategic applications.

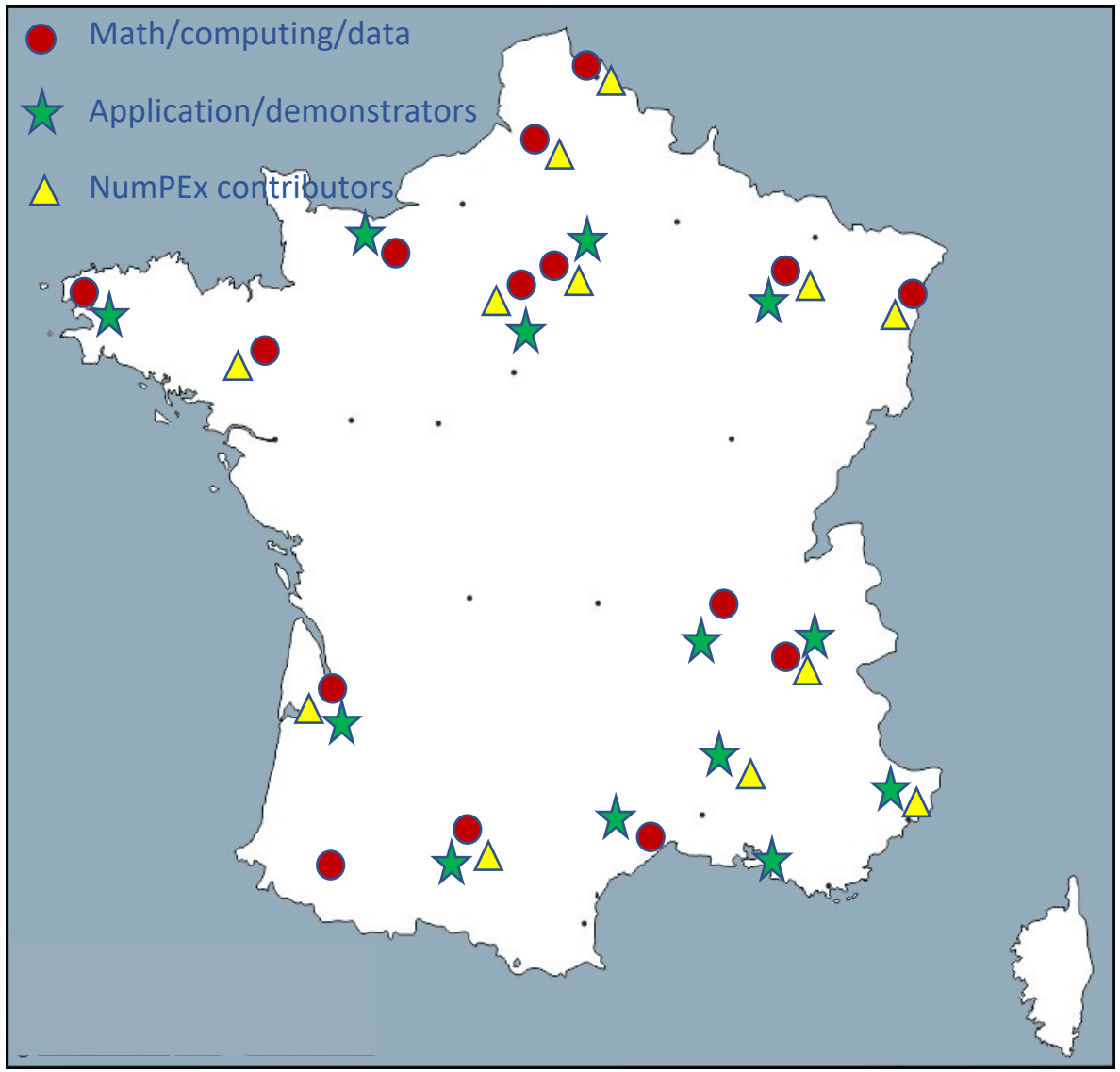
Accelerate science-driven and engineering-driven developers **training and software productivity**

Foster national and international collaborations to prepare for the Exascale and post-Exascale era



Help aggregate the French Edge/Cloud/HPC/HPDA/IA community

NumPEx by numbers



6 Years
41 M€*

2023-2028
* Funding 41M€=500 person.year non permanent staff
+ 170 person.year permanent staff
Total cost : 81 M€

Core Research Institutions

Core national Research Institutions: CNRS, CEA, INRIA, Universities, Engineer schools, Industry

3 Focus Area

Software stack development (PC 1-3)
Wide-area workflows and architecture (PC 4)
Integration and application development (PC 5)

80 R&D teams
500 Researchers

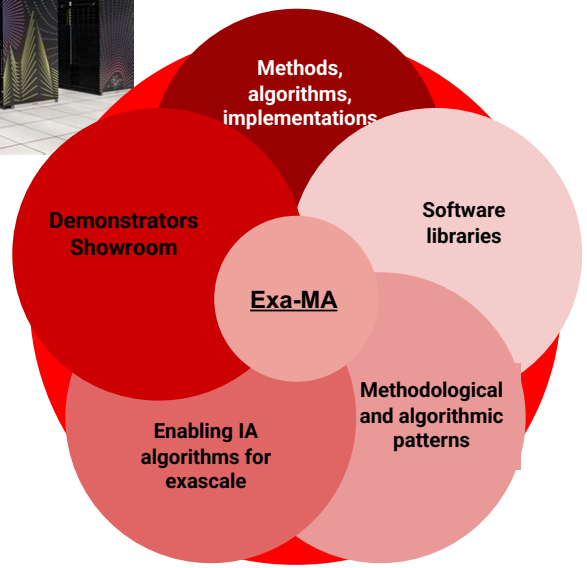
The French NumPEx Program - workplan

Applications

ExaMA

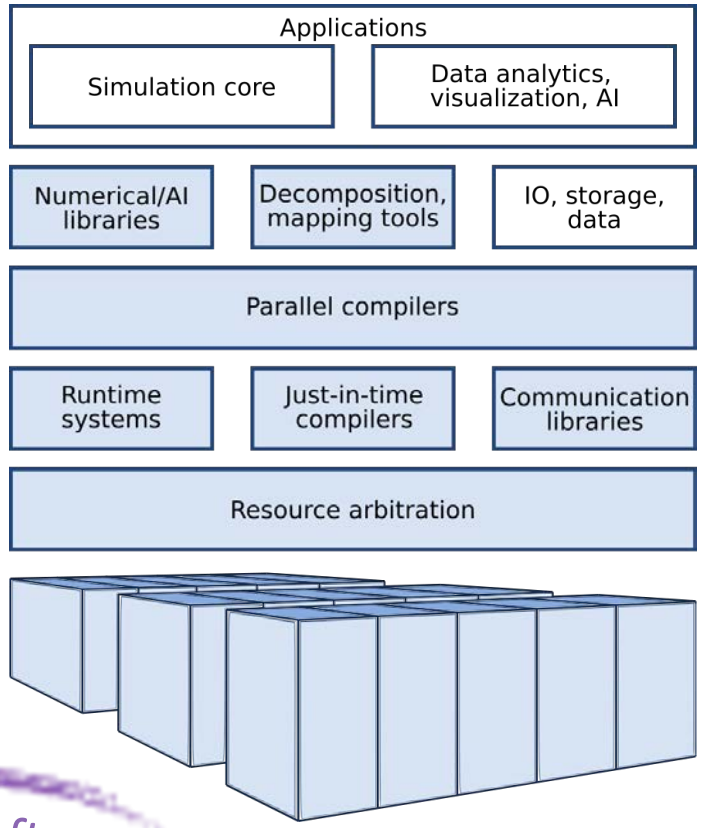
Numerical methods and solvers

C.Prudhomme/H.Barucq



The French NumPEx Program - workplan

Applications



ExaMA

Numerical methods and solvers

C.Prudhomme/H.Barucq

ExaSoft

Computing

R.Namyst/A.Butari

The French NumPEx Program - workplan

Applications

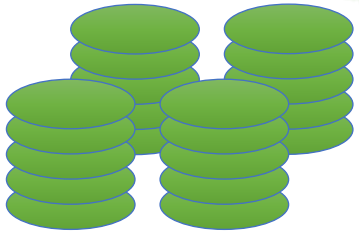
WP1: Exascale I/O and storage

WP2: Exascale in-situ data processing

WP3: Exascale ML-based data analytics

WP4: Shared building blocks & integrated illustrators

WP5: Management, dissemination and training



ExaDost
Data
G.Antoniu/J.Bigot

ExaMA
Numerical methods and solvers
C.Prudhomme/H.Barucq

ExaSoft
Computing
R.Namyst/A.Butari

The French NumPEx Program - workplan

Applications

ExaAtow
Digital continuum

F.Bodin, T.Deutsch, M.Asch

ExaDost
Data

G.Antoniu/J.Bigot

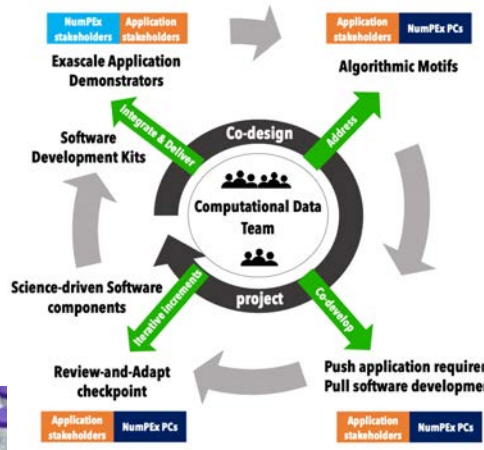
ExaMA
Numerical methods and solvers

C.Prudhomme/H.Barucq

ExaSoft
Computing

R.Namyst/A.Butari

The French NumPEx Program - workplan



Applications

ExaDI

Application co-design and software integration

JP.Vilotte/V.Brenner

ExaMA

Numerical methods and solvers

C.Prudhomme/H.Barucq

ExaAtow

Digital continuum

F.Bodin, T.Deutsch, M.Asch

ExaSoft

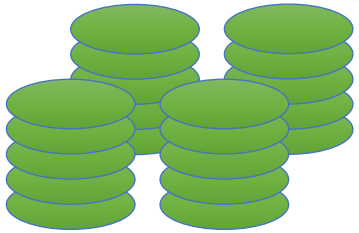
Computing

R.Namyst/A.Butari

ExaDost

Data

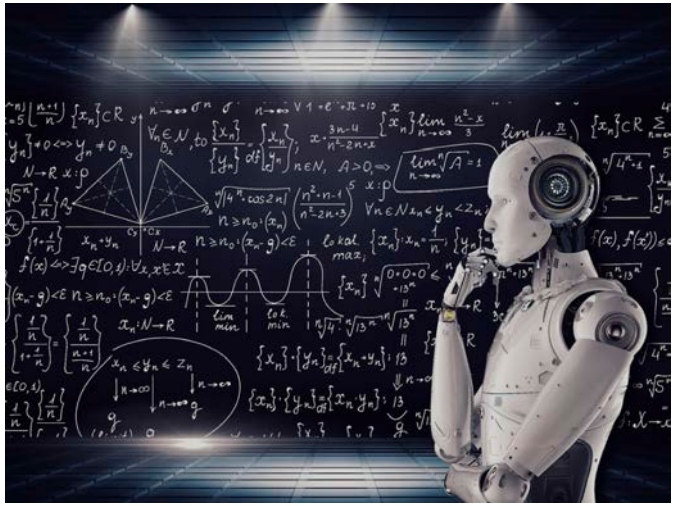
G.Antoniu/J.Bigot



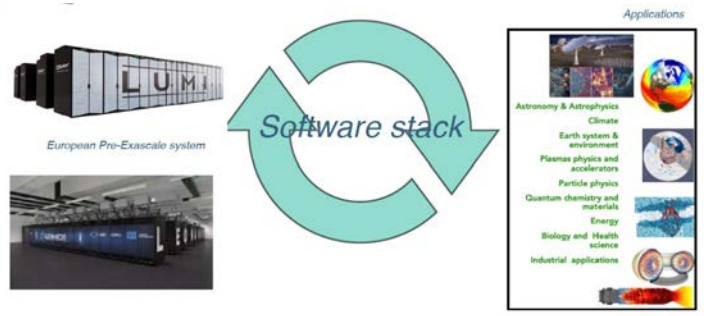
Focus on some post-exascale challenges

From edge to HPC systems
The digital continuum

AI4Science – Science4AI



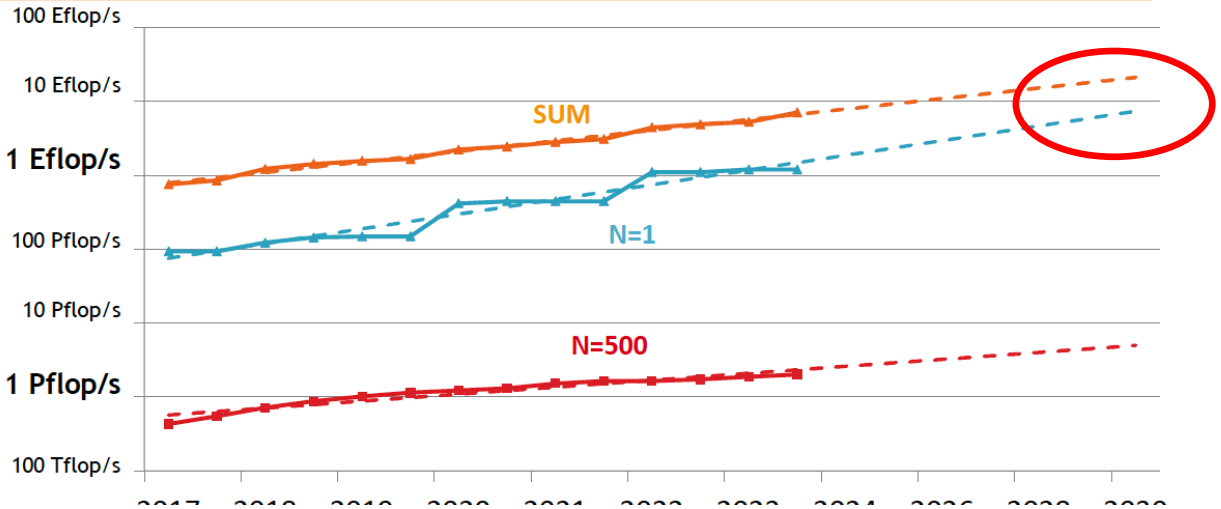
Software/application co-design



Software, the new frontier

Software, the new frontier

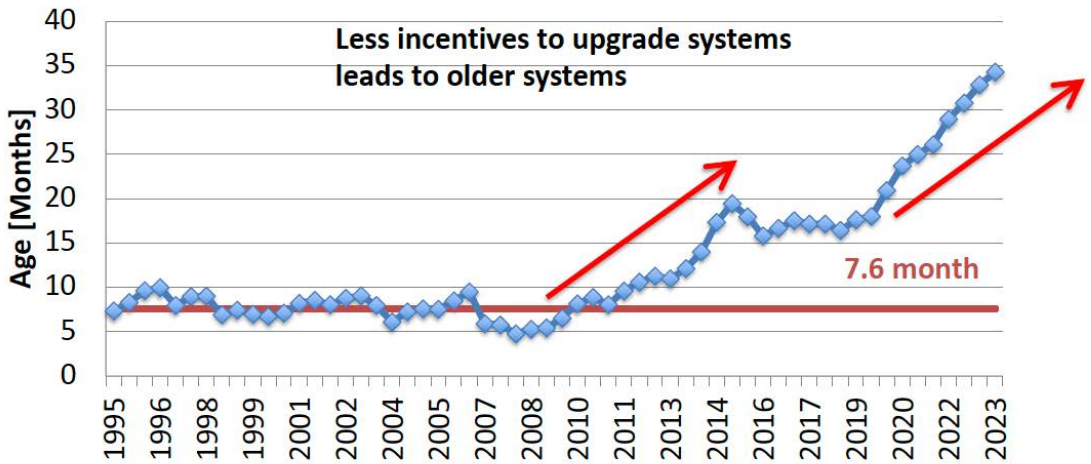
PROJECTED PERFORMANCE DEVELOPMENT



SC23 TOP500 HIGHLIGHTS

- Frontier is #1 and remains only ExaScale system in the TOP500 - for now
  - 4 new systems in TOP10! + one upgrade
  - And 9 additional new systems in TOP50
- First Intel Sapphire CPUs (25) and GPUs (4) on the list
- HPC systems are used longer and replaced less often
  - Due to an increasing number of technological limits
  - Leads to strong concentration at the top (research as well as commercial)
- TOP500 shows new reduced growth-rates since 2017!
  - End of original Moore's Law scaling
  - Unlikely to achieve 10 Exascale by the end of the decade unless we fundamentally change "business"!

AVERAGE SYSTEM AGE



Software, the new frontier

Software is the new power

- The race for power is slowing down: we're aiming for 10 Exaflops in 2030, not 1000

=> Importance of **Math&software developpment to gain effective power**: new digital schema, precision mix, innovative discretization methods, energy aware algorithms, …

- Capitalize on and consolidate software production in Europe and worldwide
- Importance of software sustainability
- Offer an "industrial" production environment converging towards "As a Service".
- Broaden the potential user base towards new AI/HPC/HPDA/Cloud/Edge uses
- Enable and facilitate, or even make possible, what was not possible before, through the coupling of components produced in different contexts (e.g. Kokkos/StarPU/HWLOC) and thus enable the emergence of new services, uses and capacities to move towards exascale

At stake: consolidating and accelerating the construction of a **sovereign European exascale software stack**

Software, the new frontier

Software Integration challenge

Context:

- **HPC applications and machines are gaining in complexity** (think compute continuum, HPC/HPDA/DL hybridization), leading to high costs to build and deploy applications on supercomputers, **impairing portability and reproducibility**.
- The standard software installation process on supercomputers is reaching its limits

  module load + cmake/make install

At Stake:

- Need for a HPC dev-ops methodology for fast deployment while ensuring performance and supporting reproducibility efforts

NUMPEX Strategy: Modern Software Packaging (SPACK and GUIX)

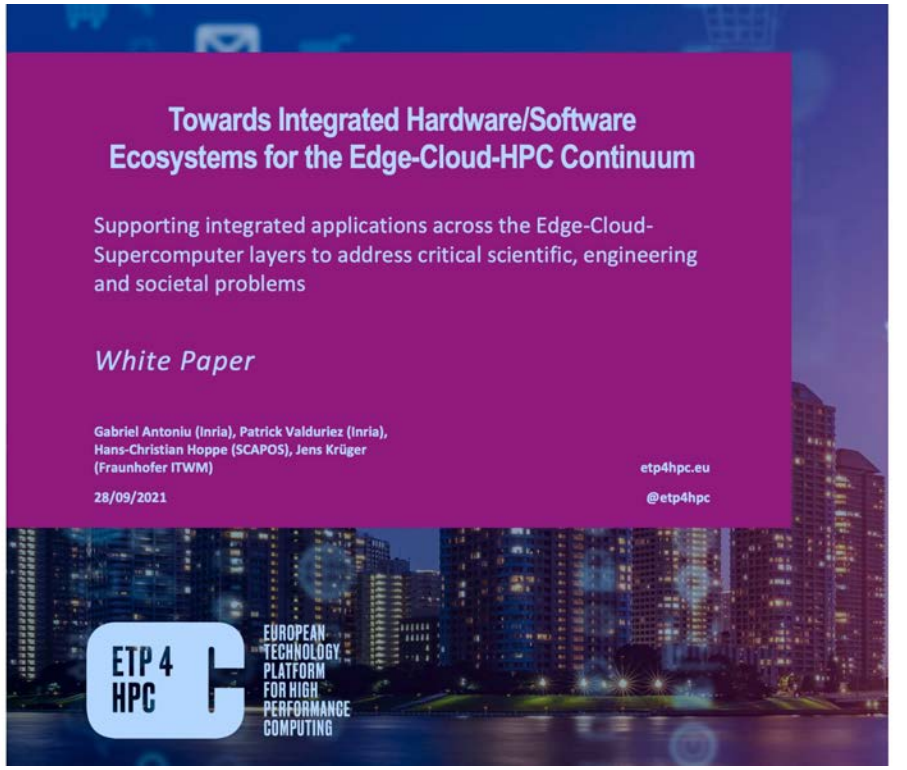
- Have users work with package managers to capture software dependencies with compilation process to automate deployment.
  - Can generate containers or be used directly if installed on supercomputer.

    guix pack -f docker mysoft

    guix install mysoft
  - Can also be used by system admin to generate modules

The Edge-Cloud-HPC continuum

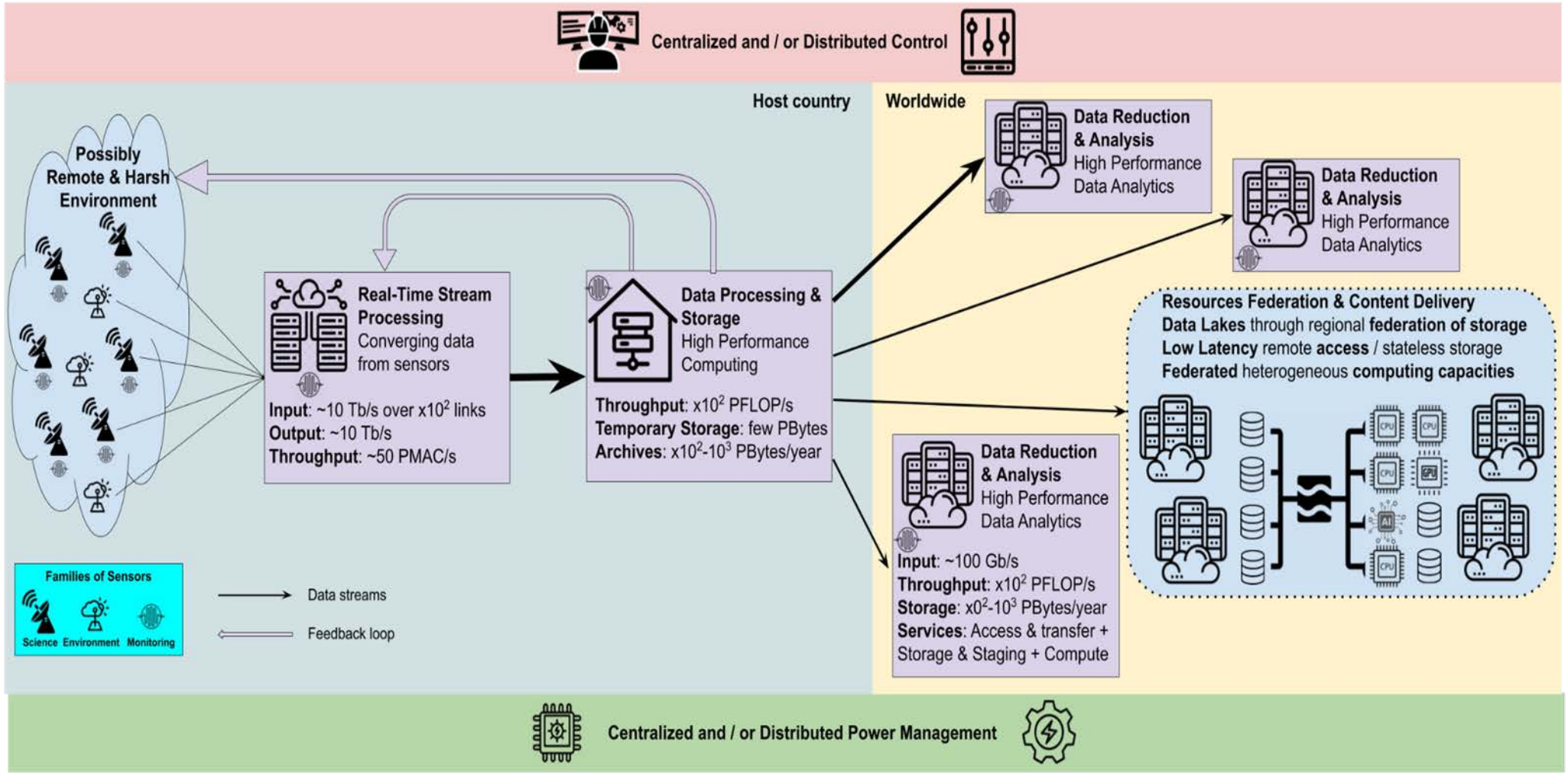
Distributed workflows
across edge, cloud and HPC systems



TCI
TransContinuum Initiative

- The Exascale machine is part of a “datasphere”
  - Data from large scientific instruments
  - Data from sensors (IoT)
  - Data from simulations
- The primary challenge: developing integrated software ecosystems across this continuum
  - At an extreme scale with stringent cybersecurity constraints

The SKA data workflow from sensors to HPC centers



Observation data product delivery
- Data streaming reduction and processing,
- Edge computing and content delivery network

Data Science analysis: R&D component & software development
- Emerging technologies compute, I/O, storage, wide area workflows
- Distributed & heterogeneous architecture, content delivery network
- Power management

The digital continuum: open challenges

Unification of HPC Simulations/Big Data/AI towards a data-centric view

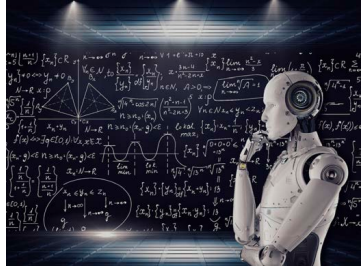
Moving, storing and processing data across the continuum: how to deal with the 3 Vs of Big Data?

- **Extreme Volume across the continuum**
  - Support the access and processing of “cold”, historical data and “hot”, real-time data + (virtually infinite) simulated data
- **Extreme Velocity across the continuum**
  - Unified real-time data processing (in situ/in transit, stream-based) in a common software ecosystem
  - Need disruptive reduction in data movement cost with new devices, packaging
    - All real-time data may not be storable in archives => real time training (bandwidth-oriented)
- **Extreme Variety across the continuum**
  - Unified data storage abstractions to enable distributed processing and analytics across the continuum
  - Interoperable data formats, "Semantic interoperability” through shared ontologies
- Digital Continuum is a multi-tenant and multi-owner environment.
  - Collected data used with multiple purposes
  - Computing Infrastructure is also shared

The Post-Exascale digital continuum: action plan

- Coordinate efforts to share workflows solutions and services for the convergence of AI/HPC/HPDA/Cloud/Edge
- Develop the concept of EaaS, Exascale as a Service, for Tier-0 European systems
- Contribute to the development of a data-everywhere, FAIR, ecosystem in Europe
- Identifying new emerging usages including urgent computing and digital twins

AI4Science / Science4AI



AI everywhere, a game-changer at the post-Exascale era

- **AI at all stages** from numerical methods to workflows, etc.
- **Trend towards AI-centric HW**, which also puts **huge constraints on the HPC SW stack**

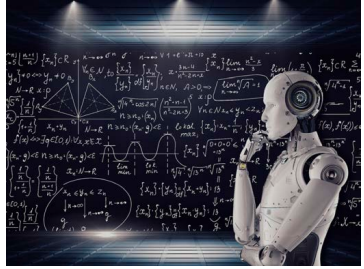
International initiatives have started, Europe must stay in the race

- US AI program, Trillion Parameter Initiative (www.tpc.dev).
- Initiatives to produce extreme scale LLMs, AuroraGPT, FugakuGPT, …
- **Pushing for European initiatives, with a strong focus on Science**

Challenges of HPC/AI convergence

- How to develop **validated/robust/trustworthy AI models** for Science
- In the post-Exascale era, data-centric era, should **benefit from AI-centric software**
- **HPC/AI software stack convergence ?**

AI4Science / Science4AI



AI for science – towards HPC/AI hybridization

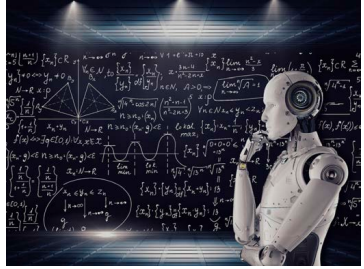
- End of Moore's law, **develop hybrid approaches based on AI** to accelerate parts of scientific computing applications

=> 10 Exascale performance with x25 EDP+Data science pp and x42 HW improvements, some AI-based solvers can be sped up by 6 orders of magnitude, etc., weather forecast with Graphcast

- **Hybridization of HPC SW with AI :** physics-informed AI models for simulation codes, observational data reduction, digital twins.

- **Push forward a post-Exascale-ready SW stack embedding AI solutions** that answer the needs of the application communities

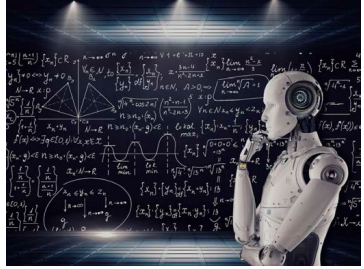
AI4Science / Science4AI



AI at scale at the post-Exascale era

- **Rapid emergence of extremely large AI models** (*e.g. LLM, multimodal GPT, diffusion models, etc*), which require huge compute capacity, massive storage.
- **Challenges in data management and computing** on Exascale/post-Exascale grade systems, with massive use of heterogeneous accelerated architecture.
- **Push a post-Exascale SW stack** that further covers AI-based libraries, with **interfaces between traditional HPC** librairies and standard **AI/Machine Learning framework**.

AI4Science / Science4AI



Accelerate AI for Science

- **Accelerate the diffusion and application of AI at all stages**, with a very strong focus on AI for scientific and engineering applications.
- Build upon an AI-centric co-design activity, focusing on **shared flexible open AI-centric application use-cases/mini-apps**, with curated data and well-defined application dependent metrics.

Fostering international initiatives and collaborations

- **Foster AI-centric European initiatives**, with a specific focus on **enabling the development of AI at scale for Science, not only focused on large fondation models**.

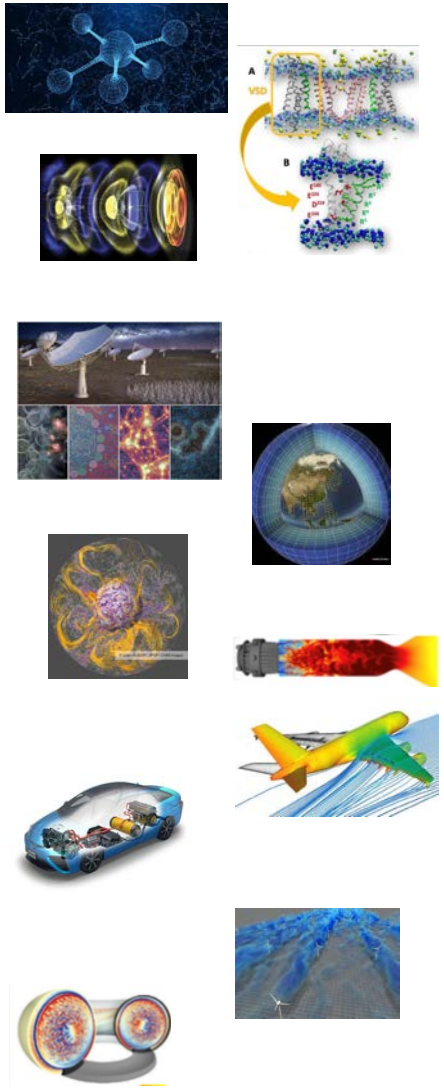
Software/application co-design

Key challenges

- How to get post-Exascale ready applications ?
- How to expand an application-driven SW stack ?
- How to make applications portable and sustainable at the post-Exascale era ?

International context

- Early-binding HW/SW/application co-design approach at Rikken (Japan)
- In the USA, DOE co-design centers were a key component of the Exascale Computing project (ECP)
- Inspired by ECP, co-design is central in the NumPEx project (FR)



Astronomy & Astrophysics

Earth System Models & Climate

Earth sciences & environment

Computational biology & Life science

Laboratory laser-plasma physics

High-energy particle physics

Quantum chemistry and materials

Digital health

Environmental & societal risks

Urban systems planing

Magnetically confined fusion plasma

Sustainable Transport & mobility

Energy production & transport

Software/application co-design

Help the applications get prepared for post-Exascale challenges

- Towards **shared and common mini-apps/proxy-apps**, centered about technical functionalities/bottlenecks (AMR, I/O, AI, etc.)
- **Shared benchmarks** are central to test/evaluate portability/performance/deployability on different HW.

Towards a post-Exascale application-driven software stack

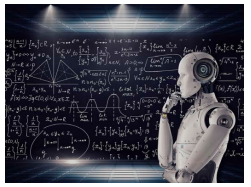
- Application code development is a long and costly journey, greater **reuse** of **composable/interoperable SW components**
- **Long-term visibility/sustainable development for applications is key.**

Take away messages

Software, the new frontier

Consolidating and accelerating the construction of a sovereign European **exascale software stack** (portable, interoperable, reproducible, sustainable)

Support and foster the developpement of disruptive Math & models



AI4Science – Science4AI

Push an **hybrid AI/HPC software stack**, to accelerate HPC and provide AI at scale

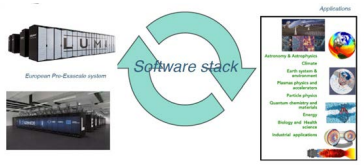
Support AI for Science, foster fully open AI use-cases/benchmarks, **not restricted to GenAI**

From edge to HPC system: the digital continuum

Coordinate efforts to share workflows, solutions and services for the convergence of HPC/Cloud/Edge

EaaS: **Exascale as a Service**, for Tier-0 European systems

Develop a data-everywhere, FAIR, ecosystem in Europe



Software/application co-design

HW/SW/application co-design to help the communities get prepared for post-Exascale

Foster the use/reuse of modular/interoperable and portable SW components

Push sustainable SW development model