RÉPUBLIQUE FRANÇAISE
Liberté
Égalité
Fraternité

Secrétariat général pour l'investissement

PROGRAMME DE RECHERCHE
NUMÉRIQUE POUR L'EXASCALE

HPC for AI @ exascale
Exa-DI workshop

ExaDI—workshop, Paris, 2-3 October 2024

https://www.numpex.org

J.P. Vilotte (CNRS) & V. Brenner (CEA)

anr® agence nationale de la recherche

The NumPEx Program

NumPEx: 6 years national project (CNRS, CEA, INRIA, Universities) - 40,8 M€

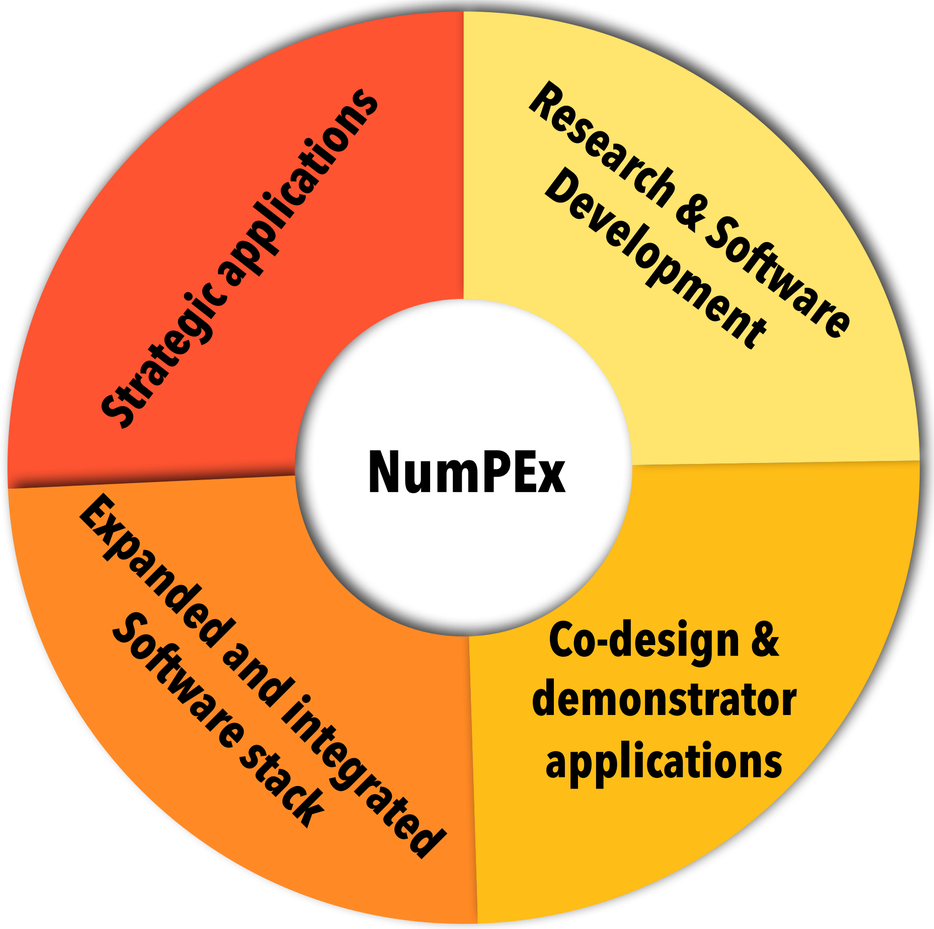
Coordinators: J.Y. Berthoud (INRIA), J. Bobin (CEA), M. Krajecki (CNRS)

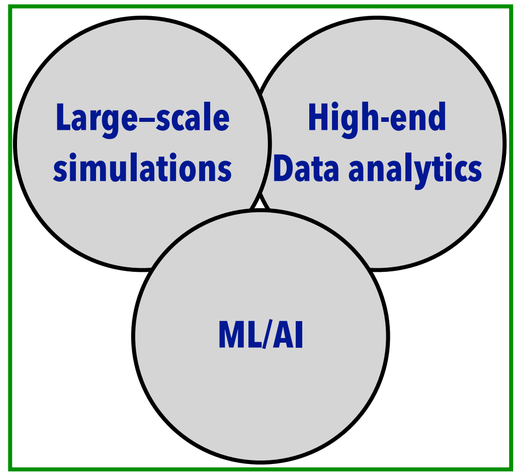
Aggregate the French HPC/HPDA/AI community, foster new collaborations and synergies

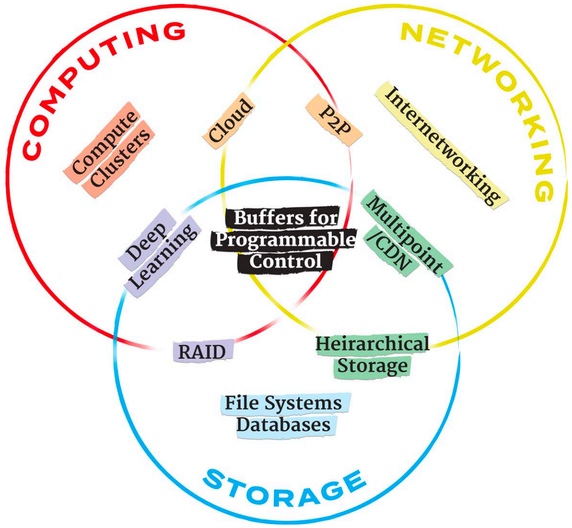
Co-develop, integrate, validate and deliver an expanded exascale software stack to accelerate productivity and sustainability of exascale applications

Contribute to and accelerate the emergence of a European sovereign exascale software stack and strategic exascale applications

Establish a multidisciplinary national workforce and develop training to improve CSE application development and software integration methodologies

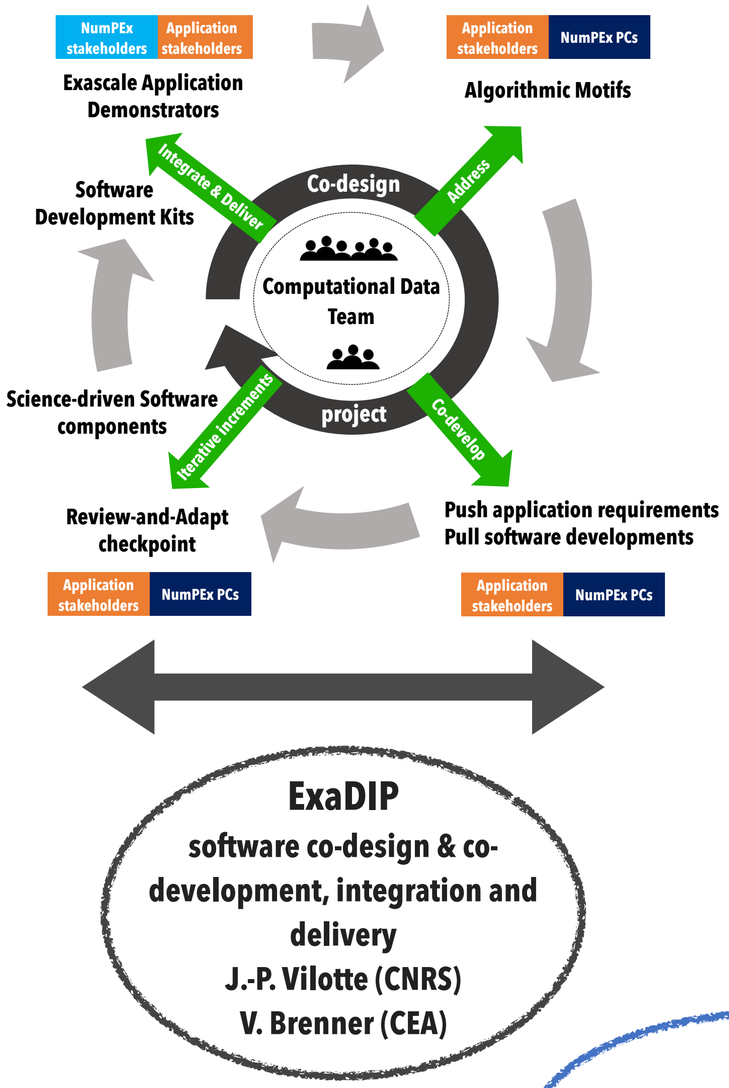




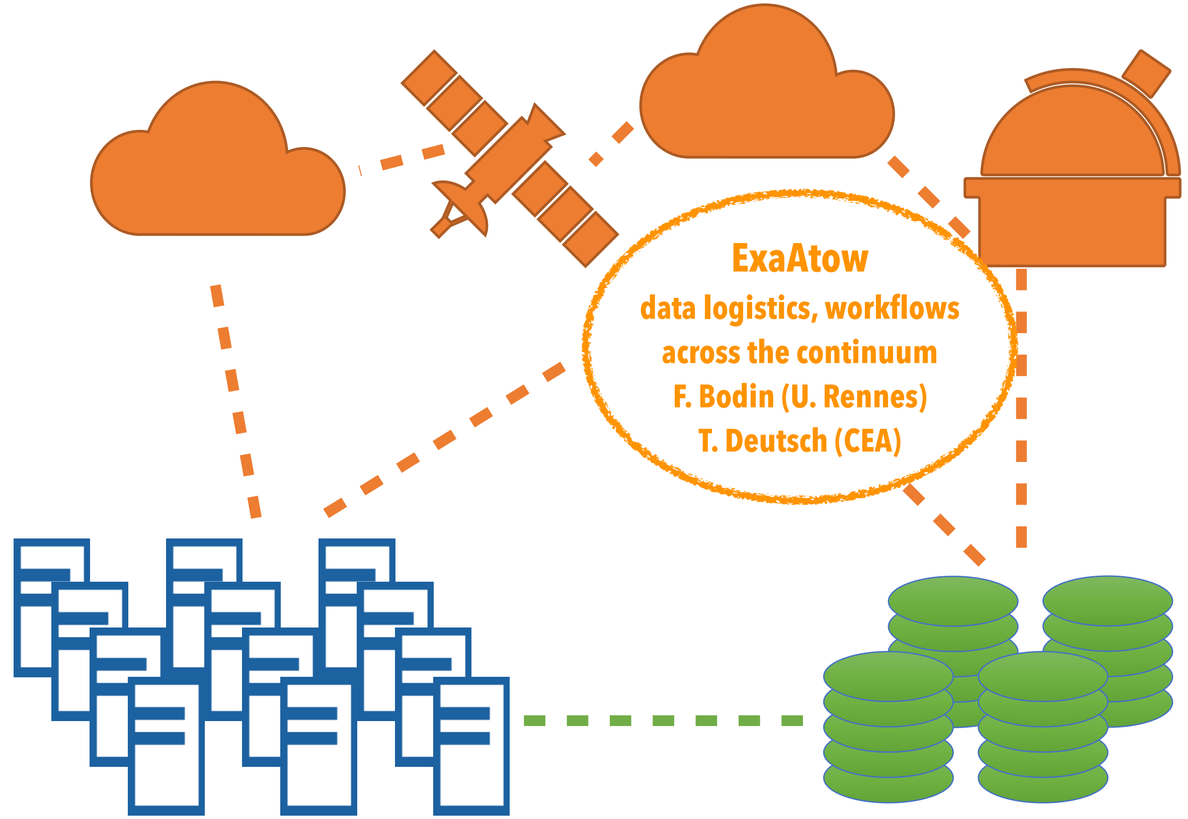


The NumPEx Projects

CSE Applications



ExaDIP
software co-design & co-development, integration and delivery
J.-P. Vilotte (CNRS)
V. Brenner (CEA)



ExaAtow
data logistics, workflows across the continuum
F. Bodin (U. Rennes)
T. Deutsch (CEA)

ExaMA
algorithms, math libraries
C. Prud'homme (UNISTRA)
H. Barucq (INRIA)

ExaSoft
parallel programming and execution environments
R. Namyst (INRIA)
A; Buttari (CNRS)

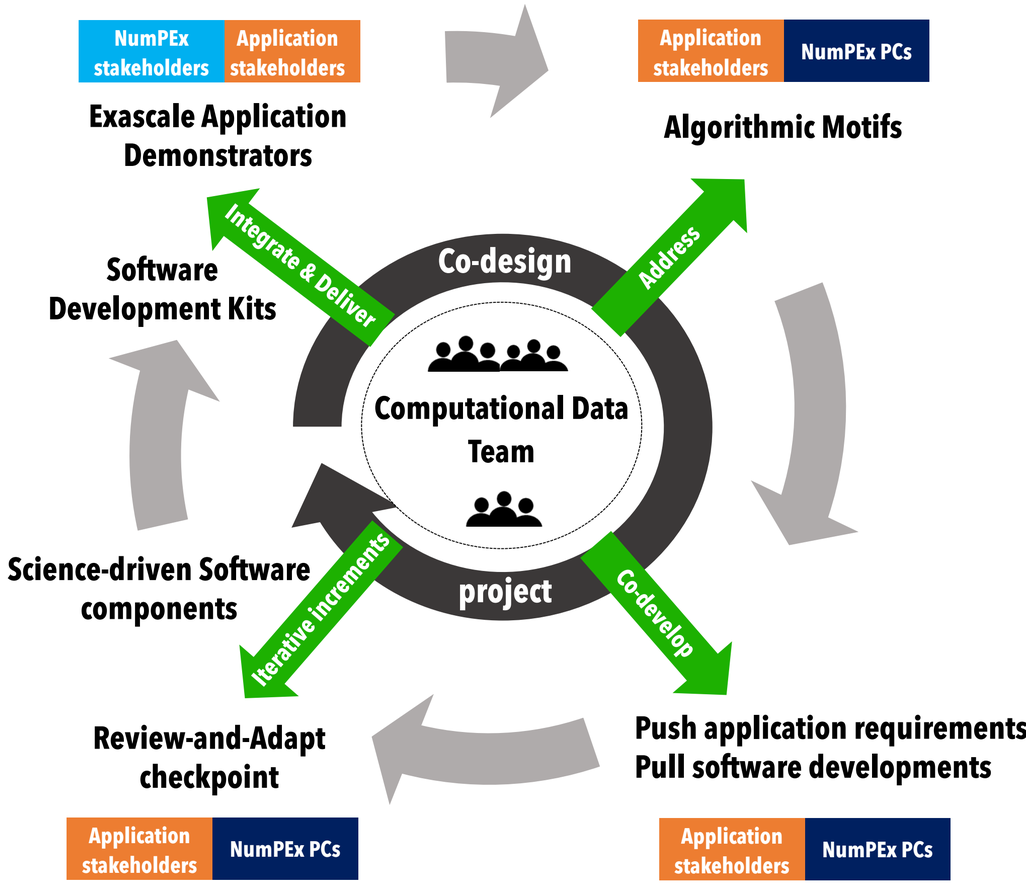
ExaDost
in-situ data reduction and analytics, storage, IO
G. Antoniu (INRIA)
J. Bigot (CEA)

The Exa-DI project: Introduction

- **Productive Exascale systems:** integrate applications and software technologies with an expanded exascale software stack easily deployable on facilities and instantiable into application environments.
- **Productivity and performance portability of exascale data-centric applications** require improvements in scientific software development methodologies leveraging high-quality software components (libraries, frameworks, workflow tools) while exploring new mathematical approaches, model improvements and execution environments.
- **Exa-DI objectives:**
  - Implement a software co-development process across NumPEx to deliver logical collections of integrated software components addressing cross-cutting computing and communication issues among a set of Application Demonstrators (ADs) covering a wide range of strategic domains
  - Promote the use of software packaging systems (e.g., Spack, Guix-HPC) across NumPeX, together with software community guidelines, functional and quality certification (badges) and CI methodologies

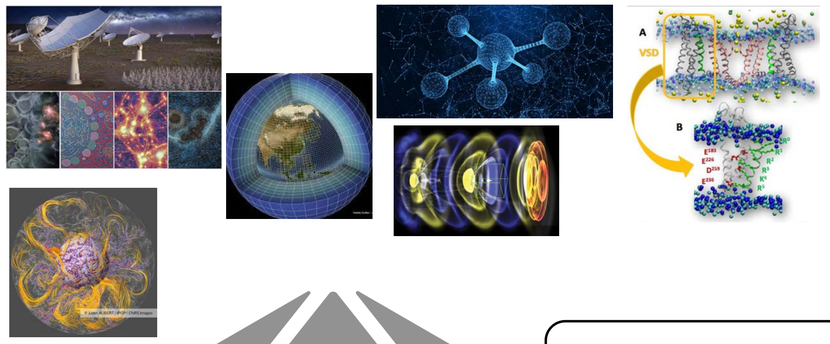
The Exa-DI iterative agile co-development process

- Co-identify cross-cutting algorithmic and communication Motifs among the ADs
- Co-develop well-documented Motifs-based proxy/mini apps for performance and portability evaluation
- Stream-aligned co-development of Motif-based software components (libraries, frameworks, tools) across NumPEx
- Integrate logical collections of Motif-based software components (SDKs) with efficient cross-layer interoperability to enable exascale application codes development
- Foster adoption of Software Community Guidelines and use of software transactional software management technologies in synergy with the national computing facilities
- Develop training materials and beacon of good practices directed to the ADs and the CSE community at large



Exa-DI strategy: important aggregation layer steering coordination between initially loosely coupled software R&D within NumPeX, while fostering longer term research.

Exa-DI CSE Application Demonstrators

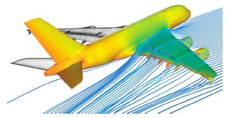


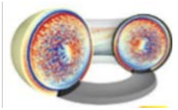
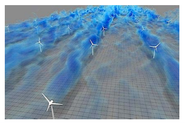
High-impact science and engineering exascale challenge problem

Criteria for assessing successful completion of the challenge problem

A Figure of merit (FOM) quantifying performance enhancement and rate of science

Demonstration and assessment of effective Software Development Kits (SDKs) integration

Astronomy & Astrophysics

Earth System Models & environment

Environmental extreme events

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 (ITER)

Sustainable Transport & mobility

Energy production & transport

…

Exascale Challenges

- Heterogeneous exascale architectures
- New multi physics and multi-scale capability
- On-line streaming data analysis /reduction
- Efficient I/Os
- AI-enabled big data analytics
- New mathematical approaches, algorithm and model improvements
- Leveraging robust and accurate logical collection of interoperable software components (libraries, frameworks, and tools)
- Improving performance portability by exercising new performance portable programming models
- New scalable programming and execution models
- Fondations for a sustainable exascale scientific software stack

Workshop objectives

Workshop Topics: Two overlapping classes of big data and ML-based problems

- Large Images analysis @ exascale; Data analysis (simulation, experiments, observation) & robust inference @ exascale

Co-identify among Applications:

- Cross-cutting computing, communication and data flow Motifs, and escale challenges
- Motifs-based proxy/mini apps enabling scalability, performance and portability evaluation,
- Software components (libraries, frameworks, workflow tools, abstraction layers, programming and execution environments) to be co-developed and integrated for evaluation in the proxy/mini apps to accelerate the exascale AI-based Application codes development and improve performance portability

Develop and share a common understanding of Exa-DI co-development process to :

- Reduce the development risk of AI-based software for exascale applications teams investigating crucial performance trade-offs related to te implementation and application of AI-based methods in science and engineering
- Produce high-performance implementation of AI-based methods and data flows through logical collection of interoperable software components that can be integrated, build and deployed through transactional software packaging technologies (e.g. Spack, Guix-hpc)
- Build cross-functional collaboration between software components development and integration with the overall NumPEx technologies, streamline developper and user workflows, CI and performance analysis methodologies
- Focus on verification, validation, reproducibility, explainability and uncertainty quantification with a solid determination of generalisation errors

Co-define agile co-development teams

➡ identified contributions of different R&D teams in-and-across the NumPEx PCs, and of the ADs,

➡ necessary resources and expertises to be gathered in the CDT to drive the agile co-development process

Set-up and organise follow-up proxy/mini apps working groups

Motif-based mini/proxy apps

Co-develop community-driven Motif-based proxy/mini apps with a high-level abstraction and programmable interface, algorithmic specifications and problem parametrisation that:

- **Capture cross-cutting complex operations or entire algorithms in workload phases** of interest *among* Applications
- **Address increasingly complex workloads, including big data workloads,** e.g., convolution, high-dimensional array transformation and data flows
- **Handle algorithm/model and data flow alternatives** for the same problem;
- **Focus on consistency, verification, validation, explainability, reproducibility and uncertainty quantification** with a solid determination of generalisation errors
- **Enable integration and evaluation of logical collections of Motifs-based software components** (libraries, frameworks, workflow tools) with more efficient cross-layer optimisation, and of performance evaluation methodologies
- **Harden the use of abstraction layers, accelerator programming and parallel execution models** to improve performance portability
- **Leverage transactional software management and deployment technologies**, together with **co-design tools such as profilers and analysing tools, simulators, continuous integration and efficiency reporting**

Connection with other projects

- PEPR
  - IA, TRACCS, ORIGINS, Diadem, Digital Health, Cloud
- EuroHPC
  - Minerva project, TPC
- National Research Infrastructures
  - SKA-France, LHC/WLCG, Data Terra / Gaia Data, CLIMERI
- European and International initiatives
  - InPEx, Simons Foundation, High-Performance Software Foundation (HPSF)