RÉPUBLIQUE FRANÇAISE
Liberté
Égalité
Fraternité

Exa-DoST: Exascale Data-oriented Software and Tools

PROGRAMME DE RECHERCHE
NUMÉRIQUE POUR L'EXASCALE

PEPR NumPEx project ANR-23-PECL-0007
1 Jan 2023 – 31 Oct. 2029 (82 months)
ANR contribution: 6 125 000 €
PI: Gabriel Antoniu (Inria), Co-PI: Julien Bigot (CEA)

The Exa-DoST "core" team

CEA/DAM - DPTA, SISR, SANL
CEA/DRF - MdlS, IRFM, IRFU
CNRS/INSU, Observatoire de Paris, Observatoire de la Côte d'Azur
Inria - DataMove, KerData, MIND, TADaaM, SODA, STATIFY
DDN

Agenda

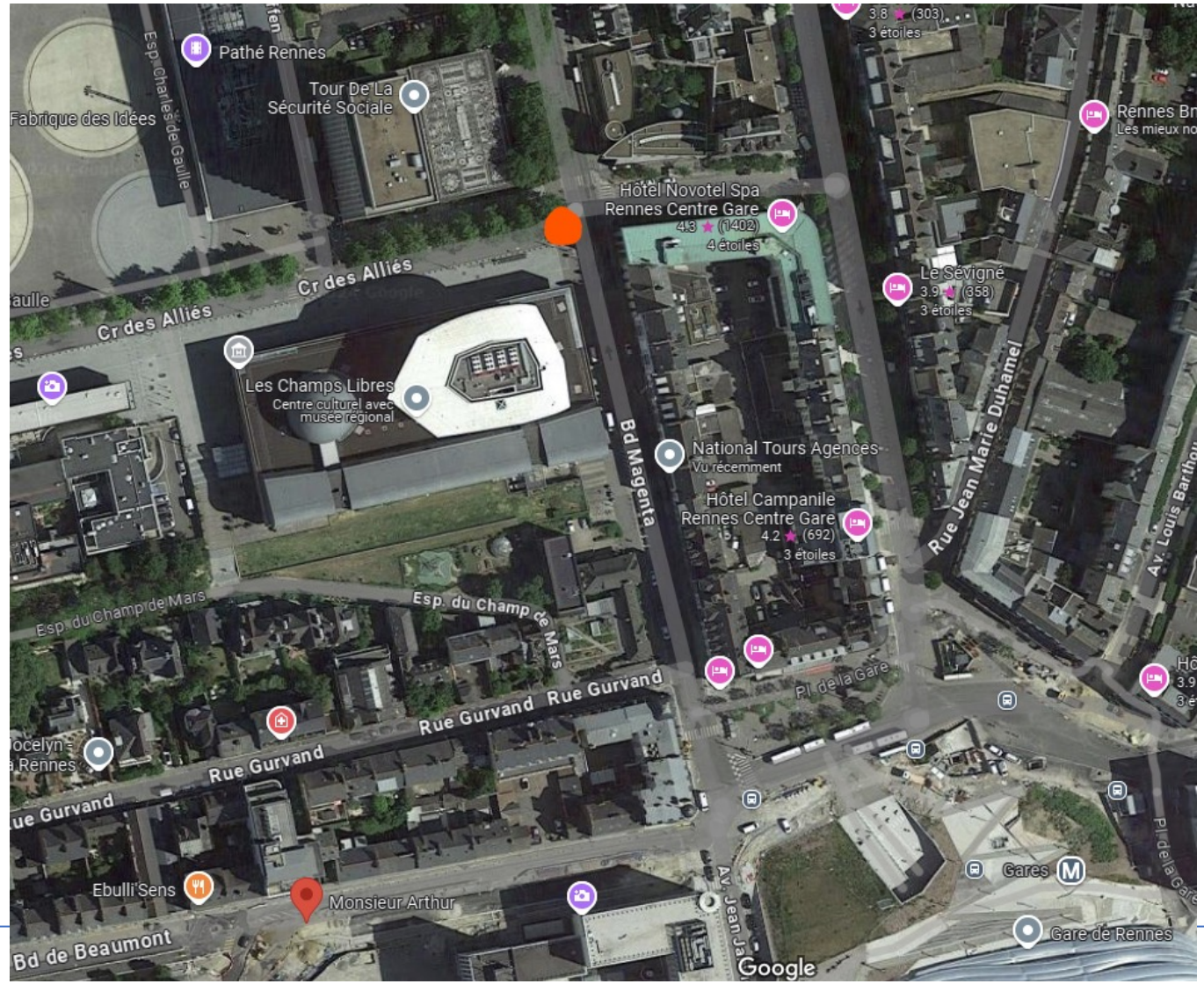
Day 1

Wednesday 18th September 2024: plenary sessions + meeting of the Board

Location	Start time	End time	Object	Speaker(s)	Minute writer	Useful info
Petri-Turing room	10:30	11:30	Welcome, coffee			
	11:30	12:00	Introduction, NumPEX and Exa-DoST	Gabriel Antoniu & Julien Bigot		Overall presentation of the project and its context (NumPEx program)
	12:00	12:30	Presentation of WP1 + discussions	Francieli Boito & François Tessier	Francieli & François & Jakob	**Content:** First results, ongoing research, challenges...
	12:30	14:00	Lunch (standing buffet within the space around the meeting room)			
	14:00	14:30	Presentation of WP2 + discussions	Yushan Wang & Laurent Colombet	Benoît	
	14:30	15:00	Presentation of WP3 + discussions	Thomas Moreau & Bruno Raffin		**Timing** : For each item, around 15 minutes of presentation and 15 minutes of discussions. The total must not be over 30 minutes.
	15:00	15:30	Presentation of WP4 + discussions	Virginie Grandgirard & Damien Gratadour	Shan Mignot & Dorian Midou	
	15:30	16:00	Break (in the meeting room)			
	16:00	16:45	How to build interactions between WPs?	Gabriel Antoniu & Julien Bigot		Open discussion
Corsica room	17:00	18:00	Meeting of the Board of Exa-DoST (in private)			Private exchange between the Board and the project leaders.
Monsieur Arthur 24 rue Raoul Dautry Rennes	19:30	22:15	Dinner	N/A		
In front of CPAM offices, esplanade Charles de Gaulle	**22:30**		**Shuttle service from Rennes to La Reposée**			

Agenda

Day 1 – Meeting place for the shuttle after dinner



Agenda

Day 2

Thursday 19th September 2024: parallel sessions

Location	Start	End time	Object	People in charge		Useful info
La Reposée	08:00		Shuttle service from La Reposée to Inria			
Amphi G	09:00	09:30	Welcome coffee (within the space around the lecture hall)			
	09:30	09:45	Introduction to the parallel sessions	Julien Bigot		
Rooms: - Direction - Corsica - Belle-Ile	09:45	10:45	Preparation of the next deliverable, on application motifs (in parallel, by application)	Damien Gratadour for SKA Virginie Grandgirard for Gysela Laurent Colombet for other apps	Francieli & François & Jakob for WP1 - Dorian Midou for WP4-GYSELA	**3 sessions in parallel**: Session 1 on motifs originating from SKA-oriented workflows Session 2 on motifs originating from Gysela-oriented workflows Session 3 on motifs originating from other applications (Coddex, ...) Read guidelines here: https://docs.google.com/document/d/19nSBNEbY4PMmrSAIqWkSHUwZbp4mJovqTZk9LispIoQ/edit?usp=sharing
Near Amphi G	10:45	11:15	Coffee break (within the space around the lecture hall)			
Rooms: - Direction - Corsica - Belle-Ile	11:15	12:15	Preparation of the next deliverable, on application motifs (in parallell, by WP)	Francieli Boito & François Tessier for WP1 Yushan Wang & Laurent Colombet for WP2 Thomas Moreau & Bruno Raffin for WP3	Francieli & François & Jakob for WP1 Benoît for WP2	**3 sessions in parallel**: One session for the identifcation of motifs handled by each WP (from 1 to 3) Read guidelines here: https://docs.google.com/document/d/19nSBNEbY4PMmrSAIqWkSHUwZbp4mJovqTZk9LispIoQ/edit?usp=sharing
Near Amphi G	12:15	13:45	Lunch (standing buffet within the space around the lecture hall)			
Amphi G	13:45	14:15	Workshop feeback: WP1 (storage and I/O)	Francieli Boito & François Tessier	Francieli & François & Jakob	**Timing** : For each workshop: 10-15 min of feedback + 15-20 min of discussions. The total must not be longer than 30 minutes.
	14:15	14:45	Workshop feeback: WP2 (in situ processing)	Yushan Wang & Laurent Colombet	Benoît	
	14:45	15:15	Workshop feeback: WP3 (ML-based analysis)	Thomas Moreau & Bruno Raffin		
	15:15	15:30	Conclusion	Gabriel Antoniu & Julien Bigot		
Room: Direction	15:30	16:30	Gysela/Damaris exchange (for those it concerns)	Gabriel Antoniu + Virginie Grandgirard		

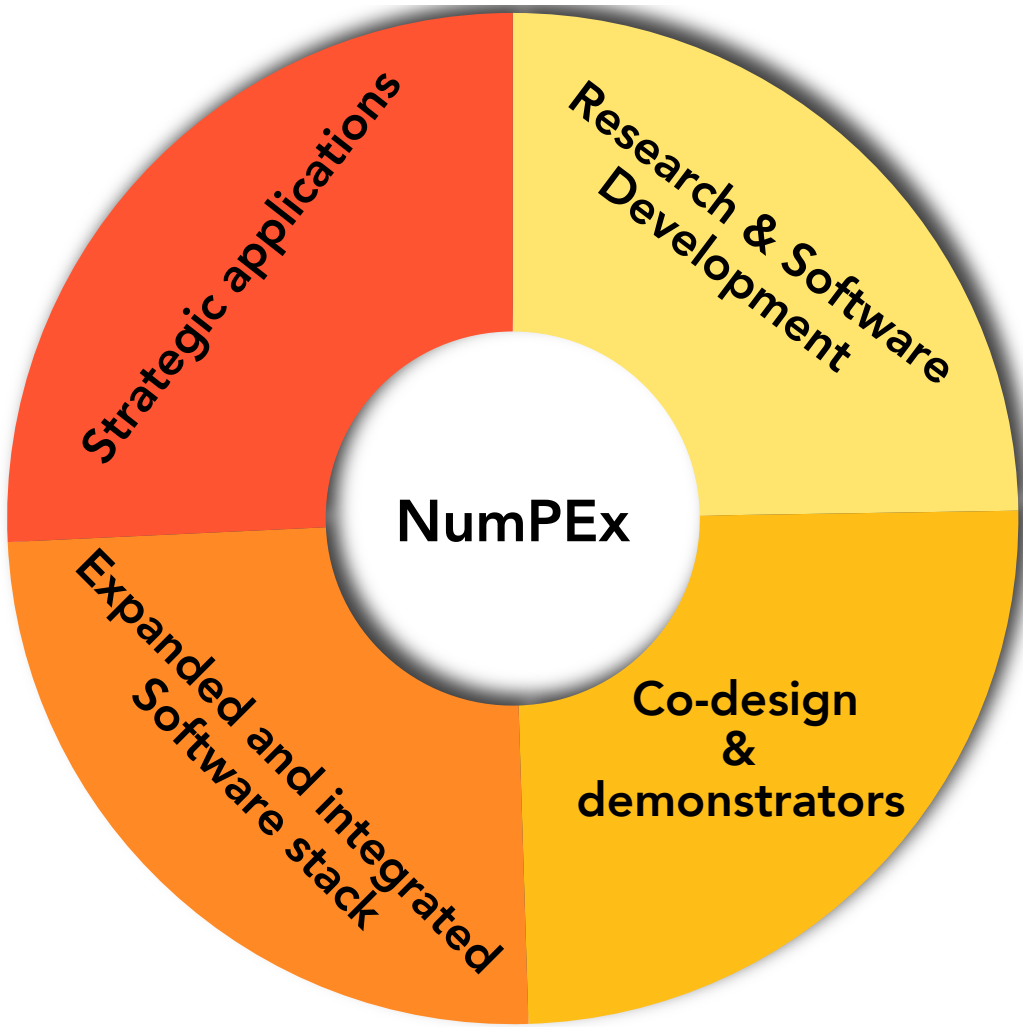
The French NumPEx Program

Consolidating and accelerating the construction of a 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**

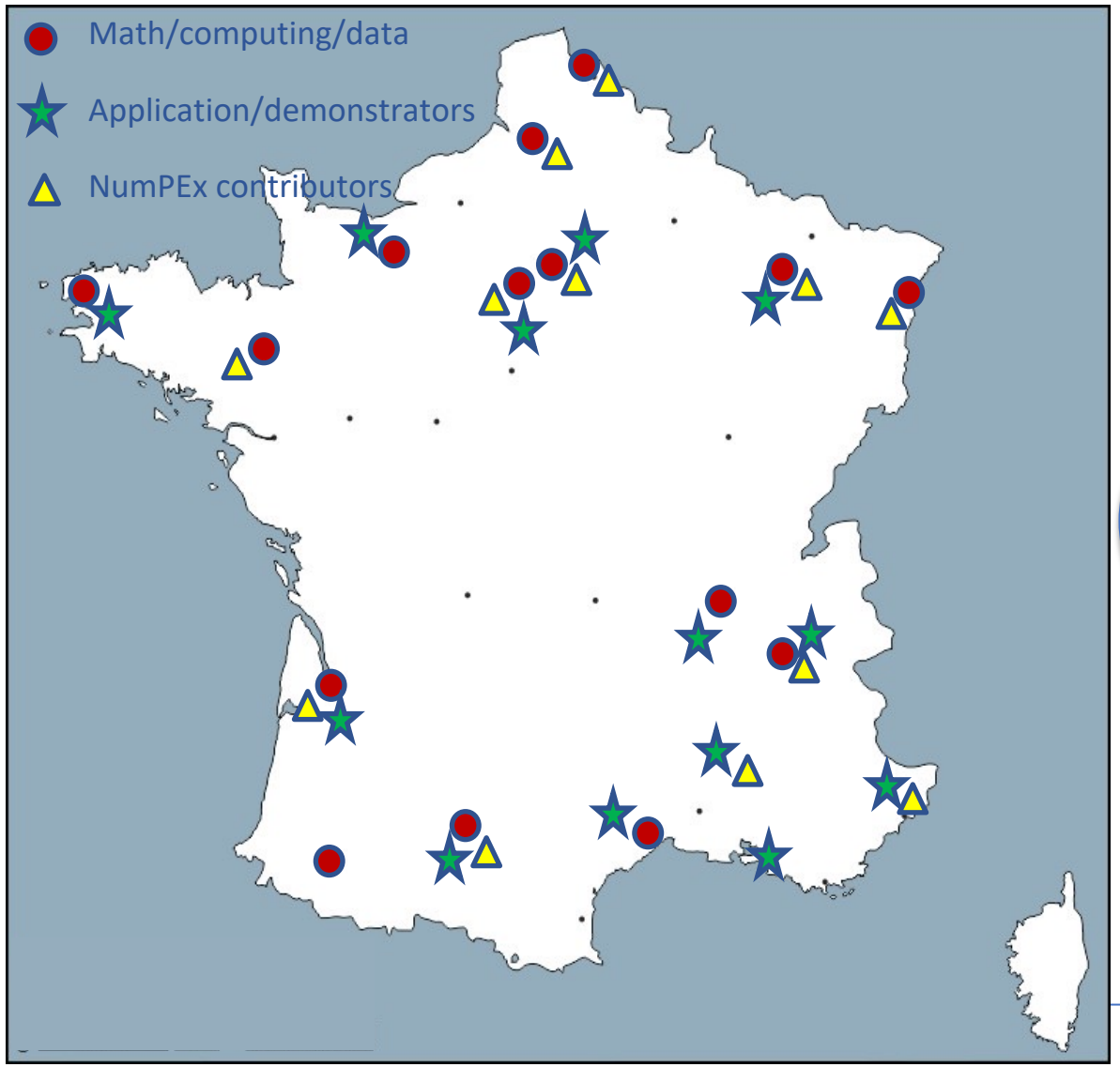
Foster national and international collaborations to prepare for the Exascale and post-Exascale era



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

https://numpex.org

NumPEx by Numbers



6 Years
41 M€*

2023-2029
* 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 (Projects 1-3)
Wide-area workflows and architecture (Project 4)
Integration and application development (Project 5)

80 R&D teams
500 Researchers

The French NumPEx Program: Workplan

Applications

ExaDI

Application co-design and software integration

JP.Vilotte/V.Brenner

ExaAtow

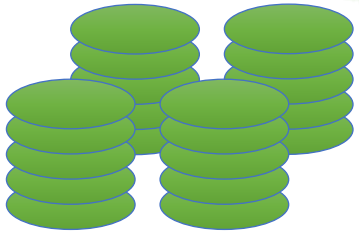
Digital continuum

F.Bodin, T.Deutsch, M.Asch

ExaDost

Data

G.Antoniu/J.Bigot



ExaMA

Numerical methods and solvers

C.Prudhomme/H.Barucq

ExaSoft

Computing

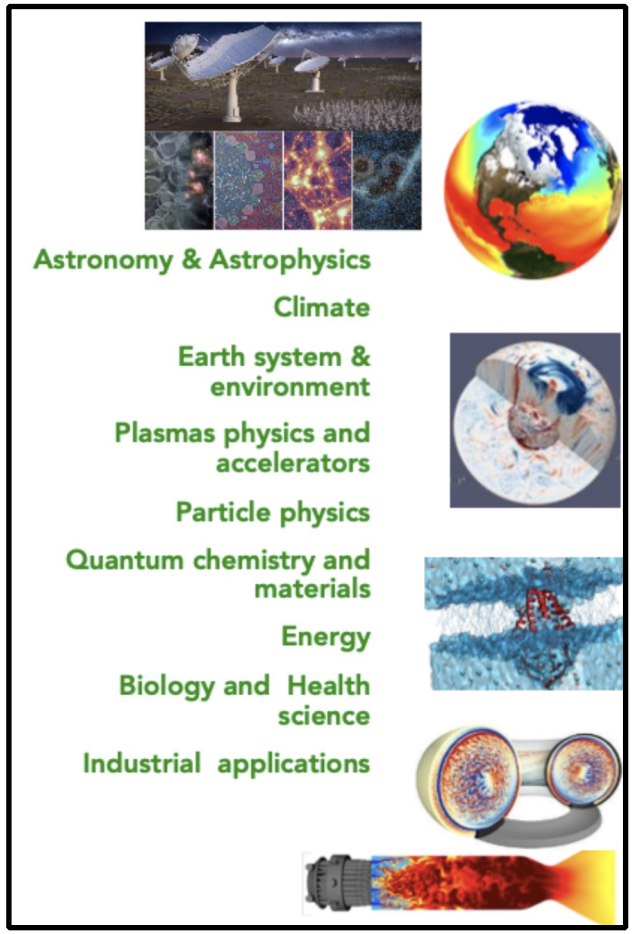
R.Namyst/A.Butari

The French NumPEx Program: Objectives

European Pre-Exascale system

Software stack

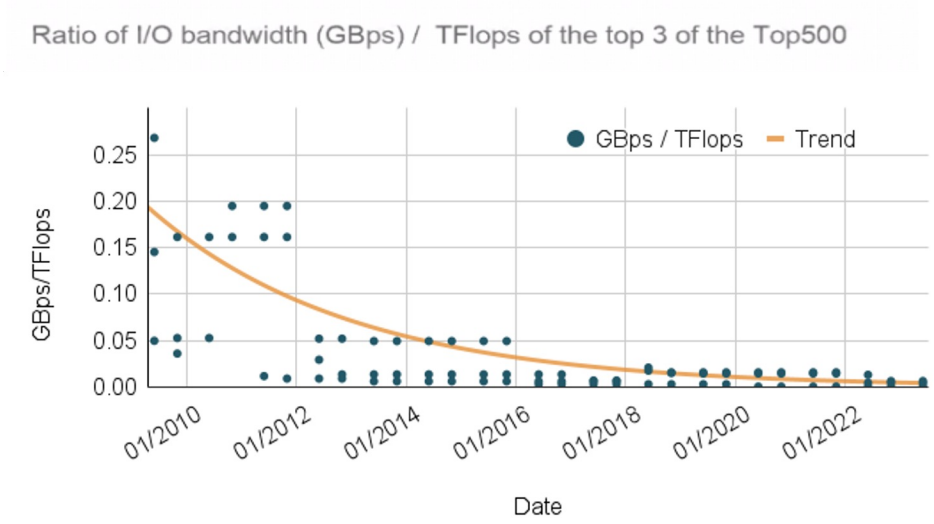
Applications



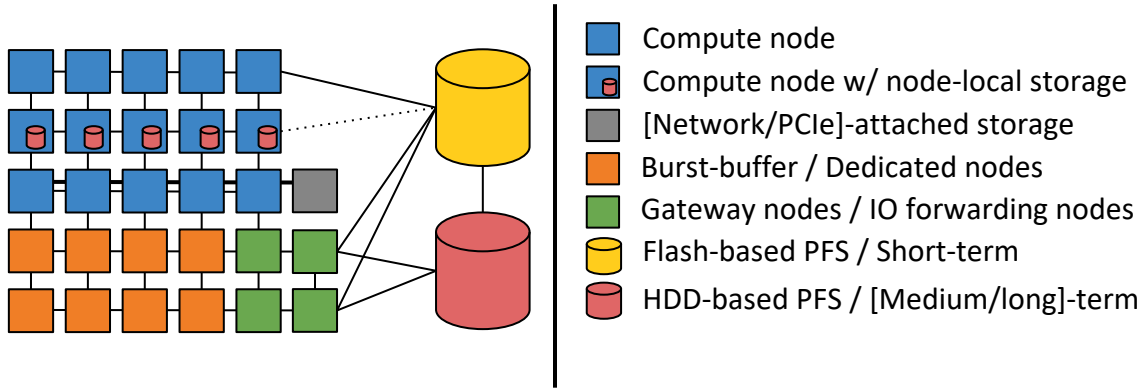
Co-design the exascale software stack
Preparing the applications for the Exascale era

Data at exascale: a challenge in hardware

- Increasing **gap between compute and I/O** performance on large-scale systems
  - Ratio of I/O to computing power divided by ~10 over the last 10 years on the top 3 supercomputers
- … and data deluge!
  - At NERSC, **data volume x41** in 10 years



- New storage tiers and advanced architectures to try to mitigate this increasing bottleneck
  - More complex on-node memory layout
  - Emerging complex applications and workflows have to adapt



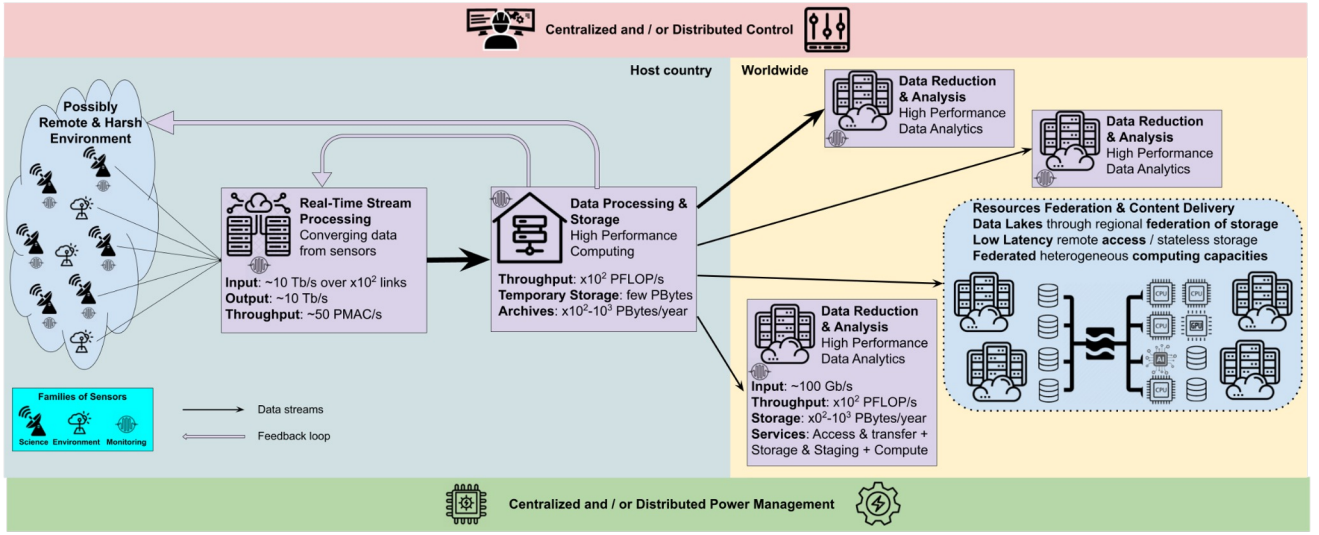
Trend in storage technologies available on extreme-scale systems

Data at exascale: a challenge in usages

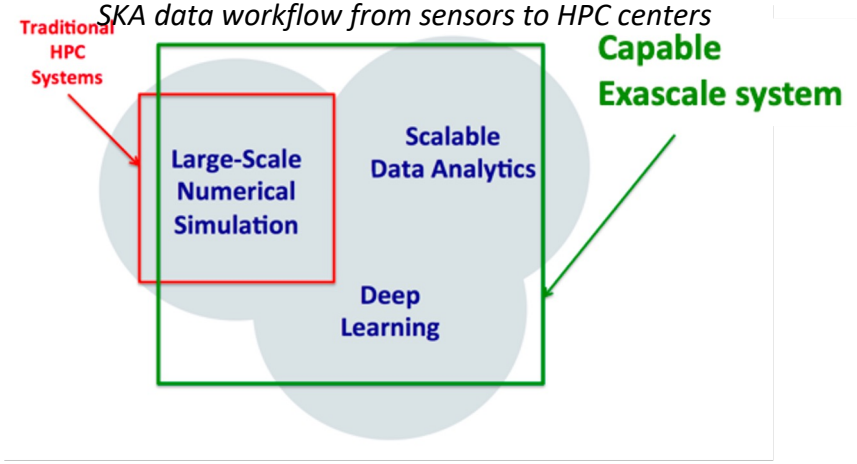
- HPC centers do not live in isolation anymore
  - Edge - Cloud - HPC continuum

- Emerging workloads are hybrid
  - High-performance simulation
  - High-performance data analytics
  - Machine learning and artificial intelligence

- Interaction with data from the outside world sensors
  - Large scientific instruments
  - …



SKA data workflow from sensors to HPC centers



Our ambition

Fill the gaps in the existing software stack designed by previous projects (e.g. ECP)

Take into account French & European specificities

Ensure French & European needs are taken into account in roadmaps

Approach:

- **Research** on data-oriented tools for HPC
- Transverse, **re-usable tools**
- Usable **in production** at exascale

⇒ ExaDoST will produce:

- **New approaches** to handle the data challenge at exascale
- Transverse **libraries & tools** that implement these approaches

Validated in illustrators at full scale

Fully application agnostic

Fully open-source

Work Packages in Exa-DoST

WP1: Exascale I/O and storage

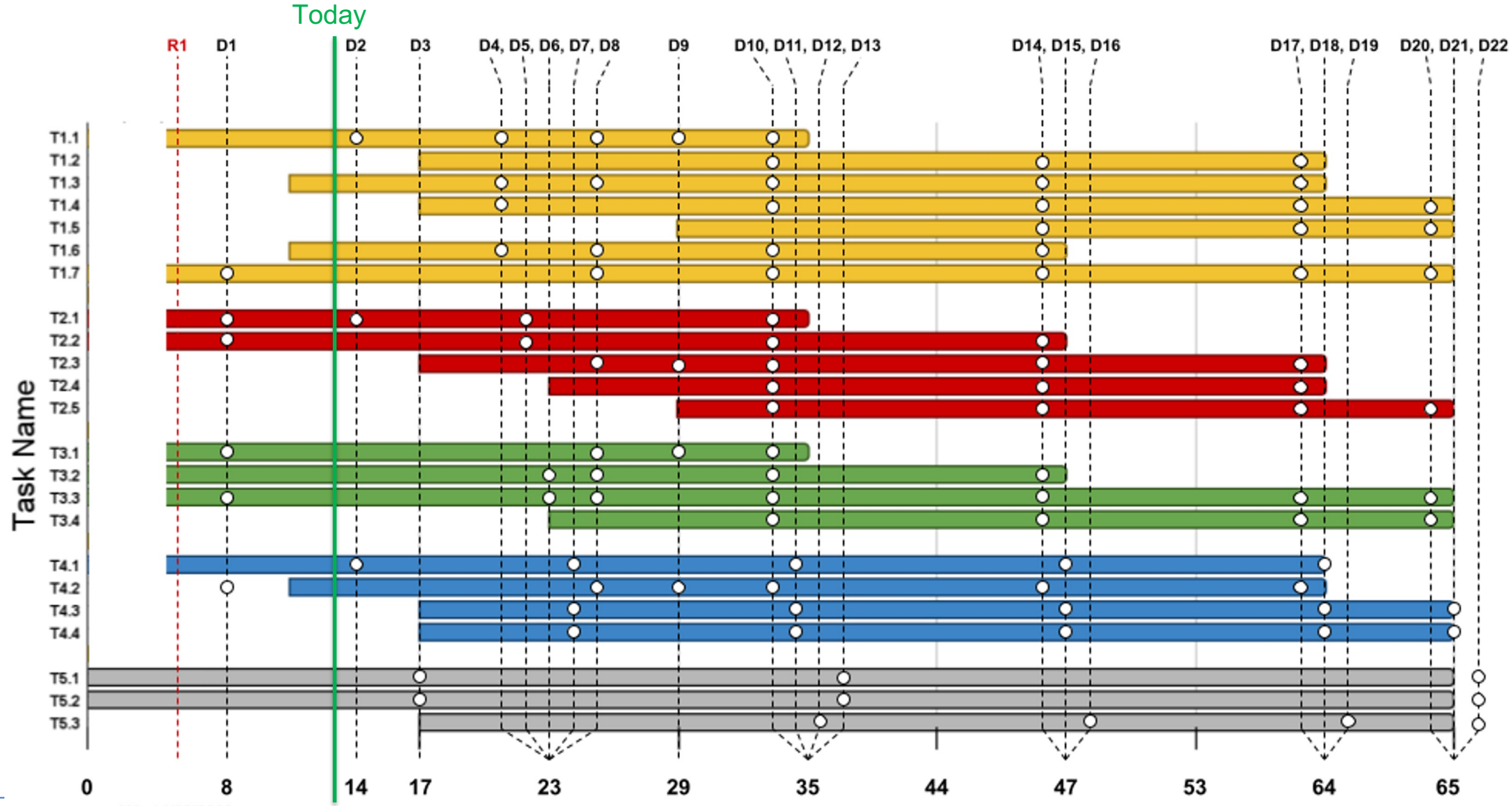
WP2: Exascale In-situ data processing

WP3: Exascale ML-based data analytics

WP4: Shared building blocks & integrated illustrators

WP5: Management, dissemination and training

Updated Gantt Chart



M0: 14/08/2023

R1: Date de fin de la période couverte par le présent rapport (31/12/2023)

Potential illustrators

- and co-design demonstrators proposals for ExaDIP (PC5)

Integrated transverse illustrators

- Gysela (CEA/DRF/IRFM)
- SKA (CNRS + …)

Motif-specific motivators

- Coddex (CEA/DAM)
- CROCO (CNRS + Inria + …)
- Dyablo (CEA/IRFU)

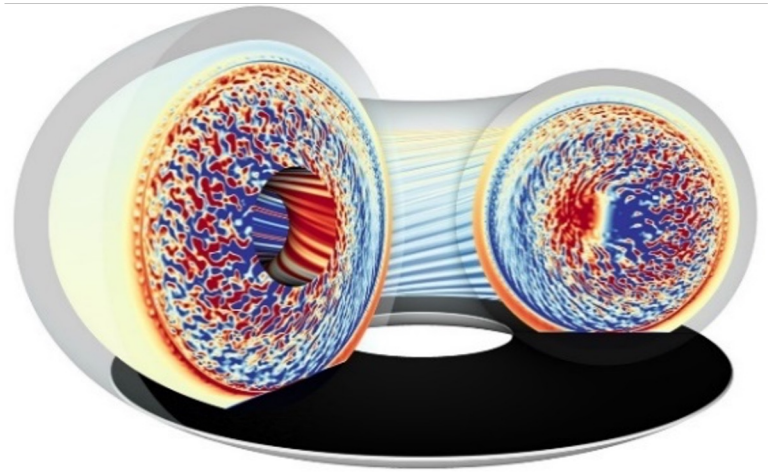
Relevant co-design demonstrators related to ExaDIP (PC5)

- Metalwalls (Sorbonne + CNRS)
- Parflow (UGA/IGE)
- RTA-France
- Many more…
- …

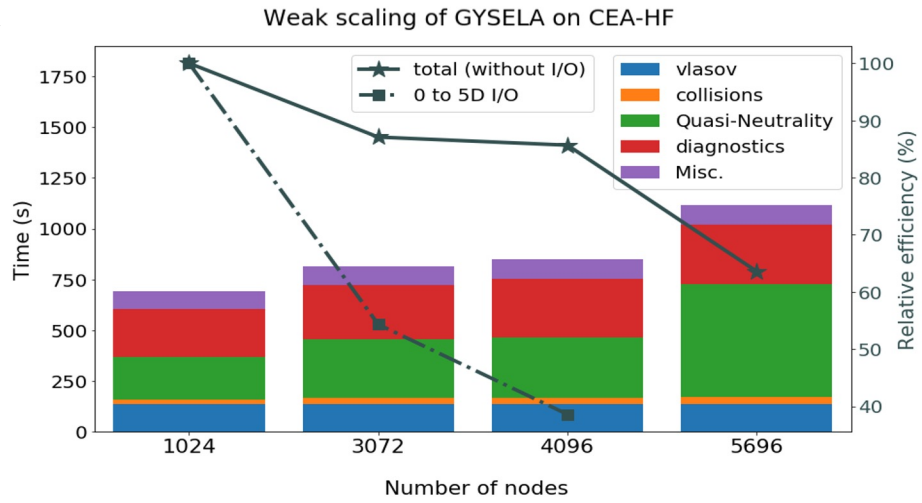
Illustrator 1: GYSELA towards exascale

Main challenge : optimized management of huge amount of data with in-Situ AI-based diagnostics

- GYSELA a non-linear 5D gyrokinetic code developed for 25 years at CEA/IRFM to simulate plasma turbulence in tokamaks.
- Optimized up to 730k CPU -> Intensive use of petascale resources (~150 millions of CPU h / year)



Typical simulation:
- **100 billion points** (5D mesh: 3D space + 2D velocity)
- **~7 million of CPU hours** (3.5 days / 65k cores)



Relative efficiency of 85% on more than 500k cores

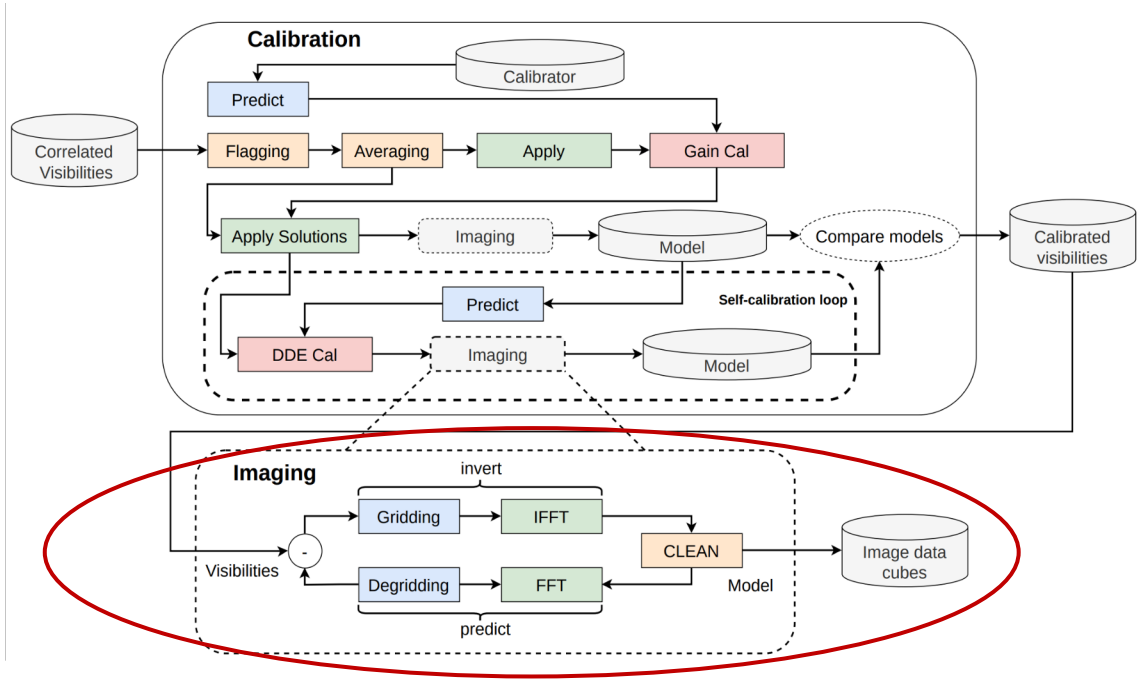
[V. Grandgirard et al., PASC 2022]

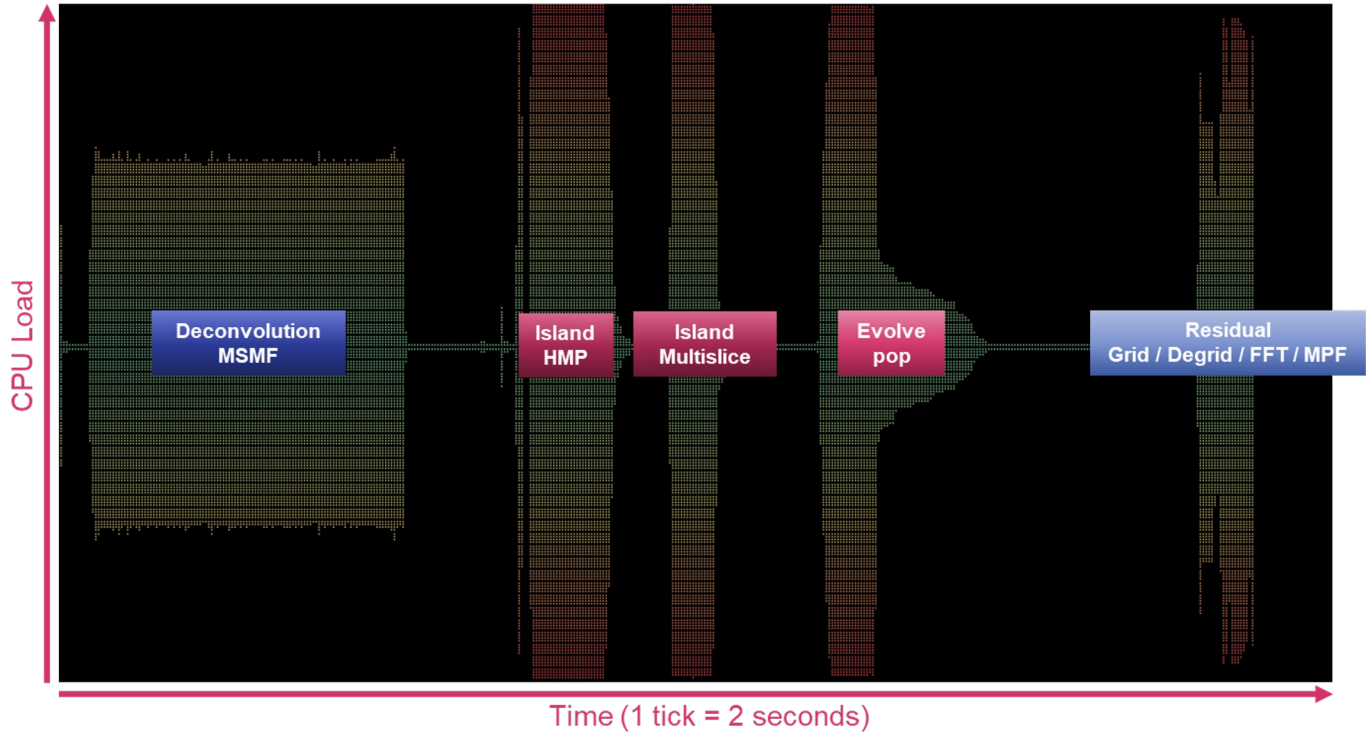
- **I/O scalability** is **an issue**: ~50% for 3072 nodes and ~38% for 4096 nodes. Crash on 5696 nodes
- **Need to be solved for exascale** ITER-like **simulations**

Illustrator 2: Square Kilometer Array (SKA)

Based on SotA software used to process large surveys (DDFacet)

- Complex iterative pipeline: optimize I/O footprint and upscale
- Current work: initial benchmarks and motifs identification





Advisory committees

Exa-DoST Scientific Advisory Committee

- Rosa Badia (BSC)
- Franck Cappello (ANL)
- Yann Meurdesoif (CEA)
- Kento Sato (Riken)
- Frederic Suter (ORNL)
- Chiara Ferrari (OCA)

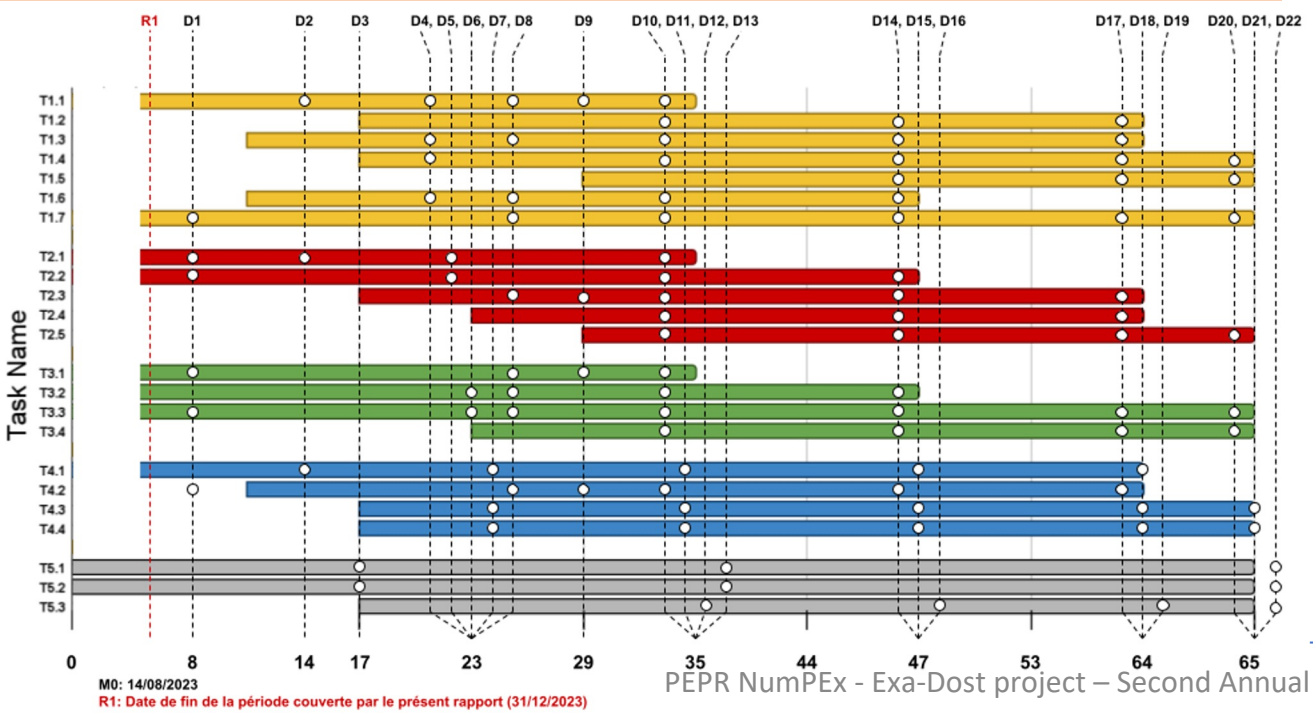
Exa-DoST Industrial and Technology Advisory Committee

- Stéphane Requena (GENCI)
- Nicolas Lardjane (TGCC)
- Pierre-François Lavallée (IDRIS)
- Gabriel Hautreux (CINES)
- Sai Narasimhamurthy (ParTec)
- François Mazen (Kitware)

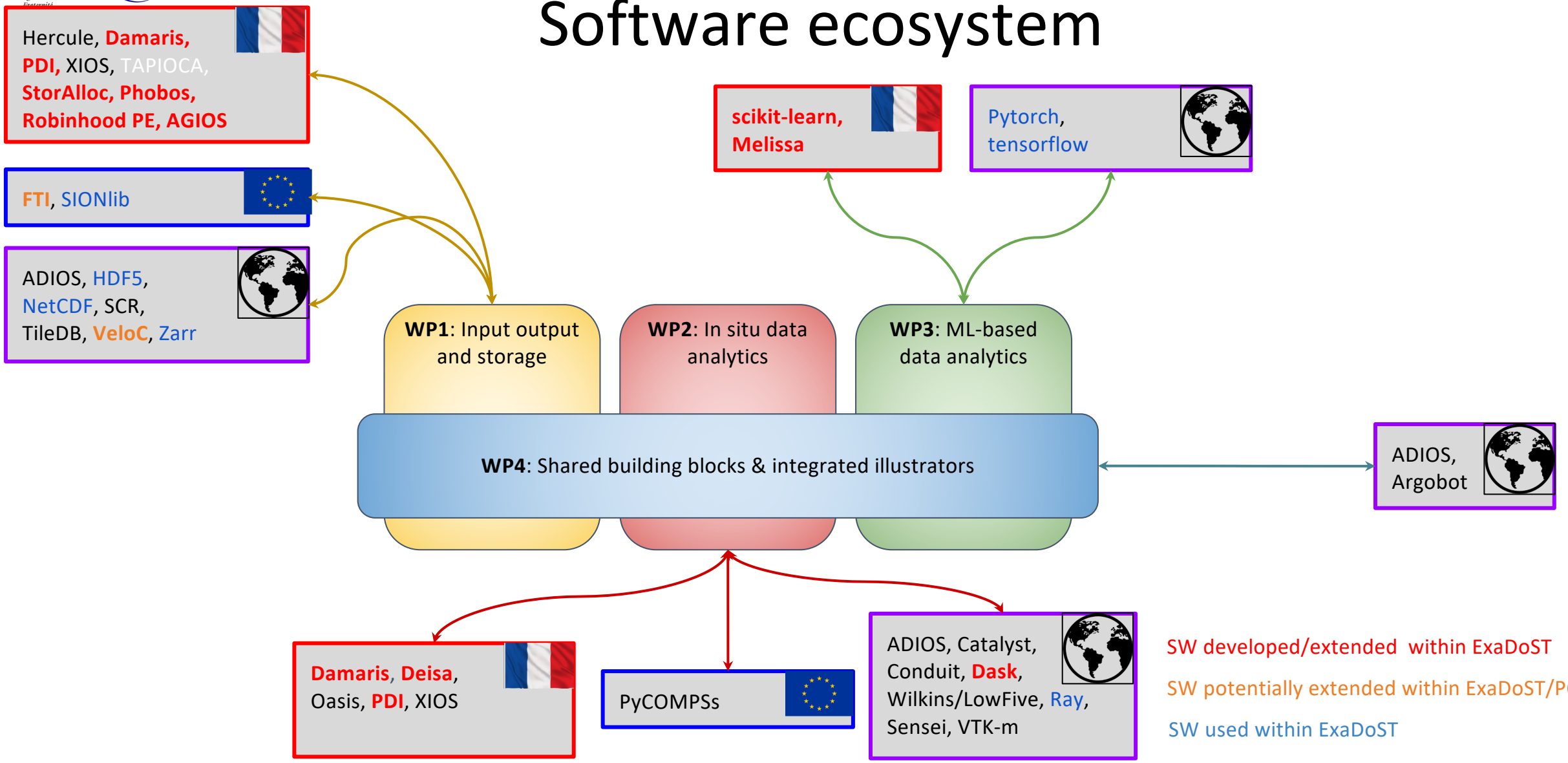
Planned deliverables

14 August 2023 : Technical T0 for NumPEx

1. [MdlS, R] (M0+08) **WP1,2,3,4**: Selection of the initial release of the libraries and tools that will make up the ExaDoST software stack.
2. [LESIA, R] (M0+14) **WP2,4**: Analysis of relevant application motifs and their covering by the project illustrators.
3. [KerData, R] (M0+17) **WP5**: Report on the project management and selected co-development strategy.
4. [TADAAM, R] (M0+23) **WP1**: Report on the solutions selected in ExaDoST to answer the storage and IO challenges at Exascale
5. [DPTA?, R] (M0+23) **WP2**: Report on the solutions selected in ExaDoST to answer the in situ challenge at Exascale
6. [MIND, R] (M0+23) **WP3**: Report on the solutions selected in ExaDoST to answer the ML-related challenge at Exascale
7. [IRFM, R] (M0+23) **WP4**: Design document for illustrators
8. [KerData, C] (M0+23) **WP1,2,3**: Intermediate coordinated release of all tools and libraries produced by ExaDoST, including documentation
9. [DataMove, R] (M0+29) **WP4**: Report on data-oriented software modularization and component mutualization between libraries
10. [MdlS, C] (M0+35) **WP1,2,3**: Intermediate coordinated release of all tools and libraries produced by ExaDoST, including documentation
11. [LESIA, C] (M0+35) **WP4**: Intermediate coordinated illustrators release, including documentation
12. [MdlS, R] (M0+35) **WP5**: Intermediate training material release for all tools and libraries produced by ExaDoST
13. [KerData, R] (M0+35) **WP5**: Mid-project report on project management and co-development strategy
14. [SANL, C] (M0+47) **WP1,2,3**: Intermediate coordinated release of all tools and libraries produced by ExaDoST, including documentation
15. [JLLL, C+R] (M0+47) **WP4**: Intermediate coordinated illustrators release, including documentation
16. [MdlS, R] (M0+47) **WP5**: Intermediate training material release for all tools and libraries produced by ExaDoST
17. [DataMove, C] (M0+59) **WP1,2,3**: Final releases of all tools and libraries produced by ExaDoST, including documentation
18. [IRFM, C+R] (M0+59) **WP4**: Final illustrators releases including a report on the integration of libraries
19. [MdlS, R] (M0+59) **WP5**: Final training material release for all tools and libraries produced by ExaDoST
20. [DataMove, R] (M0+65) **WP1,2,3**: Report on the final design of the tools and libraries produced by ExaDoST and design solved
21. [LESIA, R] (M0+65) **WP4**: Full exascale evaluation of illustrators leveraging libraries developed in the project
22. [KerData, R] (M0+65) **WP5**: Final report on project management and co-development strategy



Software ecosystem



Milestones

1. (M0+14): Identification of relevant application motifs
2. (M0+17): Initial coordinated software components released for the first cycle integration phase
3. (M0+23): Illustrators designed
4. (M0+29): Identification of mutualization potential of data-oriented software components
5. (M0+35): Intermediate coordinated software components released for the second cycle integration phase
6. (M0+35): First release of illustrators based on Gysela and SKA integrating contributed components from WP1-3
7. (M0+49): First run of illustrators based on Gysela and SKA integrating contributed components from WP1-3
8. (M0+59): Illustrators ready to run at full scale

Agenda

Day 1

Wednesday 18th September 2024: plenary sessions + meeting of the Board

Location	Start time	End time	Object	Speaker(s)	Minute writer	Useful info
Petri-Turing room	10:30	11:30	Welcome, coffee			
	11:30	12:00	Introduction, NumPEX and Exa-DoST	Gabriel Antoniu & Julien Bigot		Overall presentation of the project and its context (NumPEx program)
	12:00	12:30	Presentation of WP1 + discussions	Francieli Boito & François Tessier	Francieli & François & Jakob	**Content:** First results, ongoing research, challenges...
	12:30	14:00	Lunch (standing buffet within the space around the meeting room)			
	14:00	14:30	Presentation of WP2 + discussions	Yushan Wang & Laurent Colombet	Benoît	
	14:30	15:00	Presentation of WP3 + discussions	Thomas Moreau & Bruno Raffin		**Timing** : For each item, around 15 minutes of presentation and 15 minutes of discussions. The total must not be over 30 minutes.
	15:00	15:30	Presentation of WP4 + discussions	Virginie Grandgirard & Damien Gratadour	Shan Mignot & Dorian Midou	
	15:30	16:00	Break (in the meeting room)			
	16:00	16:45	How to build interactions between WPs?	Gabriel Antoniu & Julien Bigot		Open discussion
Corsica room	17:00	18:00	Meeting of the Board of Exa-DoST (in private)			Private exchange between the Board and the project leaders.
Monsieur Arthur 24 rue Raoul Dautry Rennes	19:30	22:15	Dinner	N/A		
In front of CPAM offices, esplanade Charles de Gaulle	**22:30**		**Shuttle service from Rennes to La Reposée**			

Agenda

Day 2

Thursday 19th September 2024: parallel sessions

Location	Start	End time	Object	People in charge		Useful info
La Reposée	08:00		Shuttle service from La Reposée to Inria			
Amphi G	09:00	09:30	Welcome coffee (within the space around the lecture hall)			
	09:30	09:45	Introduction to the parallel sessions	Julien Bigot		
Rooms: - Direction - Corsica - Belle-Ile	09:45	10:45	Preparation of the next deliverable, on application motifs (in parallel, by application)	Damien Gratadour for SKA Virginie Grandgirard for Gysela Laurent Colombet for other apps	Francieli & François & Jakob for WP1 - Dorian Midou for WP4-GYSELA	**3 sessions in parallel**: Session 1 on motifs originating from SKA-oriented workflows Session 2 on motifs originating from Gysela-oriented workflows Session 3 on motifs originating from other applications (Coddex, ...) Read guidelines here: https://docs.google.com/document/d/19nSBNEbY4PMmrSAIqWkSHUwZbp4mJovqTZk9LispIoQ/edit?usp=sharing
Near Amphi G	10:45	11:15	Coffee break (within the space around the lecture hall)			
Rooms: - Direction - Corsica - Belle-Ile	11:15	12:15	Preparation of the next deliverable, on application motifs (in parallell, by WP)	Francieli Boito & François Tessier for WP1 Yushan Wang & Laurent Colombet for WP2 Thomas Moreau & Bruno Raffin for WP3	Francieli & François & Jakob for WP1 Benoît for WP2	**3 sessions in parallel**: One session for the identifcation of motifs handled by each WP (from 1 to 3) Read guidelines here: https://docs.google.com/document/d/19nSBNEbY4PMmrSAIqWkSHUwZbp4mJovqTZk9LispIoQ/edit?usp=sharing
Near Amphi G	12:15	13:45	Lunch (standing buffet within the space around the lecture hall)			
Amphi G	13:45	14:15	Workshop feeback: WP1 (storage and I/O)	Francieli Boito & François Tessier	Francieli & François & Jakob	**Timing** : For each workshop: 10-15 min of feedback + 15-20 min of discussions. The total must not be longer than 30 minutes.
	14:15	14:45	Workshop feeback: WP2 (in situ processing)	Yushan Wang & Laurent Colombet	Benoît	
	14:45	15:15	Workshop feeback: WP3 (ML-based analysis)	Thomas Moreau & Bruno Raffin		
	15:15	15:30	Conclusion	Gabriel Antoniu & Julien Bigot		
Room: Direction	15:30	16:30	Gysela/Damaris exchange (for those it concerns)	Gabriel Antoniu + Virginie Grandgirard		

RÉPUBLIQUE FRANÇAISE
Liberté
Égalité
Fraternité

cea cnrs Inria

FRANCE 2030

PROGRAMME DE RECHERCHE

NUMÉRIQUE POUR L'EXASCALE

Retrouvez toutes nos actualités

NumPEx