1ère journée prospectives IP2I 2022 - retour sur les prospectives IN2P3
30 Juin 2022

GT09

Calculs, algorithmes et données

Laurent Vacavant
D. Zerwas(IRN Terascale.LAL), GF. Polci(GDR Inf,LPNHE), C.Smith (CSI,LPSC), M.G. Genest (LPSC), C. Ochando (LLR)

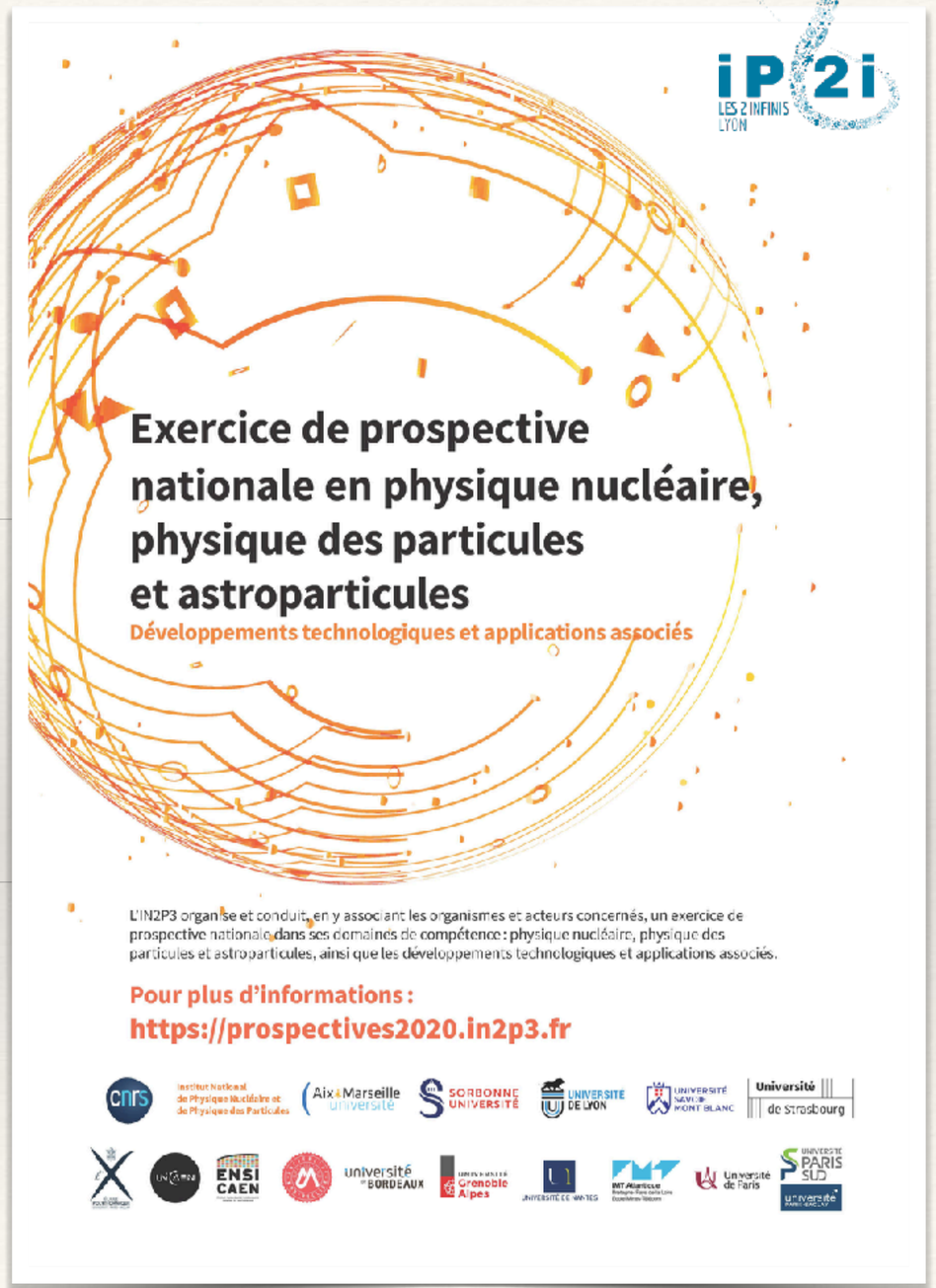
Présentation au colloque de restitution

Discussions
Questions

Rapport du Groupe de Travail 09

\+ Town Hall Meeting

+ CS IN2P3 Calculs et données 23/06/2022 : https://indico.in2p3.fr/event/27438/



Contexte … très simplifié !!

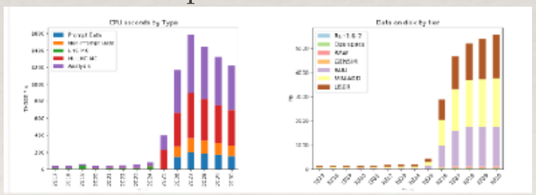
Status et besoins

A budget constant, le modèle Grille actuel ne tient pas
(Facteur 10 en computing, facteur 6 storage*)
☞ Nouveau modèle à inventer

High quality standards for software & workflow

I.A. / Machine Learning

* Prospectives IP2I 2019



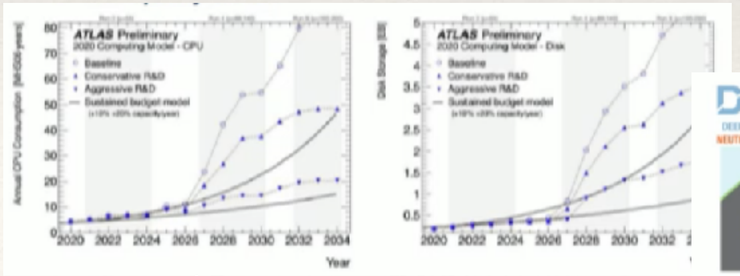
HL-LHC

Astro

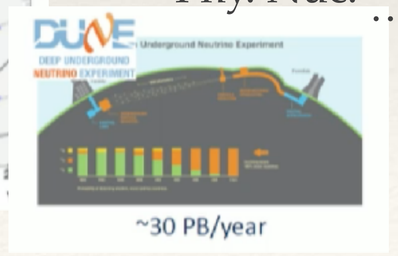
Cosmo

Phy. Nuc.

….



* Prospectives IN2P3 2021



Calcul

Algorithmes

DONNEES

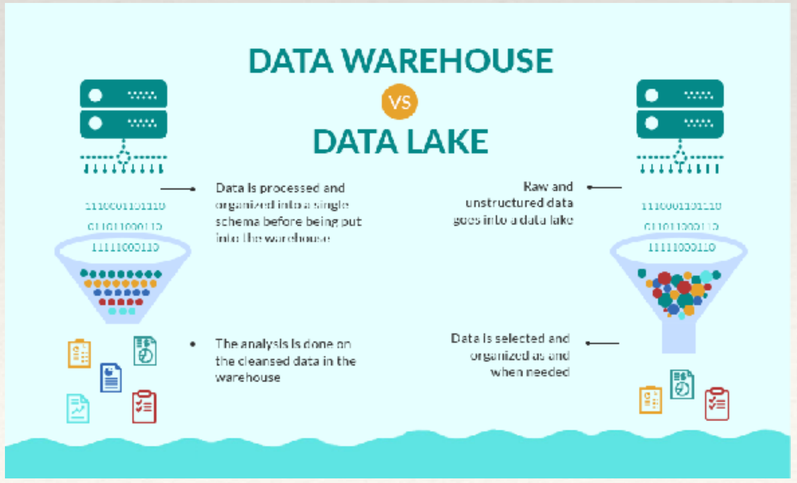
« Solutions »

Les grandes expériences ne sont plus moteurs d'innovation
ce sont les géants du web !

European Open Science Cloud (EOSC)

Accelerators such as GPU / FPGA

Data lake



Drivers et Recommandations du GT09

Drivers

- Develop a reliable, evolving, and performing e-infrastructure suitable to IN2P3 experiment needs
- Make use and develop breakthrough solutions to tackle the upcoming HEP, astroparticle, and cosmology data flow

Rapport structuré par :

Infrastructure needs & computing models
Accelerators & Emerging Technologies
Artificial Intelligence
Software development and quality assurance
Training, eduction and career development

Recommandations classées suivant trois grands 'thèmes'

Infrastructure needs and computing models for IN2P3

Improving software performance and quality

Empowering and strengthening the workforce

Infrastructure needs and computing models for IN2P3

Principalement recommandations niveau grands centres computing et IN2P3

Impacts sur notre environnement de travail

Improving IN2P3's computing infrastructure

1. **Preserve IT ecosystem:** To provide efficient IT solutions to worldwide scientific collaborations, IN2P3 needs to confirm its commitment into a dedicated e-infrastructure.

2. **Optimise the e-infrastructure:** Consolidate at the national level the organization of the resources and services providers, while maintaining the central and stable role of the CC-IN2P3. Sites could be data or CPU oriented, federated with a unique user portal bearing in mind local specifics, experiment models development and the international landscape evolution.

3. **Assure access to new type of hardware resources:** Secure access to large scale GPU resources allowing massive execution of IA and other algorithms, focusing first on the new IDRIS supercomputer converged platform.

4. **Improve synergies:** Connect experts in thematic networks enhancing their role in advertisement and technical user support. Strengthen the links between the CC-IN2P3 and the laboratories for intensified exchanges on user support or innovative solutions.

5. **Influence the international decision process:** Promote the engagement of IN2P3 in European or international initiatives to enhance our impact on the future computing landscape ecosystem and its implementation in France, including R&Ds, experiment computing models or the choice of common tools across experiments and sites.

Nos voisins !?

Landscape is Changing

Improving software performance and quality

Peut on créer ces liens ? Comment ?

Lien edaq

Intensifier l'essor de la ferme GPU

Software

Qualité des software produits

6. **Facilitate AI:** Strengthen collaboration with ML Computer Scientists through the release of open data sets, funding of co-supervised PhD theses, and collaborative projects.

7. **Make use of and extend expertise in Real Time Analysis.** This allows enhancing the scientific throughput of experiments, in particular when facing limited storage resources. This requires using in production advanced algorithms (in particular in Machine Learning) on GPU/FPGA. Make use and extend IN2P3 expertise on ML/DL in real-time applications on innovative infrastructures.

8. **Engage in evolving and emerging technologies**: for the short and mid terms, continuous follow-up of GPU/FPGA evolutions, identification of their possible use especially for online computing and first testbeds or production installations must be encouraged. On a longer term, IN2P3 could benefit from quantum computing. This requires to identify the suited applications to this new paradigm and to monitor a technology survey in order to acquire the necessary software skills and take advantages of similar initiatives worldwide.

9. **Adopt common tools for software development:** Continued effort to provide and release efficient collaborative tools for project management, and the entire DevOps lifecycle

10. **Define and apply common quality standards:** Establish common quality standards at IN2P3. These standards will be accompanied by reference documents, examples, practical recommendations, guidelines and state of the art good practices.

Empowering and strengthening the workforce

Veille et formation

Empowering the workforce

11. **Ensure an attractive work environment:** nowadays IT innovations that may attract engineers are no more coming from the research world. The attractiveness of the technology is no longer enough; it is necessary to play on other levers of career development to ensure the desired skill levels.

12. **Lifelong learning:** In the extremely fast moving field of computing and data science lifelong learning is essential. In addition to existing initiatives, IN2P3 shall further accelerate training of the workforce, e.g. through on-line options, such as tutorials, massive open online courses (MOOCs), and webinars.

Peut on faire qqch avec l'université ?

UNIVERSITÉ DE LYON

Université Claude Bernard Lyon 1

Prospectives IP2I 2019 : 2 grands thèmes

Computing Infrastructure

= Infrastructure needs and computing models for IN2P3

- Batch
- HPC
- FPGA/Hybride
- GPU
- Clouds, container, virtualisation
- Networking
- Databases
- Storage

Algorithms

≠ **Improving software performance and quality**

- Machine Learning
- Parallel Analysis
- Heterogeneous computing
- Symbolic computation

Recommendations :

Increase expertise in

- Machine learning (R6)
- GPU Usage (R8)
- Running algo & neural net in FPGA (R7)
- ‘Modern’ data processing framework (R9) (Spark, Dask, Kubernetes …)

Cooperation between physicists and engineers should also be strengthened

Empowering and strengthening the workforce

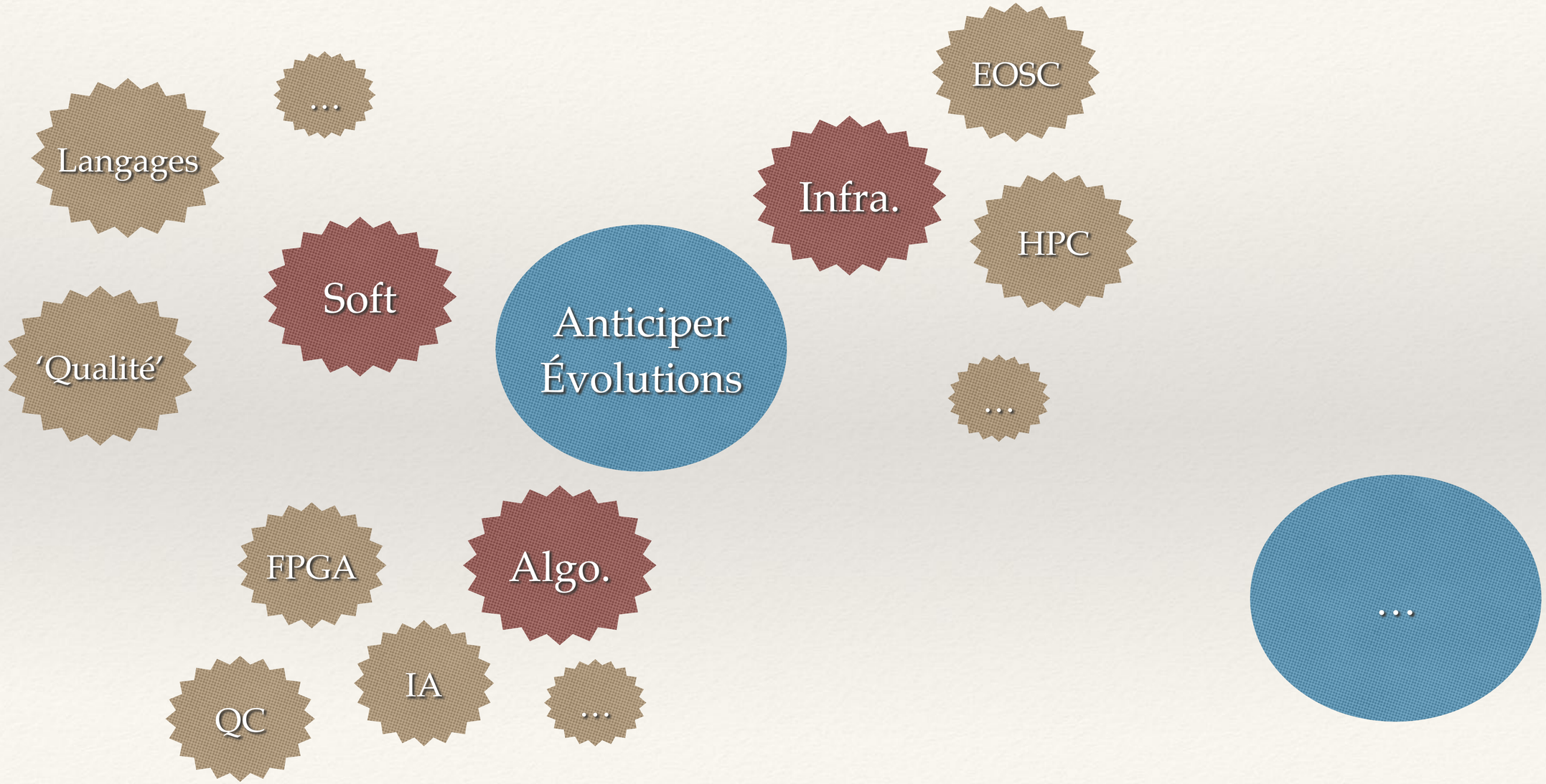
Engineers working with physicist to improve the performance and the quality of the code and its maintainability

set up an interactive analysis farm, potentially using part of the current Tier3 facility. ✔

cpu farm

important to be able to access new hardware for training, prototyping and R&D purposes ✔

gpu farm

+ Groupe Calcul !!!

Questions / Discussions ???