



ESPPU – Discussion transverse : Calcul et Données

Étienne Augé, Sabine Crépe-Renaudin, Federico Ferri

21 janvier 2025

Foreword

Computing is a key element for HEP

- Software and computing are crucial for HEP experiments : their quality has strong impact on the physics output
- Increase of needs: make full use of the fast technology evolution

Organisation

- ~10 contributions with keyword « calcul et données »
- No dedicated transverse GT to discuss these questions
- Feedback from GT1 and GT2 discussions

Proposal

- Short summary of the thematics and recommendations discussed
- Keep most of the time of the session to discuss the recommendations

Computing contributions

- Reflection sur les défis technologique dans calcul et données
- AI and Physics at IN2P3
- + GT1 contribution: [Computing Challenges for the future \(IRN Terascale\)](#)

- Open repository for Machine Learning data
- The Need for Reinterpretation of LHC Analyses
- Reinterpretation and preservation of data and analyses in HEP
- Artificial Intelligence for BSM searches
- Lattice QCD
- LPNHE contribution to the Update of the European Strategy of Particle Physics

Challenges and opportunities

Simulation, reconstruction and data analysis of present experiments and prepare future experiments

Specificities to take into account

- Hardware needs to be changed every 5-7 years => needs continuous funding
- Hardware technologies evolving at high pace, software needs to adapt accordingly and fast evolution of technics as well
 - See developments of AI in the last years for instance
 - Can't be predictive for the long term

New challenges and opportunities

- Increase of hardware an operational cost
- increase of hardware heterogeneity (different CPU, GPU, FPGA)
- Increase of computing needs in society and in science
- Fast developments of AI: contacts with AI communities at various scales
- Sustainability issues

- blue: covered by GT1/GT2
- green: contributions only
- orange : not covered

Thematics (I)

Infrastructure

- Storage an CPU requirements
- Heterogeneous architectures
- Online RTA
- Energy + hardware Costs
- Which organisation for
 - T1/T2/T3
 - Analysis facilities

Data

- Data access
- Date management

Software

- Quality
- Optimisation
- Parallelisation
- Programming language (also linked to heterogeneous architectures)
- Transverse tools to scale in parallel: generators, Geant 4...

Network

- blue: covered by GT1/GT2
- green: contributions only
- orange : not covered

Thematics (II)

Open science (FAIR)

- Data and analysis preservation
- Reproducibility
- Possibility for analysis reinterpretation
- Software
- Publications and associated data (good basis SCOAP, INSPIRE, HepData)

AI

- All levels: design, generation, reconstruction, analysis
- collaboration necessary
- engineer/researcher
- cross disciplinary (math, stat, HEP)
- funding dispersion
- Open data repository / software / models

Person-power

- Training important
- Expertise: needs both engineers and researchers implication
- Attractiveness issue
- French community undersized

Impact

- Energy consumption
- Competition/collaboration with other science and general computing needs
- Take into account implication on computing when designing detectors

- Do you agree with these recommendations ?
- Is there something missing ?
- Do you have suggestions to complete these points ?
- Priorities ?

Recommendation draft

Ensure to build/consolidate solid dedicated infrastructure

- Need to evolve the computing model to take into account hardware heterogeneity
- Increase funding ? (data preservation must be taken into account, increase of cost)

Develop online data processing close to detectors

Ensure development and maintenance of needed software with high efficiency and capability to adapt and take the best of new hardware

Strengthen the computing community and its expertise

- infrastructure, software, technology (IA, new hardware, new language)
- engineers + researchers (+ specialized postdoc ?)
- increase the training
- hiring of young scientists with strong IT skills
- => attractiveness => idea ?

Encourage and coordinate R&D effort

- Thematics : Infrastructure, software, AI and technology
- More coordination needed => pursue JENAs effort, DRD type organisation ?
- Collaborate more with other science and industry

Reinforce collaboration with EuroHPC and the use of external resources (HPC, cloud, local)

- stay opportunistic ?

Minimize the ecological impact

Interplay with instrumentation

- take into account impact on computing when designing detectors

Continue to develop Open Science tools

- Important for re-analysis and AI
- ESCAPE/EOSC