The European Strategy for particle physics

The computing requirements of the highest priority projects within the HEP programme are posing significant challenges which must be addressed head-on. The four LHC experiments (ATLAS, CMS, LHCb and ALICE) recorded almost 1 exabyte of data to make groundbreaking discoveries, from the Higgs Boson to pentaquarks. The aggregate requirements of non-LHC experiments add significantly to this challenge. Major upgrades of the LHC accelerator and experiments are due to take place in the 2020s which will result in vastly increased datasets, while in parallel we will see the build-up to operation of the next generation of neutrino experiments (DUNE and HyperK) with LHC-scale volumes of highly granular data.

This need was emphasized in the update of the European Strategy for particle physics[1]:
"The community must vigorously pursue common, coordinated R&D efforts in collaboration with other fields of science and industry, to develop software and computing infrastructures that exploit recent advances in information technology and data science. "

Developments in the UK

The diagram below shows how several projects within the STFC remit help meet these needs. In the diagram, the bottom layer is the physical hardware and its management, provided by GridPP as part of the global WLCG infrastructure. At the centre is the common software for the benefit of the HEP community, which is provided by the newly established SWIFT-HEP project (Software Infrastructure and Technology for HEP, under final review by Science Board). The top layers are experiment specific and are funded by the experiment grant lines.

Large scale computing and storage

GridPP – provides the hardware and services that form the UK distributed (or Grid) computing infrastructure that meets the requirements of the LHC experiments, and others. This is primarily funded by STFC but with significant contributions from the Universities in the form of data-centres, power, additional hardware and manpower, and network bandwidth. The scale of direct-funding has been of the order of £6m/year for the last two decades but over that time the infrastructure has grown in capacity by several orders of magnitude as the LHC experiments have ramped up to the exascale. Over the past few years, GridPP has also started to provide significant hardware capacity and services to non-LHC PPAN communities as a provider to the IRIS consortium. This has levered GridPP experience and expertise benefiting other parts of the STFC community. Iris now provides essential additional capacity for those HEP projects with significant requirements, including DUNE and eventually HyperK, as well as non-HEP projects such as the LIGO, CLAS12 (Nuclear) and CTA, the Vera Rubin Observatory (formerly LSST) and SKA. This additional capacity can only be deployed and delivered by leveraging the GridPP project.

[1] https://home.cern/sites/home.web.cern.ch/files/2020-06/2020%20Update%20European%20Strategy.pdf

Layer	Domain	Experiment 1	Experiment 2	Experiment 3
6	Physicists	Analysis code	...	...
5	Experiment Physicists programmer and software engineers	Analysis framework. Simulation, Reconstruction, Calibration Code	...	...
4	Experiment Software Engineers	Software Frameworks	...	...
3	Common Software HSF / SWIFT-HEP	Common software components (Data management, Generators, Geant4, Accelerator integration)		
2	GridPP / WLCG	Middleware infrastructure for Distributed Computing		
1	GridPP / WLCG	Physical Hardware		

Moving Software down the stack

Software. Experiments and central infrastructure

The experiments – are the primary providers of the software required in the top-layers of the diagram, and experiment specific effort is funded through the consolidated grant process and upgrade grants. It is critical that this component of the experiment infrastructure is funded to ensure the STFC investment in constructing and operating the experiments has a full return in terms of physics exploitation.

Some of the software components are non-experiment specific and development effort is needed to keep the software efficient and competitive. The experiments coordinate effort and run shared projects through the Hep Software Foundation (HSF) internationally, and in the UK with SWIFT-HEP. SWIFT-HEP – funds additional, typically joint software effort to meet the upcoming challenges from exascale and evolving architectures.

Vision and evolution

The current software and computing infrastructure has provided a sound foundation for the HEP community as a whole. The needs and requirements are changing over time, as highlighted by the European strategy for particle physics. The increase of complexity of the experiments and the necessity to record larger datasets with increasing data rates, mean that a more than linear increase in our current software and computing provision is needed.

The full solution cannot be just to increase the amount of computing hardware, which would cost too much to deploy, operate and maintain, and would vastly increase the carbon-footprint. A broad approach is required that includes:

- An increase of the amount of compute cycles (measured in HEP SpecInt 06) and storage space (in Exabyte units) that is to be expected from a flat, but inflation-protected, budget as the cost of hardware decreases
- The deployment of more modern and efficient technologies, reflecting the evolution that we are seeing within industry and other scientific fields. This requires an increased investment in skilled Research Software Engineers, as well as training of the existing users base to use these techniques.
- The hardware evolution in the coming years, such as specialised hardware for deep learning, and possibly new disruptive technologies such as quantum computing. As a community we should be prepared to take these opportunities to lead and innovate
- The community spirit across Universities, which needs to be maintained to enable innovations and new ideas to take place, such as in the current model of GridPP, the experiments, and SWIFT-HEP.