Answers of Particle Physics Community: Federated Infrastructures

1 Which infrastructures are presently used ?

- Highly federated computing model with >150 sites in the world-wide LHC computing grid;
  strong contributions by compute sites at universities;
  analysis facilities, optimised for the final steps of physics analysis at some centres and at universities provide high data-troughput
- Wide variety of jobs, ranging from low/IO Monte Carlo simulation to high I/O skimming and data distribution;
  - typical access is „write once, read often“
  - federated storage model: remote access to data via WAN possible
  - handlig of large Data volumes of tens of PB / y,
  - data transfers of tens of GB/s
  - # of jobslots, typical duration of 10h
- GPUs explored in test environments: offer large potential for track reconstruction
  - FPGAs are not used

2 Future Requirements

- Data Volume per year will increase by a factor of 10 after the year 2025 (HL-LHC), reaching the scale of one EB/y
  - higher complexity of events leads to non-linear increase of CPU requirements
  - expected performance gain through technical advancements will not counterbalance

  → need for factor ~10 more of storage capacity & network bandwidth
  factor 6 more CPUs needed if extrapolating from present computing models
- Almost all applications are batch-processing, except filter farms at the detectors or rare cases of interactive data analysis
- Access pattern remains write-once, read often; read rights of all data by every member of a collaboration, write acces restricted to small group of users.

- Exploring new storage technologies (SSDs, non-volatile RAM etc.) important, requiring R&D
- „Cloud technologies“ will enable contributions by different kinds of providers, incl. universities

2 Experience with cloud technology

- Being explored: Helix Nebula Project; experience with commercial cloud providers and usage of HPC centres with jobs encapsulated in a container
- Encapsulating jobs in containers (singularity) has become the standard by now; such applications also run in cloud environments.
- There are substantial requirements for development:
  - integration of non-WLCG resources in workflow management,
  - efficient job scheduling in heterogeneous environments,
  - light-weight creation of virtual computer centers
  - provisioning of data to virtual computer centers