Computing Frontier: Answers to Questions from other Frontiers

Computing Frontier Snowmass Group

Text answers is at:

https://github.com/SnowMassComputing

COMF1

- Theoretical physics requires increasingly more significant computing resources. How is this being incorporated into computing plans? EF-Computing: To what extent is high-energy physics still generating the world's largest randomly-accessed databases? Can we claim to be a world leader in data science? Along what dimensions?

COMF1 Answers
Theory needs

- Ongoing planning meetings like recent one at NERSC to determine computing needs of the Theoretical community.
- User requirements for exascale computing

COMF1 Answers
HEP lead in data science

- petabytes of root files (as in LHC data) would not be considered a random access database today.
- In terms of piles of data that can arguably be called databases, it is probable that the customer behavior databases amassed by eBay, Google etc., most of which would be classified as NOSQL databases, have left us behind.
- We still lead in worldwide distributed data management in a heterogeneous ensemble of independent computer centers.
- The growth in data of the Energy Frontier, in aspects not needed by the standard bit streaming providers, drives continued investment in data management and data access methods. Continued evolution will be needed in order to take advantage of new network capabilities, ensure efficiency and robustness of the global data federations, and contain the level of effort needed for operations.

COMF2

- The Grid was commissioned along with the LHC detectors. ESnet traffic has increased 10x every four years throughout the LHC lifetime. Will improvements in networking infrastructure, QoS, monitoring, etc. continue to keep up with LHC demands for distributed computing? In what directions are new enabling technologies required and when must they mature to again keep up with the LHC machine and detector upgrades?

COMF2 answers

- There are no fundamental technical barriers associated with transporting 10x more traffic in 4 years.
- Beyond that, basic and applied research is necessary to develop cost-efficient architectures
  - assure that networks and applications become more tightly integrated.
- number of cultural and operational habits need to be overcome in order for networks to continue to keep up with LHC demands:
  - campuses must deploy secure science data enclaves - or Science DMZs - engineered for the needs of data-intensive science,
  - number of cultural and operational habits need to be overcome in order for networks to continue to keep up with LHC demands: campuses must deploy secure science data enclaves - or Science DMZs - engineered for the needs of data-intensive science,

COMF3

- How do the different physics frontiers--and associated theory and physics simulation--differ in their needs for future computing technology evolution? In what respects can they benefit from common computing technology evolution?}

COMF3 answers

- Writing efficient codes is likely to become more difficult as we move to more exotic processors like GPUs or the Xeon Phi.
  - Common problem among various computing frontier projects

COMF4

- Proposed very high statistics experiments at the Z resonance require large rates -- many kHz -- at which data is written to storage. What are the limits?

CMF4 answers

- The limits are those of tolerable cost for storage and analysis. Tolerable cost is established in an explicit or implicit optimization of physics dollars for the entire program. The optimum rate of data to persistent storage depends on the capabilities of technology, the size and budget of the total project, and the physics lost by discarding data. There is no simple answer!

COMF5

- What are the requirements and opportunities for cosmological computing -- in both theoretical simulations and data analysis -- to enable the extraction of new particle physics information from astrophysical observations over the coming decade?

COMF5 answers

- a major component of the Cosmic Frontier report. We'll get you a compressed version soon.