Preparing for Discovery with the Large Hadron Collider

Frank Wuerthwein
UCSD
9500 Gilman Drive, MC 0319
La Jolla, CA 92093–0319, USA

June 11, 2010

In late 2009 a new experimental facility, the Large Hadron Collider (LHC) started its first proton proton collisions. Two large experiments, Atlas and CMS, are getting ready to explore the high energy frontier of elementary particle physics at this facility.

The known substructure of the protons colliding in the LHC, in combination with energy and brightness of the proton beams guarantee a many orders of magnitude jump in production rates for physics processes at the highest presently available energies, providing a once in a lifetime opportunity for discovery.

Each of these detectors is comprised of order 100 Million electronic channels that are designed to detect the "debris" of proton-proton collisions at a rate of 40MHz. A mix of zero-suppresion and realtime filters reduce this information rate from order 10 PetaBytes/sec to a mere 0.1 to 1 GigaBytes/sec. Multiplying this rate by 107 seconds per year of collisions, and we arrive at an annual expected data volume of up to 10 Petabytes from each experiment. To make sense of this data requires development of complex calibration, alignment, and processing algorithms, that are likely to receive refinements for the first several years of operating the detectors. One thus expects multiple passes of reconstruction of the entire data, as well as portions thereof. In addition, detailed Monte Carlo simulations are required to fully understand the detector as well as reconstruction algorithms.

Sociologically, all of this is accomplished by globally distributed collaborations of thousands of physicists and engineers. The fundamental computing challenge is thus to design and operate a distributed computing system that maximizes the productivity of these global collaborations via a mixture of

central and distributed control of petascale computing and storage resources. This challenge is met via a globally distributed multi-tiered hierarchy of computing centers.

We will present the current status of data taking with ATLAS and CMS, the expected reach for some of the new physics searches using data to be collected in 2010/11, and briefly touch upon the computing challenges that are being overcome in order to reach the physics goals.