Searching for Extreme Events in Multi-lepton Data From the LHC

Mentee: Xinyue Wu (Univ of Rochester)
Mentors: Ben Tovar, Kelci Mohrman (Univ of Notre Dame)

Motivation

- SM particles cannot represent the complete set of nature's constituents

- Particles at a higher mass scale have the potential to lead to new physics

- Associated Top Quark Production
    - Rare, heavy particles involved, not studied much
    - Various final states: **multi-lepton final state** (interest of study)
        - Clean signatures, efficient triggers, few backgrounds

Challenge

- There's no guarantee that new particles would be light enough to be produced on-shell at the LHC
- Indirect methods of probing higher mass scales become crucial in the search for new physics at the energy frontier
  - **Effective field theory (EFT)** is one of the probes
    - Relatively model independent
    - Flexible (can describe various new physics effect)
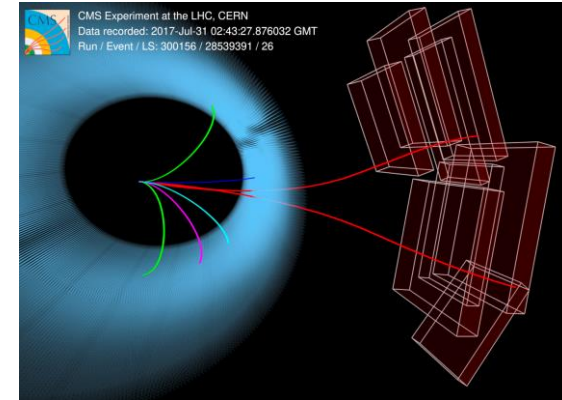
Work to do

- Large quantities of data need to be analyzed to search for extreme events

- Data from CMS preselected and classified in histograms according to the number of jets, b jets, and leptons

- Highest multiplicity categories are so far inclusive

  - E.g. 4 lepton category includes information about events with 4 or more leptons.

What were the events with the largest number of leptons? What was the event with the largest number of jets (or b jets)?...

Methodology

➜ Modify a topcoffea processor to detect events with interesting characteristics

  ◆ E.g. large number of leptons

➜ Accumulate information gathered from events into a dataframe

➜ Render the events using CMS visualization tools

➜ Gather performance data of the distributed computation

➜ Identify the bottlenecks in the computation as it is

  scaled to use hundreds of cores

  ◆ E.g. bandwidth, file access, cpus, memory

CMS Experiment at the LHC, CERN
Data recorded: 2017-Jul-31 02:43:27.876032 GMT
Run / Event / LS: 300156 / 28539391 / 26

Components

➔ **topcoffea**: An application designed to analyze particle physics data collected by the CMS experiment to study EFT.

➔ **Coffea**: Framework that provides all of the typical needs of a high-energy physics experiment analysis using the scientific python ecosystem.

➔ **Work Queue**: A system for creating and managing scalable manager-worker style programs that scale up to tens of thousands of cores on clusters, clouds, and grids.

➔ **xrootd**: Data federation for the CMS data.

➔ **CMS visualization tools**: https://cds.cern.ch/record/2751566?ln=en

Summary

❖ An analysis processor will be written to search for extreme events in CMS data.

❖ The processor will be written with the topcoffea framework.

❖ The analysis is expected to run in O(100) cores.

Backup