#### **Solution to writing MC Raw Digits to HDF5**

(Open to Discussion)

Barnali Chowdhury, Tom Junk

DOE 2022 Computing Consortium Meeting 08.30.2022





### **Challenges**

#### Fundamental and Conceptual to understand

- The idea of event based loop that art defines
- Struggling with many APAs
- Storage Limit using root with Raw Digits: 1GB limit on 1 TTree entry in ROOT.
- Performance Issue: memory needed to store all the raw digits before any of them can be written out using the current art model.

#### Few ways to solve that

Either at the art layer or in the WireCell Toolkit level

### **Consumers of Raw Digit**

#### How the Raw digits will be consumed

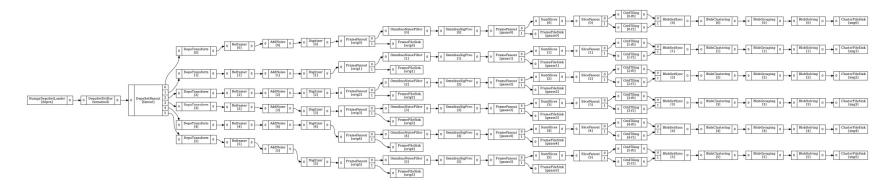
- There are 2 1st consumers of RDs, both are dealing with noise.
- Noise Filter Inside the WireCell Toolkit.
- DataPrep An Art module.

- 2<sup>nd</sup> consumer TPC Signal Processing (also in WireCell)
- This connects to the output of the 1<sup>st</sup> two consumers.
- The output of DataPrep goes to WireCell via Art event with all APAs being loaded.

## WireCell Toolkit Graph

- The RDs for simulation are created by WireCell.
- Utilize Wirecell Toolkit Graph.
- The graph has nodes that read some art event store and provide input to the graph.
- Some nodes in the graph will take the output from the graph and write it back to the art event.
- Not necessary to write to an art event. Could write to a file instead.

### WireCell Toolkit Graph



WCT flow graphs

- Shows a graph that performs simulation, noise filtering, signal processing and 3D imaging for ProtoDUNE-SP using I/O methods supported by WCT.
- Data is read from file, drifted and 6-way fanned out.
- Each "pipeline" subgraph that you see handles the processing related to one APA

## WireCell Toolkit Graph

#### When do we need to write RD objects to a file?

- In WCT, the waveforms are written in simple arrays (in NumPy).
- Provides with individual APA files.
- Collect the files.
- Could write them in a giant frame.
- Run HDF5 writer output module.

# Thanks!

