SciDAC4 “HEP Analysis on HPC”

Jim Kowalkowski

*μ*Retreat

19 April 2018

SciDAC-4: HEP Data Analytics on HPC

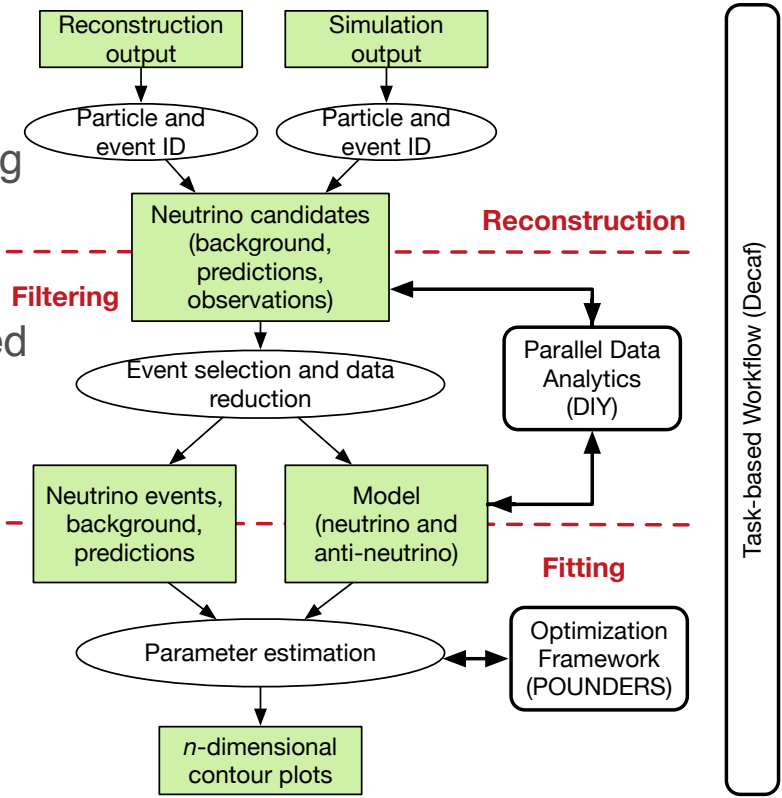
- Enable HPC facilities to meet the future HEP data analysis demands, allowing computationally expensive physics studies to be completed on time scales that are not currently feasible.
- Scientists from HEP and ASCR
  - LHC and Neutrino Science: Norman Buchanan (CSU, NOvA/DUNE), Paolo Calafiura (LBNL, LHC-ATLAS), Zachary Marshall (LBNL, LHC-ATLAS), Stephen Mrenna (FNAL, LHC-CMS), Andrew Norman (FNAL, NOvA/DUNE), Alex Sousa (UC, NOvA/DUNE)
  - Optimization: Sven Leyffer (ANL), Juliane Mueller (LBNL)
  - Storage and Data Modeling: Rob Ross (ANL), Marc Paterno (FNAL), Saba Sehrish (FNAL)
  - Workflow: Marc Paterno (FNAL), Saba Sehrish (FNAL), Thomas Peterka (ANL)
- DOE SciDAC program manager support
  - Lali Chatterjee (Computational HEP, Office of High Energy Physics)
  - Randall Laviolette (SciDAC Application Partnerships, ASCR)
- Official web site is http://computing.fnal.gov/hep-on-hpc/

U.S. DEPARTMENT OF ENERGY | Office of Science

Fermilab | Argonne NATIONAL LABORATORY | Lawrence Berkeley National Laboratory | Colorado State University | University of CINCINNATI

Neutrino physics

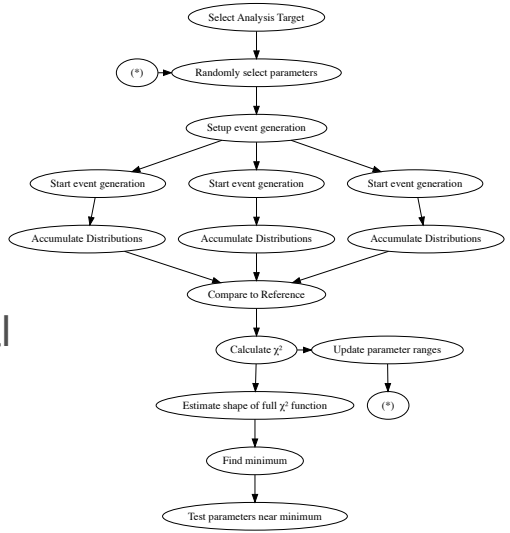
- Linking neutrino oscillation measurements back to the underlying neutrino oscillation and nuclear interaction (cross section) parameters
- Requires multi-dimensional parameter spaces and fitting the observed data to the underlying physics models
- Advancements using HPC leadership facilities
  - Analysis using full dataset across all layers, managed using tools and techniques developed by ASCR's **RAPIDS** institute
  - Expand fitting into 4D, 5D, including anti-neutrino data, utilizing multi-dimensional fitting procedures from ASCR's **FastMATH** institute
  - Include all layers in the analysis



Fermilab

Generator tuning

- Using **PYTHIA8, MadGraph, Rivet** and associated tools
- Reducing the theoretical uncertainty on Standard Model quantities extracted from collider data and on background estimates
- Automation of the tuning process
- Advancements using HPC leadership facilities
  - Modernize optimization process: **POUNDERS, Minotaur**
  - Large-scale complete workflow automation of tuning process: **DIY, Decaf**
  - Apply surrogate model optimization algorithms for finding the optimal tuning parameters
  - Adaptively select parameters in the minimization process
- Starting with the **PROFESSOR** package
  - Defines χ^2 to minimize based on $f^{(b)}$, using generator prediction $GEN_b(\mathbf{p})$
  - Producing improvements to the polynomial methods



Fermilab

Crude timeline

CMS/ATLAS Generator Tuning	Current generation analysis – new tools	Next generation techniques with automation	
NOvA/DUNE Oscillation/Cross-section analysis	Use existing code and tools	Oscillation analysis Next Generation	Cross-section analysis

- Focus on whole-dataset analysis, utilizing whole machine
  - Introduce optimization and workflow tools from HPC community
  - Incorporate simulation, reconstruction, and selection layers into analysis procedures
- Generator analysis and tuning will lead in development and utilization of new tools and techniques
- Experiment analysis will start with existing codes and application to establish baselines, next generation systems will adapt and utilize techniques from generator groups