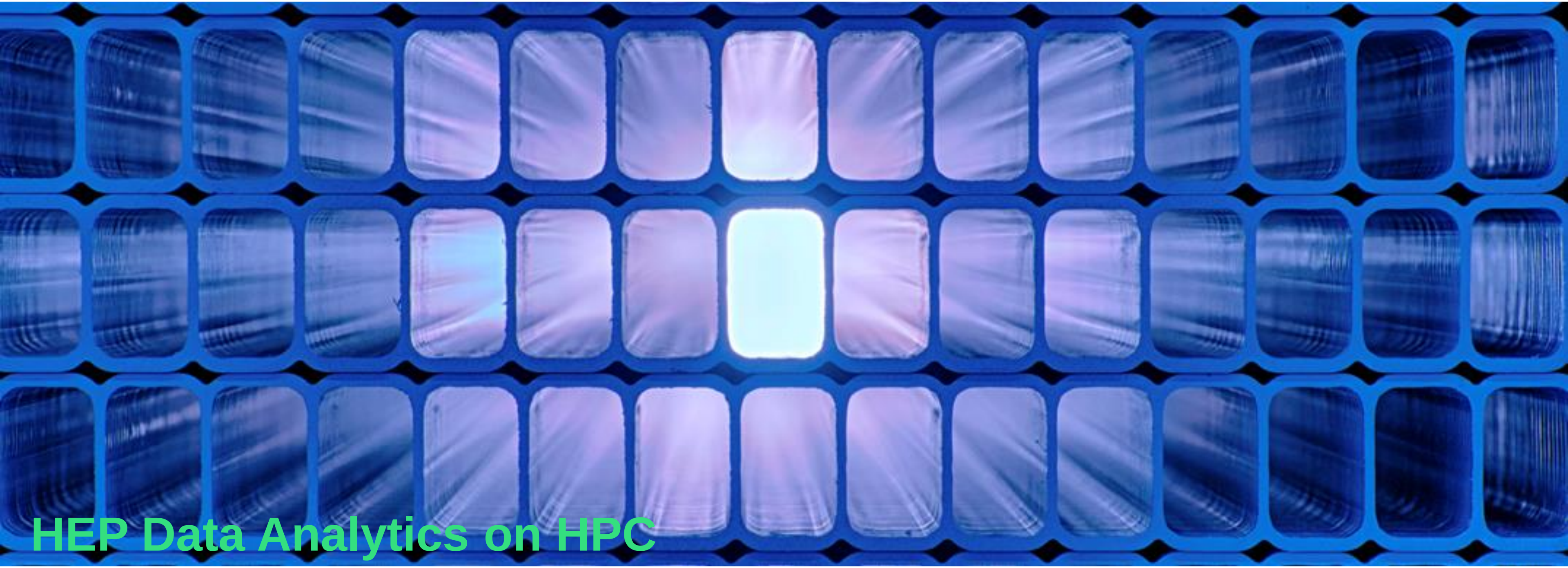


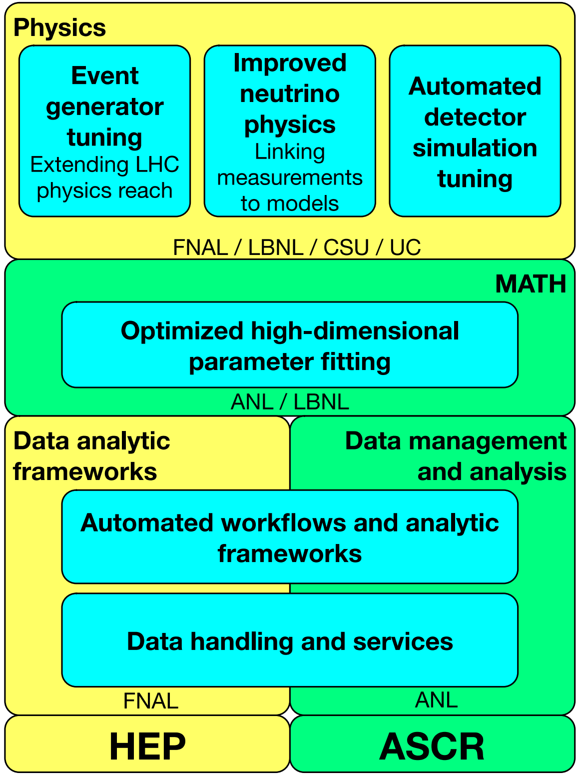
HEP Data Analytics on HPC

Jim Kowalkowski

SciDAC PI Meeting July 17, 2019

HEP Data Analytics on HPC: Goals

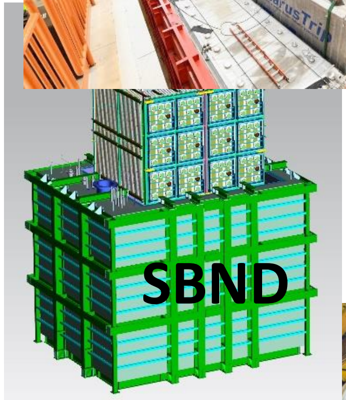
- **Physics component:**
  - Participate in getting new physics results out
  - LHC and neutrino experiments
- **Accelerate targeted HEP analysis on HPC platforms**
  - Working within experiment code and operations
  - In some cases by orders of magnitude
- **Transform how physics tasks are carried out using ASCR tools and techniques**
  - Parallel storage and access for experiment data
  - Data-parallel programming and advance mathematical techniques for HEP analysis

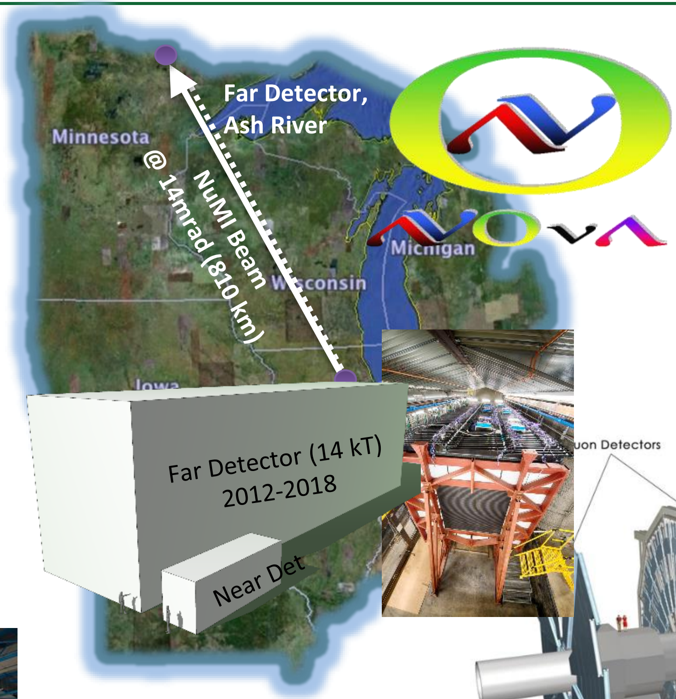


http://computing.fnal.gov/hep-on-hpc/

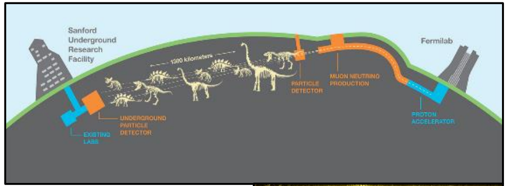
We are working with many HEP experiments

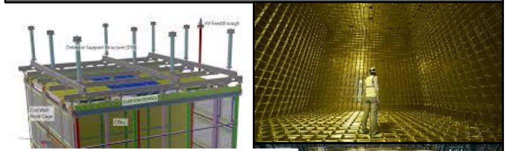
SBN LAr Program

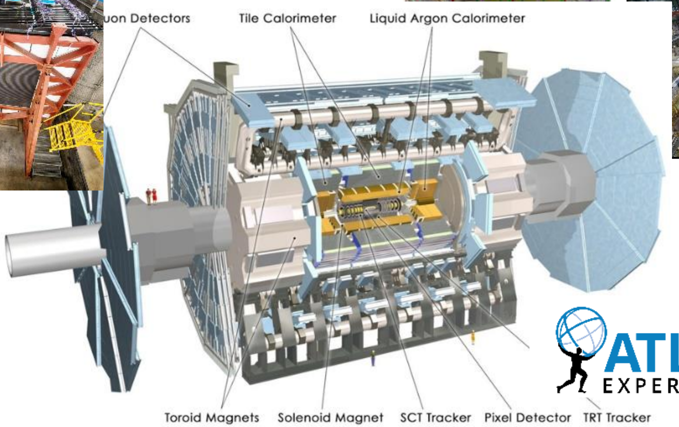




DEEP UNDERGROUND NEUTRINO EXPERIMENT







ATLAS EXPERIMENT

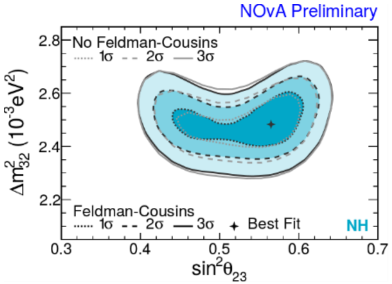
Measuring 3-flavor neutrino oscillations parameters with the NOvA experiment

Scientific Achievement

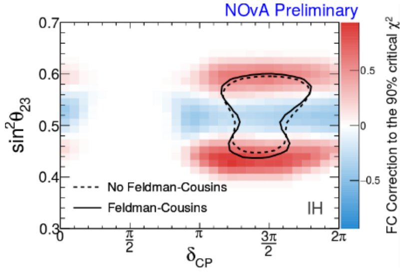
Leveraged HPC resources to perform complicated fits of neutrino oscillations data and determine confidence intervals for fundamental parameters.

Significance and Impact

World leading constraints on the neutrino mass hierarchy, the mixing angle θ_{23} and the CP violating phase δ_{CP}



Confidence intervals (1,2,3) with and without statistical corrections



Effect of the Feldman-Cousins correction on the 90% confidence contour: the blue (resp. red) areas show an increase (resp. decrease) of signficance.

Research Details

- 51 million hours consumed over two weeks for minimum disruption at NERSC.
- Would have taken one year on FNAL grid computing resources.
- Improved workflow with DIY : exploit near prefect strong scaling with hundreds of thousands of MPI ranks
- Unprecedented accuracy of statistical corrections.

Acero, M.A. et al. (NOvA Collaboration) (June 2019) arXiv:1906.04907

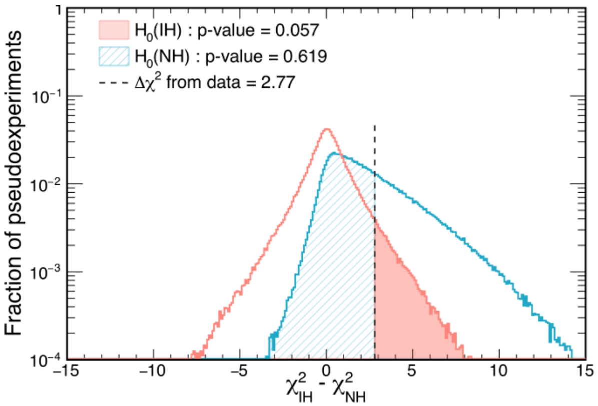


FIG. 5 CL_s pseudoexperiment distributions testing the Inverted and Normal neutrino mass hierarchy hypothesis for the 2019 NOvA datasets. Analysis required approximately 300k pseudoexperiment fits for each hypothesis to achieve a p-value of 0.057 or 1.9σ for Inverted mass hierarchy rejection.

U.S. DEPARTMENT OF ENERGY | Office of Science

Fermilab | Argonne NATIONAL LABORATORY | University of CINCINNATI | Colorado State University

Understanding neutrino-nucleus interaction in Long-Baseline Neutrino oscillation experiments

Scientific Achievement

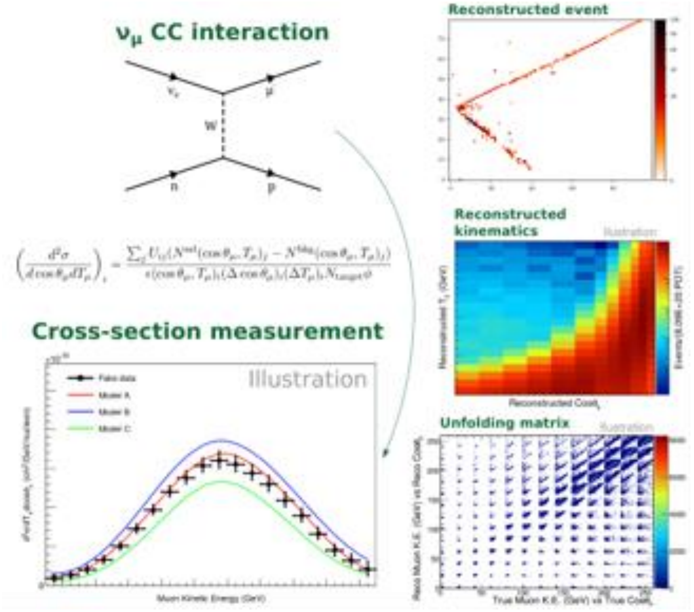
First neutrino cross-section analysis taking advantage of HPC resources (NERSC) : unprecedented accuracy in the study of systematic effects in neutrino-nucleus cross-section measurements.

Significance and Impact

Improve the understanding of neutrino-nucleus interactions for Short and Long-Baseline Neutrino Oscillation experiments (SBN, NOvA, DUNE). NOvA will report an accurate measurement of the main signal cross-section in long-baseline experiments.

Research Details

- Thousands of unfolding matrices were constructed from hundreds of thousands of Monte-Carlo simulation files and for many variations of hundreds of systematic parameters, called universes.
- 10 time more universes were studied and results were obtained 200 times faster.



Measurement of the ν_μ CC cross-section in NOvA's Near Detector: The ability to study in detail the effect of the systematic uncertainties on NOvA's measurement allows us to decrease the reported cross-section uncertainties, to discriminate between different interactions models and therefore to improve our understanding of neutrino-nucleus interaction models.

Validation of parameter estimation in 3+1-flavor long-baseline neutrino oscillations at NOvA

Scientific Achievement

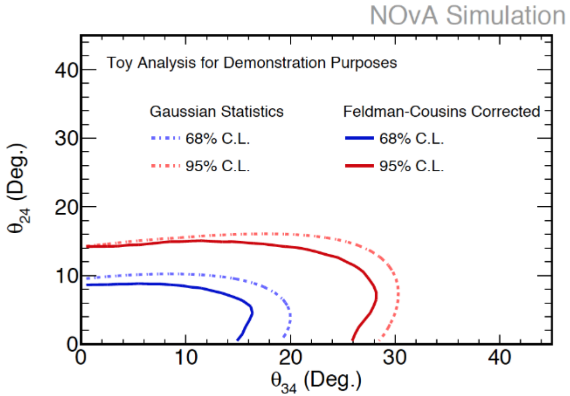
<u>Demonstration</u> of the Feldman-Cousins (FC) approach in a covariance matrix-based fit framework using the 3+1-flavor sterile neutrino oscillations model

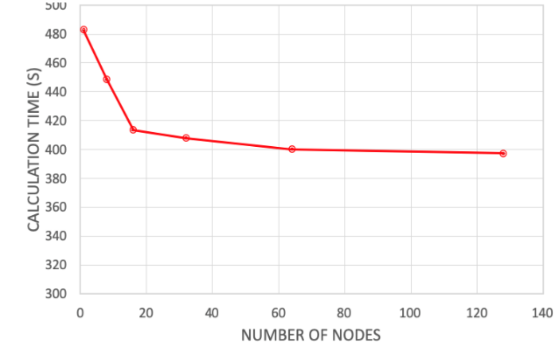
Significance and Impact

We have extended the previous 3-flavor, extrapolation-based NOvA analyses to a 3+1-flavor model, introducing 4 additional free oscillation parameters. Results of our study indicate the paramount importance of the FC approach in order to obtain maximally precise measurement of sterile neutrino parameters

Research Details

- 110K NERSC hours for test, <u>full run in July</u> 2019
- <u>5x faster</u> than current method of systematic uncertainty handling
- Over 500 pseudo-experiments are computed at each point in parameter space to determine the empirical statistical distribution
- <u>This is a new approach to parameter estimation within NOvA, enabling higher dimensions to be incorporated into calculations</u>.





Toy Demonstration of FC approach in a 3+1-flavor covariance matrix fit: (Top) Effect of the FC approach on selected intervals of fixed confidence level; (Bottom) Calculation time as a function of number of nodes used.

Rational functions yield better approximations than polynomials in HEP applications

Scientific Achievement

<u>Numerical approach for computing multi-variate rational approximations (RAs) with and without poles based</u> on linear algebra and semi-infinite programming (SIP).

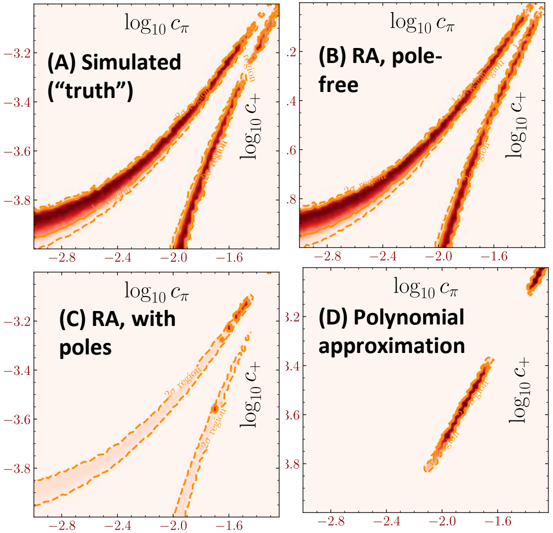
Significance and Impact

RAs are more flexible and accurate than polynomials. RAs are particularly useful for approximating expensive-to-compute HEP simulations, which enables us to gain further physics insights.

First algorithm for pole-free multivariate rational approximations.

Research Details

- Physics simulations are commonly used to understand observed phenomena, but simulations are computationally expensive, which makes direct optimization of the simulation's parameters difficult
- We use rational functions (polynomial divided polynomial) to approximate the simulated data
- Two methods for computing RAs: Stieltjes process (may lead to RAs with poles, *Fig. C*); SIP (iterative removal of poles, *Fig. B* <u>at 1/50 the CPU cost</u>)
- SIP yields significantly better approximations of the true data (*Fig. A*) than RAs based on Stieltjes process (*Fig. C*) and polynomial (*Fig. D*); shown is data of direct detection of dark matter
- Fermilab preprint **FERMILAB-PUB-19-330-CD** to be published in SIAM



Techniques for approximation that are used widely do not work well at all when there are poles

Massively parallel Monte Carlo simulations with HPC

Scientific Achievement

New event generation and analysis framework suitable for massively parallel processing at HPC facilities

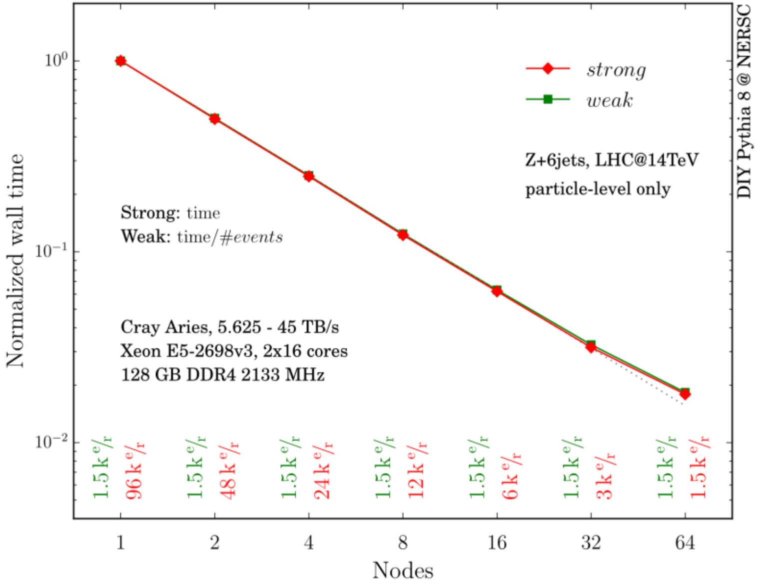
Significance and Impact

Best simulation of V+jets: arXiv:1905.05120 [hep-ph] (in collaboration with S. Hoeche)

Generic Monte Carlo event simulation with efficient use of HPC resources

Research Details

- Compute parallelism with ASCR's DIY
- Data parallelism with HDF5
- Particle level simulation with Pythia8
- Analysis with Rivet
- Supports reading standardized LHE events in HDF5 format



Scaling of particle level event simulation at NERSC. We observe very good scaling up to 2000 ranks.

Repositories:
- https://bitbucket.org/iamholger/pythia8-diy
- https://bitbucket.org/iamholger/lheh5

U.S. DEPARTMENT OF ENERGY | Office of Science

Fermilab | Argonne NATIONAL LABORATORY | University of CINCINNATI | Colorado State University

Parallelizing important algorithms

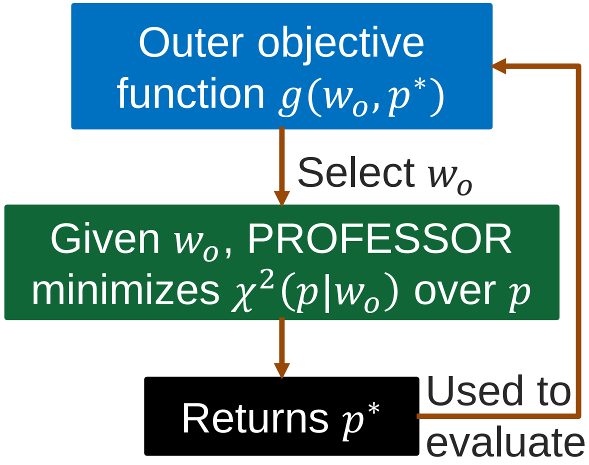
- **Feldman-Cousins prescription is ubiquitous in HEP**
  - Established procedure required for major neutrino physics results across experiments.
- **A method for producing sensitivity contours in multi-dimensional parameter fits (corrected frequentist confidence internals) accounting for non-gaussian systematic uncertainties**
  - Referred to as the *Unified Approach* by Feldman-Cousins
  - Implemented using multi-universe techniques
- **We are well-established with NOvA**
  - Requires millions to tens-of-millions of core-hours to calculate within NOvA depending on dimensionality and desired resolution
- **First demonstrations for liquid Argon (LAr) experiments**
  - DUNE using this now for Technical Design Report (TDR)
  - SBND now for TDR studies (answers question of how much data to collect for the study)
  - ICARUS very soon, real SBND (around 2021 for real thing), perhaps MicroBooNE
- **Parallelization underway**
  - DIY framework, HDF5 for parallel I/O, vectorized log-likelihood,
  - Modernization of optimizer

Bilevel optimization for automated selection of tuning-relevant observations

- Fitting to models that are known to be incomplete or imperfect
- Exploit optimization methods from finance and design of experiment

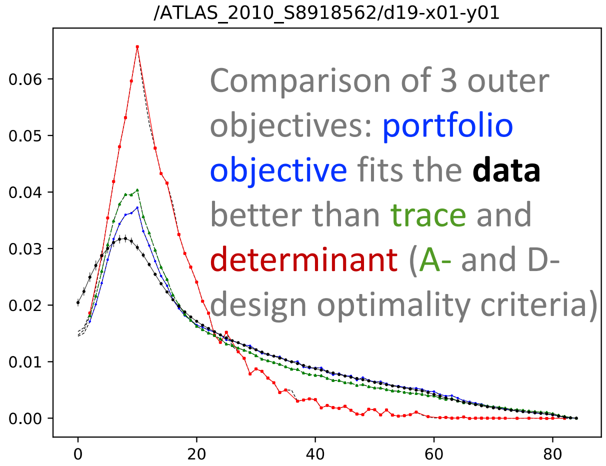
---

- Upper level: Find optimal w_o that determine which observations should contribute to the tune
- Lower level: PROFESSOR minimizes a misfit function χ^2 over a set of physics parameters p given a set of weights w_o



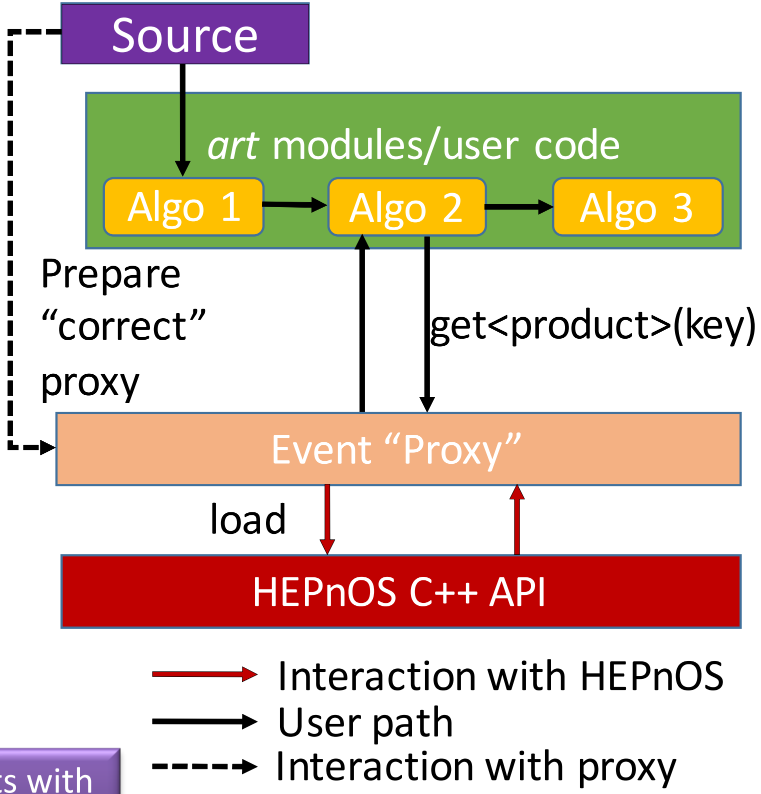
Two Nested Optimization Problems

1. Minimize objective function given data (histograms) and predictions
2. Adjust the importance (weight) of different data to remove outliers and poor modeling



Vision for managing HEP data on HPC systems

- Making high-volume reconstructed physics object data available to analysis workflows (on demand)
  - Leverage existing modular framework abstraction and extensible data models
  - Runtime replacement of ROOT with tools from RAPIDS for I/O
  - HPC facilities enable this
- Include levels (or layers) of data aggregation with metadata
  - Datasets / runs / intervals / events / physics object containers
- Exploit data independence
  - data parallel programming and implicit parallelism
- Allow facility services to distribute data at any scale, using existing abstractions



Event currently interacts with *art*/ROOT File

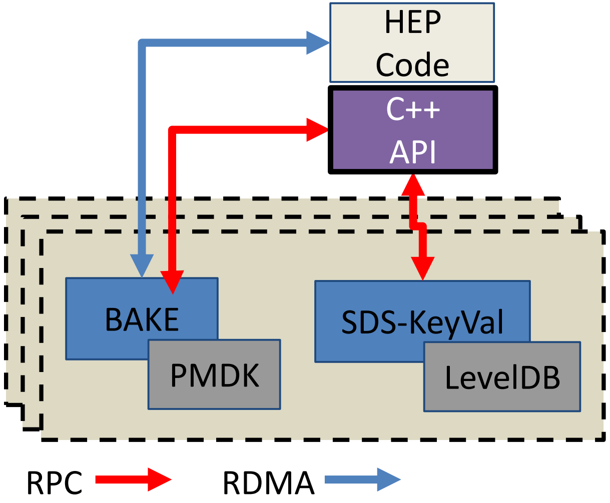
HEPnOS: Fast Event-Store for High-Energy Physics (HEP)

Goals

- Manage physics event data from simulation and experiment through multiple phases of analysis
- Accelerate access by retaining data in the system throughout analysis process
- Reuse components from Mochi ASCR R&D project

Properties

- Read in data at start of run and write results to persistent storage at the end of a campaign
- Hierarchical namespace matching physics concepts (datasets, runs, subruns)
- C++ API (serialization of C++ objects)
- Write-once, read-many



The HEPnOS event store is built using Mochi, a framework for developing specialized data services for use in HPC. Mochi allows use of state of the art libraries (e.g., LevelDB for key/value storage) while providing convenient APIs to scientists.

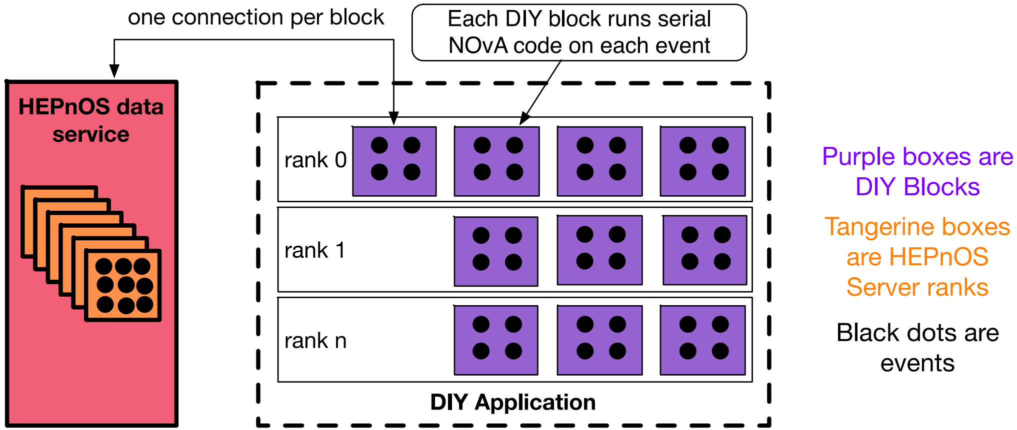
Parallel neutrino candidate selection using HEPnOS

Achievements

Faithful reproduction of NOvA's serial neutrino candidate selection procedure leveraging HEPnOS, a custom HPC data service for HEP.

Significance and Impact

Bypasses file system to accelerate data access throughout analysis process (a major bottleneck). Enables existing serial experiment code to run efficiently in highly parallel context.



This figure shows the organization of our neutrino candidate selection application (DIY) and its communication with the HEPnOS data service.

Research Details

- Demo of NOvA's "4th analysis" candidate selection
- DIY-based application, using data structures and candidate selection code from NOvA.
- Our application code provides all the parallelization, and experiment code is unaware of parallelism.
- HEPnOS data service supports the parallelism by providing global view of data and by removing software artifacts related to the filesystem.
- Using containers to deliver complex software stack to ALCF and NERSC.

Summary

- **Thank you for the support of this project**
- **Greatly appreciate the hours we have at NERSC**
  - We have been able to conduct parts of physics analysis that may never have been completed without this project
  - Working to expand into ALCF
- **Please come visit us at the poster session for more details**

This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Advanced Scientific Computing Research, Scientific Discovery through Advanced Computing (SciDAC) program.

Summary

- **Currently working with the intensity and energy frontiers**
  - NOvA, SBN, DUNE, ATLAS
- **Working with people embedded within experiments is vital**
  - Untangling algorithms & data manipulation from applications is very difficult
  - Aiming for adaption of tools
  - Building upon success: some are SciDAC4 project members, some or not
- **Working together with ASCR RAPIDS and FASTMath has been fruitful**
  - We have begun to explore high-impact areas that otherwise would not be touched
  - Want to expand discussion with other groups
- **We want to continue collaborating with other SciDAC projects**
  - Stefan Hoeche and generators

Summary of project plans and trajectory

Analysis

Year 1

- **Use existing analysis** application
- Large analysis on HPC
- Learning to use HPC

Year 2

- **Basic parallelism added**
- Simplify operations and data handling
- Several large analysis on HPC

Year 3

- Parallelizing kernel functions.
- New optimizers for higher dimensions.
- High-level workflows for sensitivity studies and tuning.
- New physics analysis campaigns.

Simulation & generators

Year 1

- Core set of HPC-capable tool for running generators

Year 2

- New techniques for increasing physics reach and performance using HPC

Year 3

- Campaigns for tuning physics targeting CMS/ATLAS

NOvA **SBN** **DUNE** **ATLAS** **CMS**

Collaboration

- **HEP and ASCR Collaboration**
  - LHC and neutrino physics: N. Buchanan (CSU, NOvA/DUNE), P. Calafiura (LBNL, LHC-ATLAS), Z. Marshall (LBNL, LHC-ATLAS), S. Mrenna (FNAL, LHC-CMS), A. Norman (FNAL, NOvA/DUNE), A. Sousa (UC, NOvA/DUNE)
  - FASTMath Optimization: S. Leyffer (ANL), J. Mueller (LBNL)
  - RAPIDS Workflow, Data Modeling: T. Peterka (ANL), R. Ross (ANL)
  - Data science: M. Paterno (FNAL), H. Schulz (UC), S. Sehrish (FNAL)
  - J. Kowalkowski – PI (FNAL)
- **Research Associates and Graduate students**
  - Steven Calvez (CSU/PD), Pengfei Ding (FNAL), Matthieu Dorier (ANL/PD), Derek Doyle (CSU/GS), Xiangyang Ju (LBNL/PD), Mohan Krishnamoorthy (ANL/PD), Jacob Todd (UC/PD), Marianette Wospakrik (FNAL/PD), Orçun Yıldız (ANL/PD)