Managed by Fermi Research Alliance, LLC for the U.S. Department of Energy Office of Science

Training and Careers: FNAL Intensity Frontier and Particle-Astrophysics

Rob Kutschke

Joint WLCG/HSF Meeting

Napoli

March 27, 2018

I am Charged to Cover:

- Topics:
  - Introductory training: current practice and needs
  - Advanced training: current practice and needs
  - Career paths
- Communities:
  - Neutrino Experiments
  - Muon Experiments
  - Particle Astrophysics – are there synergies with us?
  - (I only consider communities with significant FNAL involvement)
- All in <10 minutes!

- I have surveyed representative (I hope) people in these communities
- I have a collection of anecdotes (not data)
- I will summarize and form conclusions

Fermilab

About the Neutrino and Muon Communities' Software

- Experiments share a common S/W stack
  - ROOT, Geant4, CLHEP, boost etc
  - *art* – the common framework
    - Exception: MINERvA
  - LArSoft – common tools for LAr TPC experiments
    - See Erica Snider's talk yesterday
  - FIFE – Fermilab's tools for submitting and managing grid and cloud jobs, including data handling, storage and cataloging.
- The teams supporting the stack maintain documentation
  - It is incomplete
    - Priorities are to add the next feature and fix the current bug.
  - Mostly targeted at experts

Pre-requisites / assumptions for training materials

- Students/postdocs will learn computing languages, programming/scripting skills on their own, by taking courses or through their "home groups".
  - Experiments provide references for C++, unix …
    - But no guide to what sections are important!
  - Lab / experiments do not provide this sort of instruction
- Not pre-requisites: understanding the physics of the experiment or understanding how the detector works!

First Contact with An Experiment's Software

- A de facto requirement: In ~1 hour, follow an online cookbook:
  - Look at some event displays
  - Produce histograms that are relevant to that experiment.
  - "I want my summer student to be productive by noon."
  - Experiments have not widely adopted the *art* workbook because it does not meet this requirement.
- Experiments provide a few examples in this style:
  - Incomplete coverage
- Each experiment provides it's own discussion of the S/W stack
  - Blurs the distinction between the experiment's software and the S/W stack
- This defines a starting point from which people can drill down and explore outward.

Intro Training in Neutrino and Muon Communities

- ½ or 1 day tutorials coordinated with collaboration meetings
  - Many experiments do this
- "101" Courses
  - Fewer experiments do this
  - Up to 1 week long but only once per year.
  - Curriculum includes physics, detector, software
  - Coordinated with collaboration a meeting
- Sometimes organized by the collaboration as a whole
  - NOvA: organized by "Young NOvA"
    - Modeled on "Young MINOS"
    - Both regarded as highly successful for many aspects
- People asked to do the online exercises ahead of time.

Moving Forward in the Neutrino and Muon Communities

- How people learn:
  - People learn informally from other students/postdocs
  - SLACK channels are very popular
    - Crowd sourcing has fast turn around; advisors don't lurk on SLACK
  - To start something new, you are given an example of something close and told to perturb around it.
- This is long established pattern
  - Pros:
    - Distributes the workload and fast turn around
  - Cons:
    - Propagates bad practices and incorrect folklore
    - Computing/Software is seen as a collection of isolated tricks, not a coherent whole

Advanced Training in Neutrino and Muon Communities

- Seminar series arranged by various organizations
- *art* Multi-threading forum
- LArSoft annual workshop often focuses on an advanced topic
- Apprenticeship style
  - Student/postdoc is given a job to do; they work with the existing team and are mentored by them.
  - Often it is a very big step from “analysis power user” to being an effective developer
    - Few good candidates to take this step.
- See also page 11 about Machine Learning

Particle Astro Physics Software – Some Background

- Use Dark Energy Survey (DES) as a proxy for the field
  - Largest user of grid cycles from this community
- Data flows off of the detector
  - Detrending pipeline produces catalogs
  - Additional pipelines produce refined catalogs
  - Mostly python with CPU intensive algorithms in C
  - O(a few %) of the DES community works on these pipelines
- Science done using the refined catalogs
  - Most science papers are not signed by the entire collaboration but by a smaller group of researchers.

Training in DES

- Each research group typically has it's own analysis software
  - A variety of languages and libraries
  - Sometimes libraries are considered proprietary
    - Seen as an important asset if the author is in the job market
  - Recently some people have pushed to adopt common software but it has not caught on everywhere
- New people are introduced to the software of the group they joined, by people in the group they joined.
  - As they develop their skills, they contribute to this code base
- No program to train people to contribute to the pipelines
- See also next page about Machine Learning

More About Advanced Training

- FNAL Deep Learning Journal Club
  - Founder and chair: Brian Nord
  - Participants from many experiments; meets once per month
- FNAL Machine Intelligence Group
  - Part of the Scientific Computing line organization
  - Leader: Gabe Perdue
  - Members from: MINERvA, NOvA, DES, DUNE
  - Mandate: be a resource to all HEP communities at FNAL:
    - Develop demos and tutorials; for techniques and resources
    - Workshops
    - Consulting
    - Advertising the availability of these tools

Needs of Experiments

- Experiments see issues:
  - Too many new users experience a difficult startup
  - Too few people advance beyond “analysis power user”.
- They are not sure how to to address these issues:
  - Many individuals have ideas
  - The communities do not have a consensus
  - Some senior people are skeptical that training will pass a cost vs benefit analysis!
    - Spontaneous generation of experts?

Career Paths

- Consensus is that for most postdocs and grad students:
  - At start, almost all want to shoot for a next job in the field
    - Most understand that the odds beyond a post doc are poor
    - Options outside of the field are seen as Plan B
  - Exception: MicroBooNE reports that many grad students view their PhD as a stepping stone to a career in industry, especially in Machine Learning.
  - Experiments do not mentor people for careers outside of HEP:
    - Exception: MicroBooNE has a cadre of former colleagues, now in industry, who mentor people to write industry friendly CVs.
- Most experiments are not sensitized to the issue that individuals who have strong skills in both physics and computing have few career paths to stay in the field.

Conclusions - 1

- There are many training synergies with Particle Astrophysics
  - Material to teach underlying computing skills, including languages, especially python.
  - Training for how to use large scale computing resources
  - Machine Learning
- Areas that do not have a synergy:
  - They do not use C++, ROOT or Geant4
  - They do not have an analog to our frameworks

Conclusions – 2: In the form of Questions

- What is an appropriate distribution of computing skills on a well functioning, mature HEP experiment?
  - What are the basics skills that everyone should have?
  - What are some strawman skill levels and what fraction of the collaboration needs to have that skill level?
- What is an acceptable amount of time for someone to spend on software/computing training, including training on the software stack:
  - In their first week? In their first few months? In their first year?

Stakeholders need to develop consensus on these questions; then we can understand cost vs benefits of various training solutions.

Backup Slides

Neutrino Program – Part 1

- MINERvA: http://minerva.fnal.gov
- LArIAT: http://lariat.fnal.gov
- NOvA: https://www-nova.fnal.gov
- MicroBooNE: http://www-microboone.fnal.gov
- Short Baseline Program
  - Near: SBND: http://sbn-nd.fnal.gov/
  - Mid: MicroBooNE http://www-microboone.fnal.gov
  - Far: ICARUS: http://icarus.lngs.infn.it
    - Will moved from CERN to FNAL later this year

Red text = LAr TPC
= Liquid Argon Time Projection Chamber

Neutrino Program – Part 2

- ProtoDUNE (at CERN)
  - https://dune.bnl.gov/wiki/CERN_Prototype
  - https://web.fnal.gov/collaboration/DUNE/SitePages/ProtoDUNE Working Group.aspx
  - Two LAr detector prototypes
    - Single Phase: https://twiki.cern.ch/twiki/bin/view/CENF/DUNEProtSPH
    - Dual Phase
- DUNE: http://www.dunescience.org
  - Near detector at FNAL: technology TBA
  - Far detector at SURF: may have other technologies too

Red text = LAr TPC

Neutrino Program – Completed Data Taking

- MiniBooNE: https://www-boone.fnal.gov
- MINOS and MINOS+: https://www-numi.fnal.gov/
- ArgoNeut: http://t962.fnal.gov/
- 35 Ton Prototype: https://cdcvs.fnal.gov/redmine/projects/35ton/wiki

Red text = LAr TPC