

Summary - Training and Careers session WLCG/HSF Workshop (Naples)



Conveners

Dario Berzano - ALICE (CERN)

Sudhir Malik - CMS (Puerto Rico)

Dario Menasce - CMS (Milano)

Albert Puig - LHCb (Zurich)



10 Lightning Talks (Tuesday - 1:30- 3:30 parallel)

- ▶ Significant Milestone - first time for people involved in training in these communities under same roof, learning about each others efforts
- ▶ The session had a variety of talks about software and analysis training efforts
 - ▶ Experiment specific (6) - *CMS, ATLAS, ALICE, LHCb, Belle II, Experiments(Neutrino/Muon/Asrto)@FNAL, VIRGO/LIGO*
 - ▶ Non-experiment (2) - *Bertinoro School of Computing, CERN School of Computing*
 - ▶ *Data Science Course at Univ. of Padova/INFN (1)*
 - ▶ *Vision for Future Integrated Software Training (1)*

▶ Training efforts are mature

- ▶ trained hundreds of users - CMS ATLAS in past 8 years
- ▶ some efforts are recent and picking momentum (ALICE/LHCb/ Neutrino@FNAL, VIRGO/LIGO)
- ▶ Organized by dedicated people - official/volunteers

▶ Paradigm

- ▶ Highly distributed environment - computing and collaborators
- ▶ Students trained as physicists but asked to be data analysts
- ▶ Enormous amount of software to learn for newcomers
- ▶ Engage collaborations in meeting pre-requisites for physics analysis

▶ Motivation

- ▶ Give a solid starting point to newcomers in the most used software in their experiment (general and specific)
- ▶ Improve software literacy
- ▶ Teach good practices
- ▶ Help newcomers socialize and integrate in the collaboration

Common Observations (Experiment specific) - II

▶ Material/Content

- ▶ Common basics (Linux/Shell/Python/C++/ Git/ROOT, Machine Learning)
- ▶ Experiment Specific Basic - software for physics analysis, usage of thegrid, data visualization
- ▶ Experiment Specific Advance - data access, trigger selection, physics objects, physics analysis, simulation

▶ Methodology

- ▶ Static self-training material - twikis, vidyo, indico, jupyter notebook
- ▶ Hands-on - tutorials, workshops (1-5 days), schools (5 days)
- ▶ Hands-on training is CERN centric mostly except CMS
- ▶ User feedback - google forms, chat forums

▶ Challenges

- ▶ Facilitators (teachers/tutors) voluntary basis and a big number, does not scale very well mostly, room booking issue of CERN centric
- ▶ LHCb/ALICE issue with training off-CERN community, tutorials ties to teachers (what if they leave)
- ▶ Facilitators need not see this as charity work, needs recognition within experiment, help propel their career
- ▶ Funding for training activity is an issue, no credits for the work

▶ Commonality and Collaboration in Training

- ▶ LHCb and ALICE joined efforts for mutualize organization and lesson writing
very positive collaboration, allowed to create a common repository for basic course
- ▶ Common StarterKit blessed by both experiments
- ▶ This collaboration happened by chance by good hearted and easy to work with people from both sides

- ▶ There is desire for common training efforts by current Fermilab based experiments
 - ▶ Neutrinos, Muons, Astro,DES
 - ▶ Share common S/W stack - ROOT, Geant4, CLHEP, art (SW Framework), grid/cloud tools
 - ▶ Currently have Online Software Exercises and 1-2 day workshops aligned with collaboration meetings



CERN School of Computing Schools



- ▶ Create a common culture in scientific computing among young scientists and engineers involved in particle physics or other sciences, human networking
- ▶ Started in 1970
- ▶ Visited 21 countries
- ▶ 2600 students of ~80 different nationalities attended CSC usually 60-80 per year, postgraduate engineers and scientists, often PhD students
- ▶ Organized by CERN together with the hosting University
- ▶ Summer University style (not a conference or training session), focus on persistent knowledge, less on know-how, exam/diploma given
- ▶ Main School - 2 weeks, 55 hours of hands-on and lectures
 - ▶ Base Technologies (architecture, SW design, security, networking)
 - ▶ Physics Computing (intro, tools+techniques, analysis, MV classif.)
 - ▶ Data Technologies (storage, data workflows, visualization)
- ▶ Thematic School - 1 weeks, more advance, focussed
 - ▶ Technologies+platforms, Parallel and optimized scientific software development, Effective I/O for scientific applications
- ▶ Inverted School - 2-4 days as proposed by lecturers, alumni



Bertinoro School of Computing



- ▶ Goal - Introduce (young) researchers to the world of modern computer architectures and heterogeneous computing - how to deal with them in terms of software code and tools, the audience includes also non-particle physicists
- ▶ Content - Modern C++, Memory-friendly programming, Floating-point computation and vectorization, Parallel programming (multi-core, GP-GPU, multi-node) and related tools
- ▶ 5-full day program - lectures + hands-on, student lightning presentations, feedback
- ▶ Challenges - keeping up with technology advances and user needs, have a unique thread through all the lectures, no repetitions, C vs C++, ..., find the right balance between theory and practice, don't overload students with information, adapt to the audience

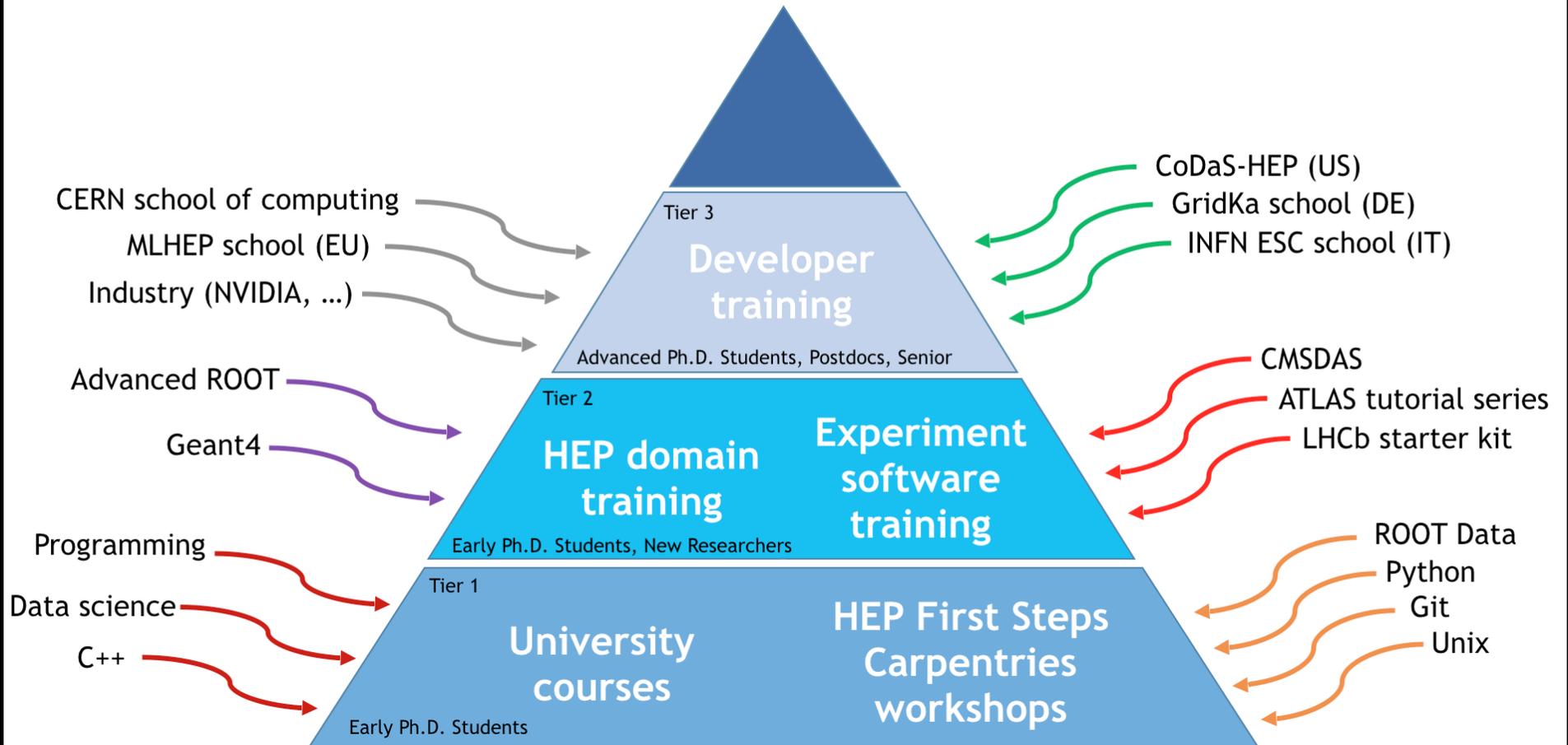


Career Guidance

- ▶ CMS established career committee in 2012 to connect prospective job seekers in non-academia with its alumni in industry
- ▶ Career events now organized jointly with ALICE, ATLAS, CMS and LHCb
- ▶ CMS maintains a job twiki to advertise jobs by CMS alumni regarding positions in their companies
- ▶ MicroBooNE Experiment has a cadre of former colleagues, now in industry, who mentor people to write industry friendly CVs
- ▶ MicroBooNE Experiment has a cadre of former colleagues
- ▶ Data Science course (Master Level - 2-year) established at Padova University is a good effort to match job market and consolidate HEP research with career, push basic software trying to university

Proposed Vision of Training Chain

HEP Software Training



- ▶ Training efforts are in full momentum well appreciated in all running LHC experiments
- ▶ Realization in the HEP community that software training is important to achieve goals of experiments and provide alternate (non-academic) careers to the users
- ▶ Common training can happen before entering HEP field and some common software skills can be ported back to university curriculum (e.g. Data Science Master Level degree)
- ▶ Synergies already begin to exist between experiments (LHCb/ALICE)

Next Steps

- ▶ A good start definitely but need to meet/communicate regularly to build synergy further, design common software training and entry point (collective links to documentation) for sustainability
- ▶ Some funding agencies (NSF-US) are supportive of software training, expand this support from other funding sources
- ▶ Individual Experiments need to mandate and invest in training efforts - money, common responsibility, manpower, credits, career path to personal involved
- ▶ HSF can be a key hub for future training path



THANKS !!



- ▶ Thanks to all the speakers and attendees for participation and input to this session
- ▶ Thanks to the meeting organizers for giving a platform to training efforts and great hospitality and care