Summary of the Computing Frontier Discussions

LATBauerdick/Fermilab
Steven Gottlieb/Indiana U.
Working Group Conveners and Participants

Community Planning Meeting
Fermilab October 13, 2012

Organization I

- ✦ We have subgroups for User Needs and for Infrastructure.
- ✦ User Needs subgroups:
  - CpF E1: Cosmic Frontier
  - CpF E2: Energy Frontier
  - CpF E3: Intensity Frontier
  - CpF T1: Accelerator Science
  - CpF T2: Astrophysics and Cosmology
  - CpF T3: Lattice Field Theory
  - CpF T4: Perturbative QCD
- ✦ Each subgroup with interact with the other frontiers to assess the computing needs to advance the science.

Organization II

- Infrastructure subgroups:
  - CpF I1: Computing, including special purpose hardware
  - CpF I2: Distributed Computing and Facility Infrastructures
  - CpF I3: Networking
  - CpF I4: Software Development, Personnel, Training
  - CpF I5: Data Management and Storage
- The infrastructure groups are supposed to project computing capabilities into the future and see how the user needs map onto the trends.
- If the trends indicate that research is needed to meet some computing needs, we will point that out to the funding agencies.

Computing Frontier

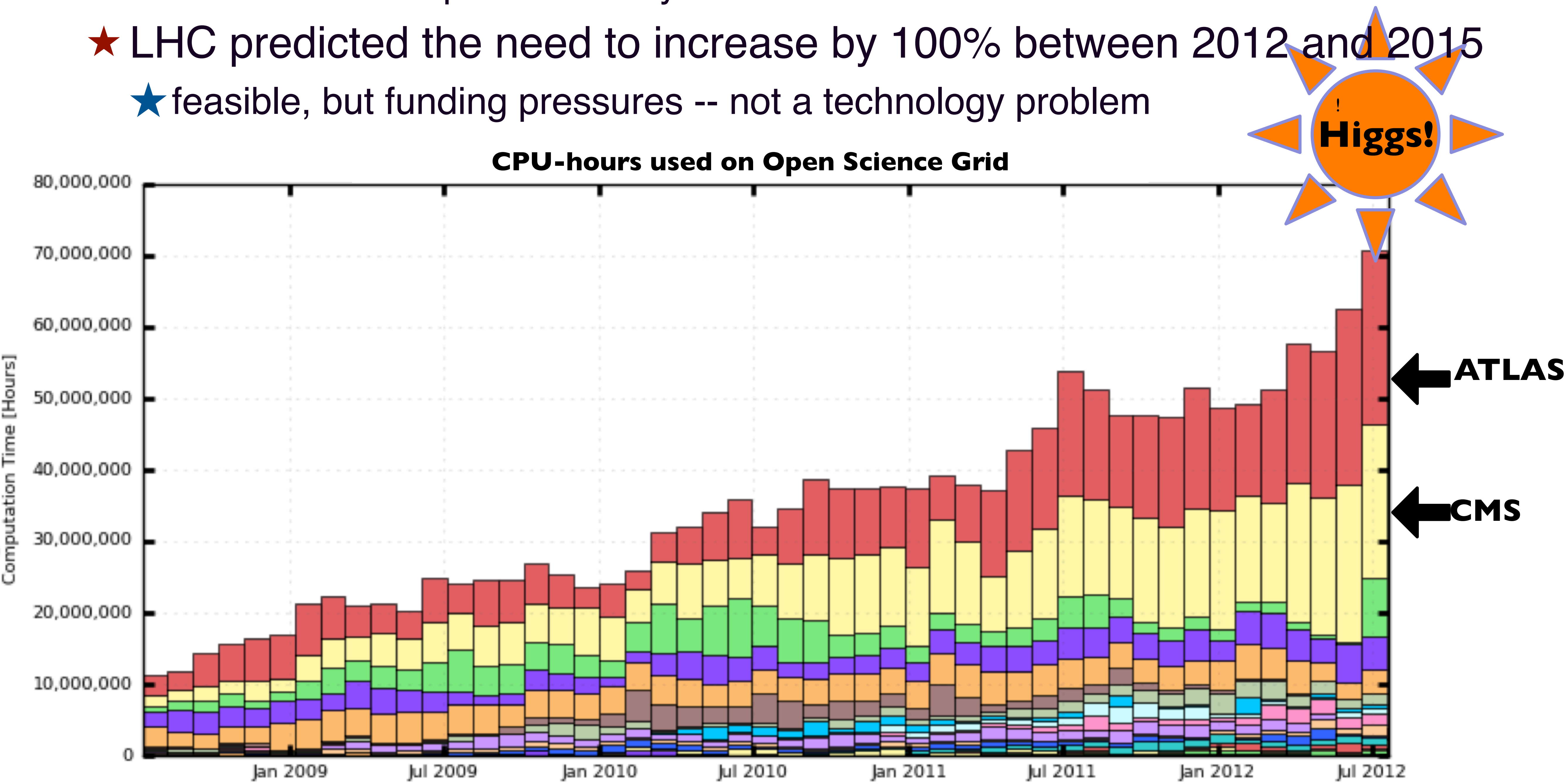
- **User Needs** groups look at computing as tool to enable science
  - what computing/software is needed to get the science done
- **Infrastructure** groups will project computing capabilities into the future and see how the user needs map onto the trends
  - identify where research is needed to meet computing and software needs
    - e.g. in Open Science Grid we documented a number of areas where computing research is needed to sustain the needs of communities for this decade
  - Identify advances and investments needed in computing infrastructure
  - Identify where computing capabilities open physics opportunities
- Approaching computing and software problems in a creative and innovative way could give leadership advantages to US groups
  - quite possibly US leadership position in LHC science is helped and enabled through computing and support for collaborative research
- In any case, innovation and continuous engineering needed to keep up w/ the increasing demands and expectations from the physics program

We start to answer some of the questions :-)

- ★ Is the current LHC computing model robust into the future? How far?
  - ✦ probably yes, by doing the required engineering, the limit is cost, not scalability
- ★ Are other compute models needed to support the other Frontiers, or can 'one size fit all'?
  - ✦ there is a lot of interest to collaborate closely, share ideas, even resources!
- ★ Can we imagine transformative breakthroughs? Yes, there are examples!
  - ✦ e.g. analysis 10X faster? — actually, LHC achieved factor ~10 in reco performance
  - ✦ e.g. all data access over "intelligent" networks, globally accessible datasets!
  - ✦ e.g. location-independence of computational resource so we're ready for cloud computing prices to go down (but need a factor ~10 for this to be economical!)
- ★ Are there new architectures that need to be developed to help with the DAQ/triggering etc?
  - ✦ yes, large interest across the community in developing skills and techniques to use highly parallel systems, GPUs etc, both in trigger/DAQ and "offline"
  - ✦ need to make sure that we get close liaison b/w instrumentation and computing
- ★ Interest in new hardware across areas, bring together know-how
  - ✦ also interest in looking at the "leadership computing" facilities to play a role

Energy Frontier Computing

- ✦ At the LHC computing increase ~< Moore's law
  - ✦ on Open Science Grid, factor 5 increase in CPU-hours over 4 years
    - ★ non-LHC use expanded ~in synch with LHC use
  - ★ LHC predicted the need to increase by 100% between 2012 and 2015
    - ★ feasible, but funding pressures -- not a technology problem



Energy Frontier Computing limited by cost, not technology

- at 2021 within Moore's law could sustain LHC rates of ~10kHz, even if reconstruction CPU goes up x3 and event size x3
  - someplace in the neighborhood of 10-20k Hz sustainable rate and bandwidth of O(50GB/sec) should be sustainable, question of cost
  - so, it is an option to take a higher rate and perform more of the complete chain of processing and analysis and writing of synthesized output
  - can we afford it? to be looked at across detector/daq and computing cost
- While we don't expect to hit scalability limits, scaling remains a hard problem in the real world
  - The multi-core problem is being solved
    - re-engineering workflow management, framework parallelization etc
  - network, CPU and storage are improving in synch
  - we are learning how to use "computing as a service" (like clouds)
    - it will become a matter of cost and "business model"
  - data management remains a hard problem, but storage management issues tend to become easier as connectivity between sites improves

Intensity Frontier Computing

- Siegrist's question of "does one size fit all"?
  - So far have benefitted enormously from advances in computing at energy frontier
- but, community more fragmented, also there are experiments that are mostly University based and do not have a large lab participation
  - needs to be addressed, and organizations like OpenScienceGrid can help
  - Cosmic Frontier has a similar issue
- computing will need to get a lot of information from the experiments and even to get some idea of the scope of the experiments and efforts that we need to include in order to really understand how to proceed.
- Some of this is easy - the Fermilab IF experiments running and approved - but others are tougher and require that a wide search be made with the help of the IF conveners.
  - that will be a task that we will take on soon to get things off the ground
- Have a list of questions to be addressed and answered
  - see backup slides and Twiki

Cosmic Frontier Computing

- Computing plays three major roles at the Cosmic Frontier:
  - Pipelines
    - Real-time computing (dedicated/specialized h/w), capacity computing (clusters)
    - specialized hardware and software, specificity to particular projects
  - Computational cosmology
    - The theory and modeling aspect of the Cosmic Frontier (analog of lattice QCD), poses major requirements for large-scale capability computing
  - Analysis computing
    - Scientific inference from observations as well as data-drive discovery, in effect a “telescope for data”
      - issues include short and long-term data archiving/analysis/curation for both simulations and observations
    - Growing need for nontraditional data-intensive scalable computing; hybrid of capacity/capability systems to deal with unusual access patterns and different usage models
    - Serves as a bridge for science teams to do/connect with big science (bringing the “HEP way” to survey science)

Cosmic Frontier Issues to Address

- Hardware Requirements
  - exploit heterogeneous/next-generation architectures
  - plan response to the many PB needed for sims/experiments, different data life-cycles, access patterns
    - fast sequential, parallel, localized index random access, needs for very large contiguous memory systems, and interactive viz/data views
- Infrastructure/SoftwareFacilities
  - S/W, algorithms: key for simulations, do same for analysis computing
  - Develop national and local facility partnerships to combine large-scale facilities with agility and diversity (build a tiered system)
  - R&D timescale is slow compared to e-industry, needs to be sped up
- Personnel/Training
  - Recruiting and training young talent crucial, need alternative pathways in academia, make R&D cycles shorter and more intellectually attractive

Cosmic Frontier Computing Plan (CpF E1 and T2)

- Work with Cosmic Frontier team to provide a statement of needs for future computational physics Infrastructure/Software Facilities

- Will work with other computational frontier areas about pipeline computing partnerships (HEP meets astro detectors)

- will need more input from the community, plan to piggyback on workshops

- Many possibilities for cross-specialization interactions (e.g. with Energy Frontier)

CpF T1: Accelerator Science

- started gathering requirements from the machine physicists in energy and intensity frontier at this meeting
  - confident that the computational and computing technologies aspect can be covered from the members of the subgroup (and utilizing experience and expertise from our SciDAC and other computational collaboration)
  - explored commonalities in requirements for the different frontiers and different technologies
    - more detailed account of Accelerator Science Modeling needs in backup slides
- Next steps, we will discuss these requirements among the members of the group, and make the connections between the physics driver and need for algorithmic / capability development
  - have established a mailing list to facilitate these discussions. When appropriate, we will involve our machine physicist partners.
- are aiming to have a mini-workshop (either virtual or face to face) midterm through the process to bring things to focus and begin writing.

CpF T3 Lattice Field Theory

- Non-perturbative lattice field theory calculations are needed to maximize scientific output of the current and planned experimental HEP program
  - The Lattice-QCD community has been working at the scientific computing frontier in terms of algorithms and massively parallel architectures which has enabled reliable computations with controlled systematic errors.
  - US Lattice QCD community has an established program of weak-matrix element calculations for quark-flavor physics, is developing in other areas such as calculations of BSM theories relevant for high-pT collider experiments or of the hadronic light-by-light contribution to muon g-2.
- most important charge is to map out long-term physics goals.
  - For example: What non-perturbative calculations will be needed to interpret measurements made at future frontier facilities in the next 10 years? To what precision do we need these calculations to make them relevant? Because of the long time frame please consider difficult quantities that we don't know how to do right now, e.g. SM matrix elements with large long-distance contributions or multi-hadron final states.
- welcome input from/interaction w/ experimentalists, phenomenologists!

CpF T4: Perturbative QCD

- ★ subgroup had one-hour meeting with about 15-20 people
- ✦ short term goals
  - ★ provide collider experiments with state-of-the-art theoretical predictions, make this process as automated/fast/efficient as possible
  - ★ work in close contact with the QCD, Top, EW, Higgs groups of EF
- ✦ longer term: facing serious computing issues
  - ★ benefit from working with computing community, taking advantage of large-scale computing facilities and already existing computer-science knowledge an technologies (ex: use of GPU, networks, etc.).
  - ★ This may imply restructuring the way we build codes for calculations and simulations, and will therefore take some longer term planning, since we need to convince and educate the pQCD community
  - ★ proofs of concept to demonstrate efficiency and effectiveness of using new approaches, for example to show how parallelization of existing codes can really help improving their speed, and how it can be achieved with relatively minor changes of existing codes

"Infrastructure" Issues

- ★ HEP has large experience in both Distributed High-Throughput computing (experiment program) and High-Performance computing (mostly theory/simulation/modeling) — we are good at "collecting together" compute resources from whereever we can get them
- ★ emerging network capabilities and data access technologies improve our ability to use resources independent of location, over the network
  - ✦ this enables a large spectrum of provisioning resources: dedicate facilities, universities, opportunistic use, commercial clouds, leadership-class machines etc
- ★ with the need for more parallelization the complexity of systems continues to increase: frameworks, workload management, physics code,
- ★ we are developing the list of issues to be addressed
  - ✦ also including maintenance of software libraries, data preservation

✦ plans to address those "across the board" technology issues

- ★ e.g. "piggy-back" on planned networking workshop in December
- ★ ideas for a "advanced hardware architecture" workshop end of November
- ★ expect to run others e.g. on data management, etc

Training, Workforce development

- Two axis of training to explore:
  - A) Ongoing availability skilled developers (needed to address the long-life of the experiments in tandem with the increasing pace of technology changes) need to include computational training relevant to our field as part of the core, perhaps multi-disciplinary, course and practicum work of all colleges and universities.
  - B) Include training of the users and workforce as an integral part of an experiment/project from the start to the end.

Computing at a Good Starting Point for Moving into the HEP Future

- we have established and well-working computing models
  - the different frontiers are at some level separate in terms of facilities
    - but we are identifying many commonalities in terms problems and approaches
  - by coming together we will map out a way to go forward
- start using new technologies and approaches that are transformative
  - for sure we'll see things over the coming years we have not yet thought of
- industry caught up to us and in cases surpassed us
  - might reassure us that with hard work computing won't be a road block

- a roadmap how to best support science from the view of computing
  - assess the individual needs of different frontiers
  - identify commonalities and differences
    - where it makes sense to share facilities or grids/clouds/etc. and where dedicated installations are needed, and how all can be used efficiently together
  - look at how industry can help and when this becomes feasible

First Steps on the Road to "Snowmass on the Mississippi"

- ✦ We've only started — it was a vibrant meeting and a promising start
- ✦ Thanks to conveners and participants!
- ✦ We will need extensive input and attention from the community to help us identifying the user needs
- ✦ Please sign up on Computing Frontier mailing list
  - ★ send email to listserv@fnal.gov with contents SUBSCRIBE CPF Firstname Lastname
- ✦ Let's keep momentum to move forward

Backup Slides,
More Detailed Information

Information Needed from IF

- See Intensity Frontier Computing page
  - http://www.snowmass2013.org/tiki-index.php?page=CpF%20E3%20Intensity%20Frontier&pagenum=2
  - What are the CPU needs for simulation, reconstruction and analysis?
  - What is expected in terms of disk usage and archival storage?
  - Are there special networking needs?
  - What kind of effort (ie people) is needed to support the experiments?
  - What tools are needed to use and integrate all of the above?
  - How do the experiments expect to make use of computing resources, along with expected timescales?
  - How are common HEP tools used, such as G4, GENIE, ROOT, frameworks and any other software tools?
  - What Grid/Cloud resources are needed (which includes many of the above)?
  - What is currently limiting these experiments?
  - What is the expected balance between lab vs university computing, offsite computing availability, data transfer model, etc?

Accelerator Science Modeling Needs

- ★ In broad terms, the driver for intensity frontier accelerators is beam loss characterization and control (even better, minimization); for energy frontier (short term, i.e LHC upgrades) is beam stability; for energy frontier (long term, i.e muon colliders, plasma driven, etc) is the ability to produce end-to-end designs
- ★ All frontiers require integrated modeling. Currently, for high-fidelity modeling, because of the many degrees of freedom involved in the problem, we run “single physics” or “few physics” model simulations. This will require better algorithms with the ability to utilize massive computing resource (tightly coupled)
- ★ Intensity frontier machines would like “control room feedback” capabilities (because of the loss implications). Would it be possible with utilization of new computing technologies to deliver such fast turnaround? Such capability is also relevant to R&D experimental efforts (for example plasma driven acceleration), but there the requirements are even more stringent because, in addition to faster modeling, extracting useful information out of a simulation run involves analyzing massive data. So this leads to analysis tools development requirements that must perform in almost real time (as the experiment is running).
- ★ For intensity frontier and for experimental programs develop tools and techniques to allow for direct comparison between beam diagnostic detectors (what machine physicists see) and simulated quantities (what computational physicists use for input to the simulation and to describe the output of the models).
- ★ Energy Frontier (long term) accelerators have a lot of specific new physics model capability needs (with some overlap, for example radiation and scattering, which is relevant to muon collider, plasma and gamma-gamma options). We detailed lists for each case.
- ★ Energy frontier machines require better numerical models (less numerical noise). This has direct impact on the choice of numerical techniques for different physics implementations.

More thoughts on Training/Workforce

- Better training of the entering workforce both scientists and (software and hardware) engineers is essential. This is driven by
  - the increased specialization of physics codes
  - transitions in developers and maintainers of the code (long life of code, short availability of any one individual)
  - aims for reuse and broader adoption of the codes (see agency talks)
  - agile changes in the face of the increasing pace of new technologies (hardware and software)
- Embedding of computational thinking, theory and practice in physics college curricula is one aid. Some ideas are:
  - A recognized program in DPF to encourage the trend towards multi-disciplinary degrees and tracks in college faculty that cross physics and computing.
  - Partnering with computer engineering and computer science faculties to have courses directly relevant to the practices and tools of physics e.g. Object oriented design course examples from physics; ensuring that C++ courses are maintained and not replaced.
- Encouraging all physics faculty to accommodate (and require?) specific computing classes for credit as prerequisites for particular physics courses
- Expanding the internship (REU, GPCF) opportunities between colleges and physics collaborations and institutions.