



Computing at the Intensity Frontier

Mayly Sanchez (Iowa State/ANL)
Brian Rebel (Fermilab)
Stephen Wolbers (Fermilab)

M. Sanchez - mayly.sanchez@iastate.edu
B. Rebel - brebel@fnal.gov
S. Wolbers - wolbers@fnal.gov

Snowmass on the Mississippi
July 29, 2013

The Motivation Line

No computing
=
No physics

The Overview

- **Computing at the Intensity Frontier has many significant challenges.**
- While not as data intensive as LHC experiments, **there are significant computing requirements** for theory and modeling, beam line and experiment design, triggers and online/DAQ, event reconstruction and processing, and physics analysis.
- Intensity Frontier encompasses: quark flavor physics, charged lepton processes, neutrinos, baryon number violation, new light weakly coupled particles, nucleons nuclei and atoms.
 - The requirements and resources of quark flavor physics (as in Belle II and LHCb) are more similar to those of the energy frontier.
 - The requirements and resources of new light weakly coupled particles and nucleon, nuclei and atoms sectors are more similar to those of the cosmic frontier.
- In this exercise we **targeted the computing model** for: **charged lepton processes, neutrinos and baryon number violation.**

The Opportunities

- **Computing resources are required to maximize the impact of the physics extracted from IF experiments.**
- **IF has become the central focus of the US-based particle physics program.**
- The change of focus of the US program, coincides with the transition of Fermilab from EF to IF. **Many experiments in IF are thus Fermilab based:**
 - experiments to measure neutrino cross sections (MiniBooNE, MicroBooNE, MINERvA),
 - experiments to measure neutrino oscillations over long baselines (MINOS+, NOvA, LBNE),
 - experiments to measure neutrino oscillations over short baselines (MiniBooNE, MicroBooNE),
 - experiments to measure muon properties (g-2, mu2e),
 - as well as various future experiments (ORKA, nuSTORM, SEAQUEST).

The Opportunities

- There is also **strong US participation in several international IF experiments**: Super-Kamiokande (SK), T2K, Daya Bay, SNO/SNO+ as well US university lead experiments such as IceCube. There is significant detector/experiment design R&D as well.
- **The impact of the US contribution to the physics results of these experiments is strongly correlated to the availability of computing resources and the efficiency of the computing model adopted.**
- The groups participating in any of these experiments go from 30 to 400+ people.

The Challenges





HEP Intensity Frontier Experiments

Outside US

Taking data

List from DOE:

There are MANY

Experiment	Location	Status	Description	#US Univ.	#US participation
Belle II *	KEK, Tsukuba, Japan	Physics run 2016	Heavy flavor physics, CP asymmetries, new matter states	10 Univ, 1 Lab	55
BES III *	IHEP, Beijing, China	Running	Precision measurements charm, charmonium, tau; search for and study new states of hadronic matter	6 Univ	26
CAPTAIN *	Los Alamos, NM, USA	R&D; Test run 2015	Cryogenic apparatus for precision tests of argon interactions with neutrinos	5 Univ, 1 Lab	20
Daya Bay	Dapeng Peninsula, China	Running	Precise determination of θ_{13}	13 Univ, 2 Lab	76
Heavy Photon Search *	Jefferson Lab, Newport News, VA, USA	Physics run 2015	Search for massive vector gauge bosons which may be evidence of dark matter or explain g-2 anomaly	8 Univ, 2 Lab	47
KOTO *	J-PARC, Tokai, Japan	Running	Discover and measure $K_L \rightarrow \pi^0 \nu \bar{\nu}$ to search for CP violation	3 Univ	12
LArIAT	Fermilab, Batavia, IL	R&D; Phase I 2013	LArTPC in a testbeam; develop particle ID & reconstruction	11 Univ, 3 Lab	38
LBNE	Fermilab, Batavia, IL & Homestake Mine, SD, USA	CD1 Dec 2012; First data 2023	Discover and characterize CP violation in the neutrino sector; comprehensive program to measure neutrino oscillations	48 Univ, 6 Lab	336
MicroBooNE	Fermilab, Batavia, IL, USA	Physics run 2014	Address MiniBooNE low energy excess; measure neutrino cross sections in LArTPC	15 Univ, 2 Lab	101
MINERvA	Fermilab, Batavia, IL, USA	Med. Energy Run 2013	Precise measurements of neutrino-nuclear effects and cross sections at 2-20 GeV	13 Univ, 1 Lab	48
MINOS+	Fermilab, Batavia, IL & Soudain Mine, MN, USA	NuMI start-up 2013	Search for sterile neutrinos, non-standard interactions and exotic phenomena	15 Univ, 3 Lab	53
Mu2e	Fermilab, Batavia, IL, USA	First data 2019	Charged lepton flavor violation search for $\mu N \rightarrow e N$	15 Univ, 4 Lab	106
Muon g-2	Fermilab, Batavia, IL, USA	First data 2016	Definitively measure muon anomalous magnetic moment	13 Univ, 3 Lab, 1 SBIR	75
NOvA	Fermilab, Batavia, IL & Ash River, MN, USA	Physics run 2014	Measure ν_μ - ν_e and ν_μ - ν_μ oscillations; resolve the neutrino mass hierarchy; first information about value of δ_{cp} (with T2K)	18 Univ, 2 Lab	114
ORKA	Fermilab, Batavia, IL, USA	R&D; CD0 2017+	Precision measurement of $K^+ \rightarrow \pi^+ \nu \bar{\nu}$ to search for new physics	6 Univ, 2 Lab	26
Super-K	Mozumi Mine, Gifu, Japan	Running	Long-baseline neutrino oscillation with T2K, nucleon decay, supernova neutrinos, atmospheric neutrinos	7 Univ	29
T2K	J-PARC, Tokai & Mozumi Mine, Gifu, Japan	Running; Linac upgrade 2014	Measure ν_μ - ν_e and ν_μ - ν_μ oscillations; resolve the neutrino mass hierarchy; first information about value of δ_{cp} (with NOvA)	10 Univ	70
US-NA61 *	CERN, Geneva, Switzerland	Target runs 2014-15	Measure hadron production cross sections crucial for neutrino beam flux estimations needed for NOvA, LBNE	4 Univ, 1 Lab	15
US Short-Baseline Reactor	Site(s) TBD	R&D; First data 2016	Short-baseline sterile neutrino oscillation search	6 Univ, 5 Lab	28

More on the Challenges

- Taking two more lines from IF plenary:
 - Intensity Frontier itself is a micro-cosmos of all of HEP
 - Interconnected, but somewhat (or entirely) independent experimental efforts within Intensity Frontier
- Thus we have:
- **Many experiments** which can potentially lead to **fragmentation of efforts**, reinventing the wheel, lack of access to computing advances, more dollars per megabyte or cpu.
- **Broad range of experiments** which leads to **broad range of needs**. From experiments with 10s to 100s of people, from high intensity realtime processing but small data sets to large data sets which in the sum might be getting close to collider experiments of yore.

The Method



- We wanted to **qualitative survey the community of current and future experiments in the IF** in order to understand the computing needs but also the foreseen evolution of said needs.
- Computing liaisons and representatives for the LBNE, MicroBooNE, MINERvA, MINOS+, mu2e, g-2, NOvA, Daya Bay, IceCube, SNO+, SK, T2K, SEAQUEST collaborations all responded to the survey and provided input.
- This does not cover all experiments in all areas but we consider it a representative survey of the field.
- More input is of course welcome. Please **see/email/chat Brian Rebel and myself** over these days or the next few weeks.

We want to thank the people that took the time to give well thought answers this survey.

The Categories

- Responses to the survey can be summarized in five aspects:
 - support for software packages,
 - support for software frameworks,
 - access to dedicated and shared resources,
 - access to data handling and storage, and
 - overall and evolution of computing model.

Survey still live: [click here](#). Please contribute and let us know.
There is also space for input that does not fill neatly into these.

Software Frameworks

- **Broad range of experiments** inevitably leads to broad range of frameworks.
- The Fermilab-based IF experiments (from g-2, NOvA to LAr experiments including MicroBooNE and LBNE) have converged on **ART as a framework** for job control, I/O operations, and tracking of data provenance.
 - Developed and maintained by the Scientific Computing Division at Fermilab by computing professionals. It has perhaps the largest user base within IF at this time.
 - Increased resources for this framework could enable some of the needs experiments such as more accessible parallelization of experiment's code, for example using standard thread libraries (OpenMP, TBB).
- Experiments outside of Fermilab (or before ART) use **LHC derived frameworks** such as Gaudi **or homegrown frameworks** like MINOS(+), IceTray and RAT.
 - The level of support for development and maintenance of such frameworks varies depending if the experiment is a significant stakeholder and/or significant human resources are available.

Software Packages

- **ROOT** and **Geant4** are the bread and butter of all HEP experiments. They are critical to all experiments in IF. Support for these packages is essential.
- Geant4 has traditionally focused on EF experimental support. More ties/stronger support to IF experiments is a requirement.
- As an example, Geant4 is barely suitable for large scintillation detectors; given a complex geometry and large number of photons to track.
- Community desires improved efficiency for both of these packages. For example better ROOT I/O and Geant multi-threading (latter in beta?).
- Neutrino experiments use **specialized packages for neutrino interactions: GENIE and Neut**. GENIE is a public package that would benefit from continued support as it is heavily used in US experiments.
- Comprehensive simulation of neutrino-nucleus interactions with state-of-the-art generators (GENIE) has been added GEANT4.

Software Packages

- LArSoft is a **common simulation, reconstruction and analysis toolkit** for use by experiments using liquid argon time projection chambers (LArTPCs) that is managed by Fermilab. All US experiments using LArTPCs currently use LArSoft. Similarly the LAr and NOvA experiments share a simulation toolkit.
 - **Joint efforts** where possible **make better use of development and maintenance resources.**
- There are a number of other **specialized physics packages** in use by the community, for example: FLUKA for beam line simulations, CRY for simulating cosmic ray particles, NEST for determining ionization and light production in noble liquid detectors GLOBES for experiment design.
 - By no means a comprehensive list can be discussed here. More complete responses to the survey will be made available.
- Less specialized packages but **equally important are those relating to infrastructure access for code management, data management, grid access, electronic log books and document management.** Experiments differ much more on these than on the specialized ones (unless lab centralized).

Access to Dedicated Resources

- **Hardware demands of IF experiments and IF R&D** modest compared to those of EF experiments. However the needs are **NOT insignificant**.
 - As an example, each Fermilab based IF experiment will require 1000s of dedicated batch slots. NOvA will need ~5M CPU hours to generate just simulation files. LBNE expects to use PB of storage, even smaller experiments like MicroBoone and MINERvA expect to use close to a PB.
- **Efficient use of available grid resources has had/could have a huge impact on IF experiments and IF R&D.**
 - As an example experiments like T2K run intensively on grid resources in Europe and Canada. The US groups use a combination of local university and international resources (albeit with lower priority on the latter).
- **Dedicated storage resources are needed for internationally or university run IF experiments as well as IF R&D.**
 - Data storage in T2K will also get to PB scale is distributed in Japan, Canada (Triumf), UK (RAL). SNO+ similarly uses Canada and UK grid storage. Slow link to the US for local analysis of data/simulations.
- **Access to grid resources is all important and high on every experiment's list.**

Access to Shared Resources

- **Fermilab provides access to the grid to all experiments on site** and it is considered a fundamental part of the computing model.
 - Issues are mostly in efficient data handling and script optimization. Resources for computing professionals is provided through Fermilab and would be extremely useful if increased.
 - On site grid access is however not sufficient, offline Monte Carlo generation is common among experiments. Use of resources could be improved by suggestions below.
- **US participation** in international IF efforts **uses a combination grid resources based mainly outside of the US** and smaller local clusters.
 - A common theme in this category described how US institutions are at a great disadvantage due to the lack of grid resources, and unable to participate in primary data processing.
 - Even though they are recognized by the Open Science Grid, US groups have no US-based grid computing resources.
 - Canadian and UK grid support was cited several times as a model both for grid computing and grid storage.
- **It was widely noted that the lack of dedicated US resources has a detrimental impact on the science.**

Data Handling and Storage

- **The use of Fermigrid and Open-Science Grid are essential to all experiments responding to the survey.**
- **Fermilab-based experiments indicated that all data is stored on site.** The infrastructure there handles active storage as well as archiving of data. SAM was noted as the preferred data distribution system for these experiments. Heavy I/O for analysis of large numbers of smaller sized events is an issue for systems like BlueArc.
- **Professional support is required for methods to seamlessly use Fermilab and non-Fermilab resources through job submission protocols.**
 - For Fermilab-based experiments university and other national lab resources are used in the production of Monte Carlo files. A common protocol to access these resources such as OSG is in the current plans.
 - All international efforts would benefit from an ATLAS-like model where institutions can set up official, verified mirrors of data and simulations. For these experiments, only UK and Canada grid sites are available to store data.

Computing Model

- We found a **high degree of commonality among** the various experiments' **computing models** despite large differences in type of data analyzed, the scale of processing, or the specific workflows followed.
- The model is summarized as a traditional **event driven analysis and Monte Carlo simulation using centralized data storage** that are **distributed to independent analysis jobs running in parallel on grid computing clusters**. Peak usage can be 10x than planned usage.
- **For large computing facilities** such as Fermilab, it is useful to **design a set of scalable solutions** corresponding to these patterns, with associated toolkits that would allow access and monitoring. Provisioning an experiment or changing a computing model would then correspond to adjusting the scales in the appropriate processing units.
- **Computing should be made transparent to the user**, such that non-experts can perform any reasonable portion of the data handling and simulation. Moreover, all experiments would like to see **computing become more distributed across sites**. Users without a home lab or large institution require **equal access to dedicated resources**.

Evolution of Computing Model

- **The evolution of the computing model** follows several lines including taking **advantage of new computing paradigms**, like storage clouds, different cache schemes, GPU and multicore processing.
- As regards computing technology, there is a concern that as the number of cores in CPUs increases, RAM capacity and memory bandwidth will not keep pace, causing **the single-threaded batch processing model to be progressively less efficient on future systems** unless special care is taken to design clusters with this use case in mind.
- **There is no current significant use of multi-threading**, since the main bottlenecks are GEANT4 (single-threaded) and file I/O. However **there is interest in real parallelization** at the level of ART for example.
- **Greater availability of multi-core/GPU hardware in grid nodes would provide motivation for upgrading code to use it.** For example currently we can only run GPU-accelerated code on local, custom-built systems. A proposed example for GPU use included “repeated frequent tasks like quick down-going cosmics identification for pre-reconstruction filtering”.

Computing Technology for the Intensity Frontier

- There are many efforts at each of the frontiers driving technology to improve computing.
- We choose to highlight here three efforts that have interesting commonalities among the three frontiers and that have the potential to have a high impact:
 - Intensity Frontier effort - Stan Seibert (UPenn)
 - Energy Frontier effort - Tom LeCompte (ANL)
 - Cosmic Frontier effort - Amanda Weinstein (Iowa State)
- But first let's talk about a problem across frontiers...

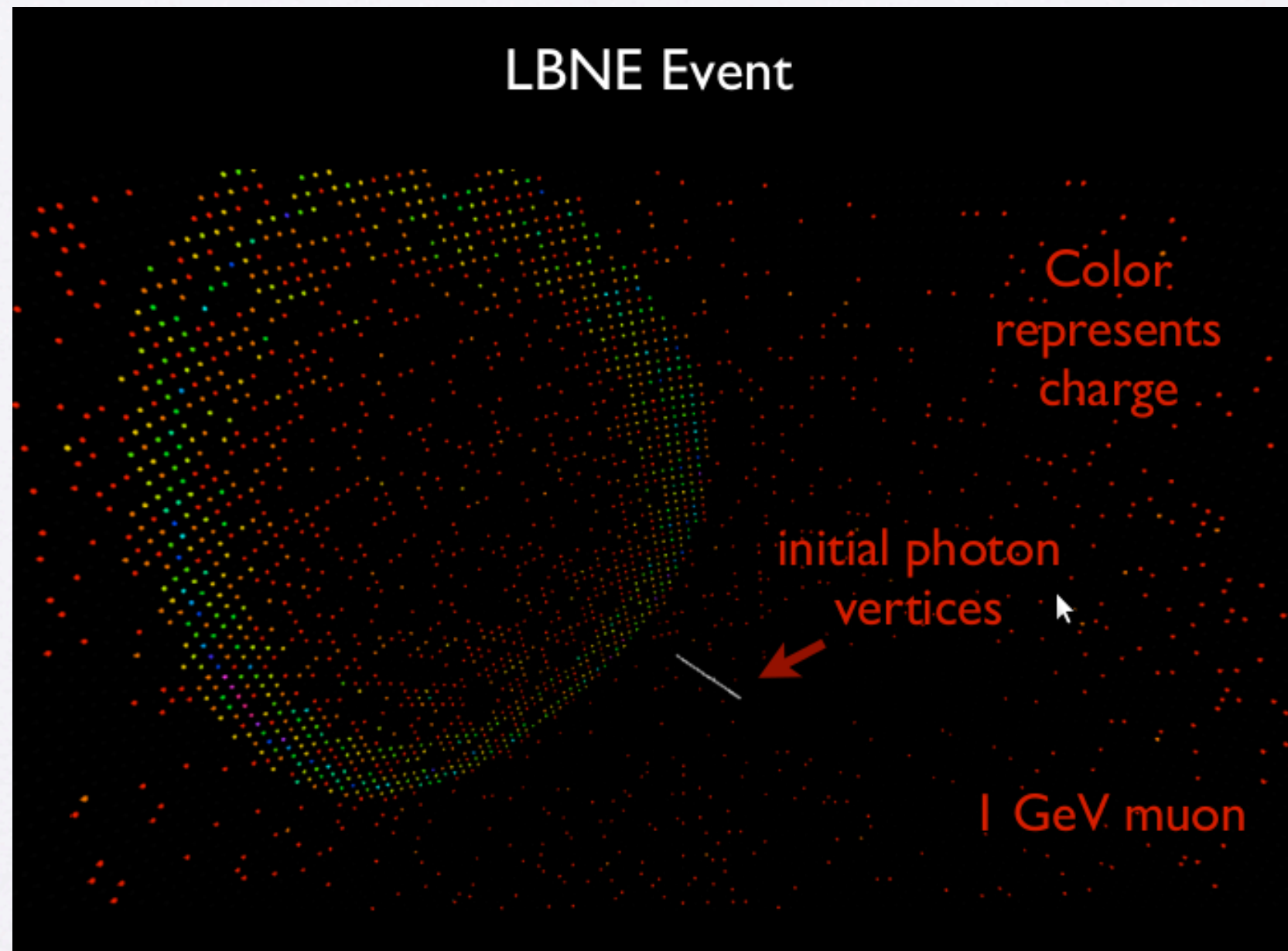
Open Data Policies

- Agencies have recently begun requiring open data policies, however our community does not have a plan to address this issue which is clearly shared across frontiers.
- For us, there is no clear avenue for sharing multi-TB or PB data samples, no clear acceptable format, no guidance for how we release these to the public.
- There are no additional resources to support these data sets once they are released, or to curate these samples over the long term.

IF: Chroma

Stan Seibert (UPenn)
Anthony LaTorre (UChicago)

- **Chroma is an open source optical photon Monte Carlo** which simulates standard physics processes such as diffuse and specular reflections, refraction, absorption, Rayleigh scattering, and scintillation.
- **Photons are propagated in parallel on many-core modern GPUs.**
- Chroma can propagate 2.5 million photons/sec in a large detector with 29,000 PMTs. **This is roughly 200 times faster than the same simulation with Geant4.**



- Geometries are defined by triangle mesh representing the surface between surfaces.
- The core track propagation can be use to visualize and navigate the detector in real time.
- Could this become a GEANT4 feature?

EF: Geant on HPCs

Tom LeCompte
(ANL)

- Atlas uses ~800M CPU-hours per year. They have 30K cores pledged for simulation at any given time, but they need and use much more (40-50k). Big backlog.
- Using these machines in ATLAS will require a front-end which accepts jobs to OSG, starts the job, does the initialization and db access, and then accepts the output and finalizes the job. Back-fill idle time.
- So where are they at? Geant, ROOT and ALPGEN all have been run at ANL's Intrepid.
- For a taste of performance: **generating 100M toy MC** using ROOT starting with a polynomial distribution, poisson fluctuation for each bin and fitt to a Gaussian. **Takes 18 mins on 0.6% of Intrepid which would take 500 CPU-hours on x86.**
- Right now **10,000 CPU hours of toy MC** which is **typically run in 200 nodes x 200 hours** which gets to the answer in **about a week**.
 - **Using HPCs you could do this in a day.**
- What would it be with GEANT multithreading?



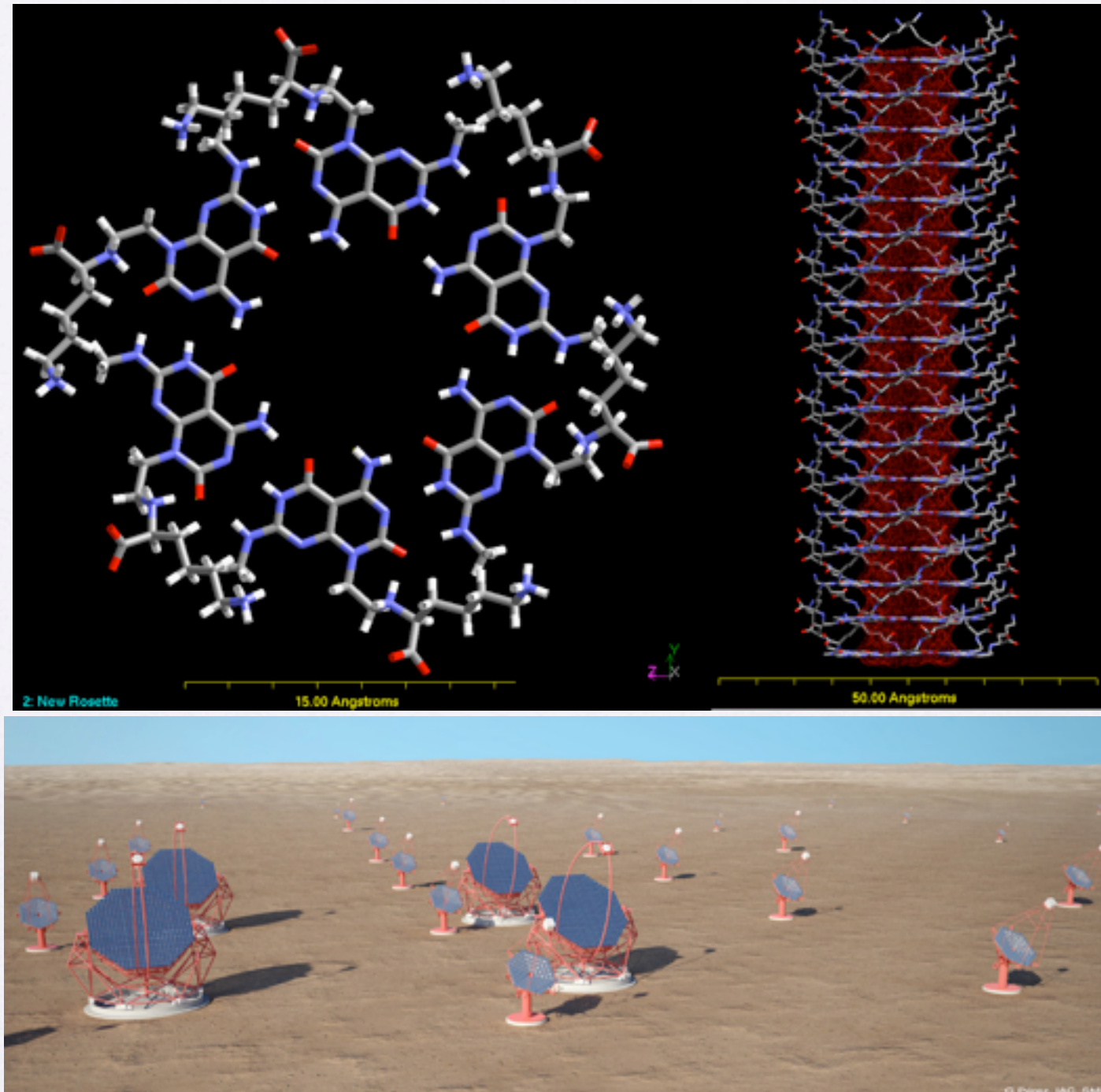
Meet ANL's Intrepid

- 40 racks, each with 1024 quad-core PowerPC 450 compute nodes = 163,840 cores
 - Roughly equivalent to 23,000 x86 cores.
- Snowmass on the Mississippi - Computing Frontier

CF: Self-assembling data

- The problems of packaging, transporting, and processing large volumes of data in real time are of critical importance throughout the frontiers.
- CTA must gather ~ 30 GB/s of data from ~ 100 telescopes distributed over a $\sim \text{km}^2$ area and process it in real time so that observation strategies can be modified in response to transient phenomena.
- A multidisciplinary group at the CF proposes to **design a fault-tolerant real-time association of information** across their large-scale experiment containing distributed sensors **by creating a self-assembling data paradigm**.

Applications to more transparent and efficient data distribution/storage



Summary

- There are significant computing requirements by current and future IF experiments and IF R&D.
- The quality and impact of the IF effort depends heavily on efficient and transparent access to dedicated computing resources.
- While resources are available for Fermilab-based experiments, all efforts will benefit from dedicated and transparent access to grid resources.
- Dedicated grid resources for the intensity frontier (in the form of intensity frontier VO?) would have the largest impact on our international efforts.
- Computing professionals are in demand as support for key software frameworks, software packages, scripting access to grid resources and data handling.
- There are efforts (and problems) that are shared across frontiers, significant investments in ROOT and GEANT4 optimizations, HPC for HEP, transparent OSG access and open data solutions.

Draft summary will be circulated by the end of the week!