Incorporating High Energy Physics Data Capabilities into Joint Forces Simulations

Julian J. Bunn and Thomas D. Gottschalk
Center for Advanced Computing Research, Caltech
Pasadena, California 91125
julian@cacr.caltech.edu, tdg@cacr.caltech.edu

ABSTRACT

The data management and data exploitation issues for large-scale, distributed DoD simulations have striking parallels within a number of existing large-scale High Energy Physics (HEP) projects, in particular, the experiments associated with the Large Hadron Collider (LHC) in Geneva, Switzerland. The significant commonalities include: data rates of 10-100 GBytes/day, data distribution and database operations over very large scale, high-speed networks, and sophisticated data exploitation objectives. In this regard, the lessons learned over the past decade of preparations for LHC operations have obvious significance and relevance for operational (fielded) DoD information exploitation systems. The requirements for persistent, scheduled, secure data access and data mining within the HEP environment are similar to many aspects of future large-scale DoD simulation environments, such as the Sentient World Simulation (SWS).

This paper explores three particular areas of DoD data exploitation needs with significant parallels within existing HEP/LHC work. The first involves robust, scalable database design and management, such as the distributed simulation and data system within the Joint SemiAutomated Forces project now under development within the US Joint Forces Command. Important aspects here include operational transparency and efficiency from the perspective of a single user/analyst at a workstation. The second general area involves support for “user toolkits” - significant additional computational subsystems such as data-mining/knowledge-discovery procedures and “what if” Monte Carlo excursions that go well beyond straightforward queries of a distributed database. The final area has to do with “real-time” considerations, where this term is to be understood in the more general sense of legitimate, possibly urgent user needs that exceed available computational resources. Strategies are discussed for leveraging the demonstrated HEP expertise toward DoD data management and exploitation problems, using SWS, on occasion, as a template for some specific DoD requirements.

ABOUT THE AUTHORS

Julian Bunn is a Member of the Professional Staff at Caltech, where he works in the Center for Advanced Computing Research. He held a Physicist position at CERN, the European Centre for Particle Physics for 15 years before joining Caltech. His research interests include high performance network and computing systems and Grid architectures for solving the data distribution and analysis challenges posed by CERN's Large Hadron Collider experiments. He gained his B.Sc (Hons) in Physics from the University of Manchester in 1980, and his Ph.D. in Experimental Particle Physics from the University of Sheffield in 1983.

Thomas D. Gottschalk is a Member of the Professional Staff, Center for Advanced Computing Research (CACR) and Lecturer in Physics at the California Institute of Technology. He has been with CACR for nearly a decade. Much of his research has been on the use of parallel computers to simulate various physical phenomena. His instructional duties include his upper division course on Statistics for Physics Graduate students. He also consults for the Aerospace Corporation and other scientific, financial and technical organizations. He received a B.S. in Physics from Michigan State University and a Ph.D. in Theoretical Physics from the University of Wisconsin.

Incorporating High Energy Physics Data Capabilities into Joint Forces Simulations

Julian J. Bunn and Thomas D. Gottschalk
Center for Advanced Computing Research, Caltech
Pasadena, California 91125
julian@cacr.caltech.edu, tdg@cacr.caltech.edu

I. CONTEXT: FORCES MODELING AND SIMULATION

The long-term goals of Forces Modeling and Simulation (FMS) within the DoD community are undeniably ambitious. The Distributed Continuous Experimentation Environment (DCEE) concept (Ceranowicz, 2003) addresses the essential role of reliable, very-large scale simulations for concept development and training, and notes the significant number of technical hurdles that must be overcome in bringing this vision into an effective operational reality. The more recent Sentient World Simulation (SWS) proposal (Cerri, 2005) presents an extremely ambitious vision for a continuously operating, integrated environment of simulation and real-world data, as would be essential to predict, analyze, and evaluate events and strategies incorporating military, political, and economic influences. (A brief, admittedly biased overview of the SWS concept is contained in the appendix.)

Considerable progress towards the DCEE objectives has been demonstrated within the context of the Joint Forces Command Experimentation Directorate's (JFCOM/J9) Joint Urban Operations (JUO) (Ceranowicz, 2002) and Counter Mortar Rocket (CMR) simulation experiments. These experiments have clearly demonstrated the utility of very large scale distributed simulations, using the JSAF applications suite and RTI-s communications among computational assets distributed from Virginia to Hawaii. High performance computing concepts are essential within the evolving JUO/CMR framework. Computation advances in support of these projects include Scalable Parallel Processors (SPPs) for the required computational power (Lucas 2003), scalable communications procedures (Barrett 2004 and Gottschalk 2005) and distributed, interactive database components for managing the enormous generated data sets (Graebner 2003 and Yao 2005).

An idealized (and, admittedly, somewhat biased) picture of the emerging information processing system for JUO/CMR is indicated by the schematic in Fig.(1). SPP resources at the Maui High Performance Computing Center (MHPCC) and the Aeronautical Systems Center Major Shared Resource Center (ASCMSRC) at Wright-Patterson AFB OH drive most of the entity simulation, with additional workstation clusters at TEC, J9/Norfolk, and SPAWAR/San Diego driving operator-controlled additional entities. The blue components of the figure indicate the real-time, interactive Simulation Data Grid (SDG) component (Yao 2005), providing analysts with real-time access to detailed information during the event itself, managing data production at rates nearing the GB/sec range.

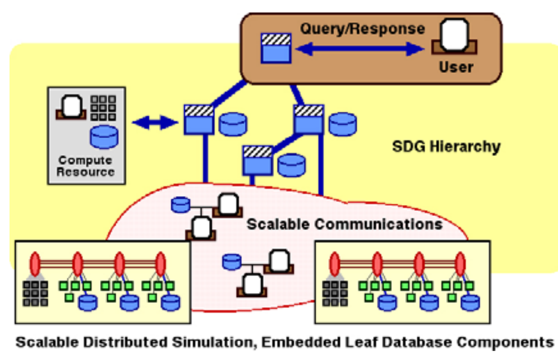

Figure 1. Integrated Information Management System for JUO/CMR

Steps towards the development and deployment of the system indicated in Fig.(1) have been done within the constraints noted in (Ceranowicz, 2003) for DCEE. In particular, deployment and utilization of new technology has, of necessity, been done in a strictly "vertical integration" fashion, with the requirement that new features never interfere with or impede ongoing operations of the system. This means that some potential information exploitation capabilities from a computing science perspective are integrated rather slowly into the entire system. One example, indicated by the grey components of Fig.(1), involves CPU-intensive "knowledge discovery" engines that continually sift through the wealth of data provided by the simulation. Such capabilities are clearly desirable for DCEE, both for experimentation/training and for development of operational, fielded systems. More recent efforts to incorporate the Synthetic Environment for Analysis and Simulation (SEAS) (Chaturvedi, 2005) into the JUO/CMR simulations represent additional important steps towards development of the integrated information objectives of SWS.

The idealized Scalable Information System of Fig.(1) is hardly a unique problem of the FMS community. "Big Science" has been faced with comparable problems of scale for years. The Center for Advanced Computing Research (CACR) at the California Institute of Technology (Caltech) has been concerned with large-scale (TeraFlOPS, PetaByte, and beyond) computing issues for over a decade, always within the context of real programs, with real schedules and often competing objectives. In this paper, we begin the process of exploiting the front-line advances of Big Science, adapting lessons learned to the specific problems of FMS.

The Compact Muon Solenoid (CMS[1]) experiment for the Large Hadron Collider (LHC[2]) facility in Geneva, Switzerland provides the context for our comparisons and discussions. As described in the next section, CMS/LHC is an extremely large undertaking, involving thousands of physicists all over the world, and a total construction budget of around three billion Swiss Francs (1 US Dollar ~ 1.25 Swiss Francs). The CMS experiment's detector will produce raw data at Pbyte/sec rates, which will be reduced by a fast pipelined and buffered online system (the "trigger") to a "mere" 100 Mbyte/sec data stream for subsequent storage and analysis. CACR has been at the forefront of research into and development of data management requirements for CMS/LHC since 1997, specifically in the context of a number of demonstration projects funded by major scientific agencies in both the USA and Europe (PPDG[3], GriPhyN[4] etc.). The results of these studies have directed the design and implementation of the information retrieval and management system for CMS/LHC.

After a brief description of the overall CMS/LHC system in the next section, this paper focuses on three aspects of the LHC Data Grid that appear to be particularly relevant for very large-scale DCEE/FMS:

- Robust, scalable database design..
- CPU-intensive computational tools.
- Operational issues associated with conflicting demands on an information system that is - in spite of best efforts - ultimately a limited-capacity resource.

We conclude with some brief remarks and suggestions on next steps for leveraging the demonstrated information management capabilities within the academic world to the problems now faced in DoD applications.

II. CONTEXT: HIGH ENERGY PHYSICS

The major High Energy Physics (HEP) experiments of the next decades are designed to provide new insights into the fundamental interactions, structures and symmetries that govern the nature of matter and space-time. Probing nature at this level of detail requires very large machines and very large, long-term collaborations among literally thousands of physicists at hundreds of institutes around the globe.

Figure 2: **The CERN site near Geneva.**

The largest elementary particle physics projects today are those associated with the LHC program at the CERN laboratory in Geneva, Switzerland. The LHC is constructed inside a circular tunnel that straddles the Franco-Swiss border near Geneva. The tunnel has a circumference of about 18 miles, and contains a vacuum pipe in which accelerated beams of high energy protons circulate. The superconducting magnets that maintain circulation of the beams are cooled by a vast plumbing system containing liquid Helium, pumped from the surface. In fact, the proton beams circulate in both directions around the ring, and each proton carries about 7 TeV of energy. (A TeV is a unit of energy used in particle physics. 1 TeV is about the energy of motion of a flying mosquito. What makes the LHC so extraordinary is that it squeezes energy into a

[1] CMS (Compact Muon Solenoid), http://cms.cern.edu
[2] LHC (Large Hadron Collider), http://www.cern.ch
[3] PPDG (Particle Physics Data Grid), http://ppdg.net
[4] GriPhyN (Grid Physics Network), http://griphyn.org

space about a million million (i.e., quadrillion) times smaller than a mosquito.)

The beams cross over at four points around the tunnel, causing collisions between oncoming protons, with an overall collision energy of 14TeV. Four detectors are placed at the collision region, measuring the particles produced in the collisions.

An aerial view of the CERN site is shown in Fig.(2). The main runway at Geneva International airport is visible on the right. Lake Geneva (Lac Leman) is in the upper right. The superimposed white rings show the position of the underground collider tunnels.

The very high intensity of the LHC beams will ensure enormous numbers of these proton-proton collisions. Many of them will produce spectacular bursts of particles with hundreds of charged particles: protons, electrons, and more unusual beasts such as muons, pions, and so forth.

Extracting physics from the collisions requires careful identification and measurements of all the produced particles. This, in turn, requires extraordinarily large and complex particle detectors. The detectors include thousands of electronic systems that digitize the data for collection, amounting to an aggregate of around 100 million channels for the CMS detector as a whole.

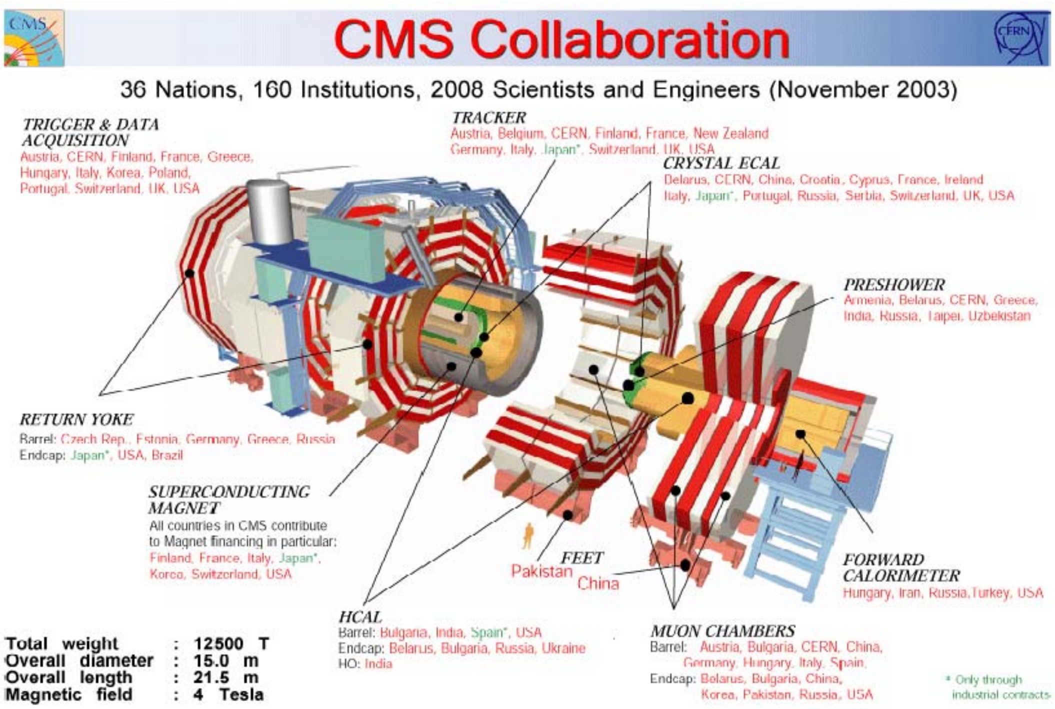

Figure 3. The CMS Experiment at CERN

Figure 3 shows a schematic of the Compact Muon Solenoid (CMS) detector (the other major detector at LHC is ATLAS, which is of a similar design). The device is huge, with an overall diameter of 15m and a length of 21.5m. (Note the "little man" near the FEET component, added to the schematic to provide a sense of size.) The collaboration designing, building, and eventually operating the CMS experiment is correspondingly huge, involving more than 2000 scientists and engineers from 160 member institutions in 36 nations.

The large sizes of both the accelerator in Fig.(2) and the detector in Fig.(3) are dictated by the huge proton energies required to explore physics at this new frontier. The Holy Grail of the LHC's physics program is the observation of the "Higgs particle", which is predicted by the currently most favored theory in particle physics, the so-called "Standard Model". If all goes well, Higgs particles will be observed at the LHC, further supporting the validity of the Standard Model, and allowing detailed studies to be made of their properties. On the other hand, if the Higgs particles are not seen, then this would be almost as exciting since it would cause a major re-thinking of accepted theory.

The detection of a usable sample with O(10^2) cleanly identifiable Higgs production events per year implies that the number of garden variety (i.e., boring) proton-proton interactions needs to be beyond one million interactions per second. Clearly, the CMS experiment is only interested in recording the new and interesting physics events, and so some way of filtering these out, or "triggering" on the interesting events, is required.

Triggering on events of interest requires a good understanding of how such events will appear in the detector. Early on in the processing chain, events are rejected if their topologies do not fit a predetermined set of desired categories. Later on, more sophisticated selection mechanisms reduce the event rate even further. The hierarchy of selections and refinements tame the huge raw data stream from the detector, and the rate of events destined for storage and analysis is reduced by at least six orders of magnitude, to around 100Hz.

Each event that emerges from the online triggering system is akin to a successfully isolated needle in a haystack that is many orders of magnitude deep. However, even after selection, the good events are still not easily handled. Not only are the basic events of interest complicated, they are typically subsumed beneath tens of "normal" events occurring within the same read-out cycle of the detector.

The data management/interpretation problems faced by CMS and other large HEP experiments are clearly enormous. Planning for the needed networks and computational facilities to process CMS data began in the early/mid 1990's, with Caltech's CACR playing an important role in both concept development and prototype demonstration. The wish list from the initial/ongoing studies was fairly daunting:

- Provide rapid access to enormous data stores: PetaBytes in 2007 to ExaBytes (10^{18} bytes) in 2015.
- Provide secure, efficient, transparent access to heterogeneous world-wide distributed data sets and computational resources.
- Match resource usage to policy goals.
- Provide a collaborative infrastructure, enabling efficient independent and joint utilization of available data by physicists across the world.
- Design/build regional, national, continental and transoceanic networks, with bandwidths rising from Gigabit/sec to Terabit/sec over the next decade.

As described in (Bunn 2003), many of these issues were addressed through a number of large demonstration projects within the HEP community over the past decade. The following sections highlight advances and issues within HEP efforts, noting similarities and differences relevant for DCEE.

III. THE LHC DATA GRID HIERARCHY

The designed data rates for CMS/LHC are truly enormous. Ignoring the sublime wizardry of the on-line triggering system, considerations of expected data rates, accumulated yearly volumes, and required processing power quickly drove plans for the LHC computing system to a <u>distributed computational grid</u>.

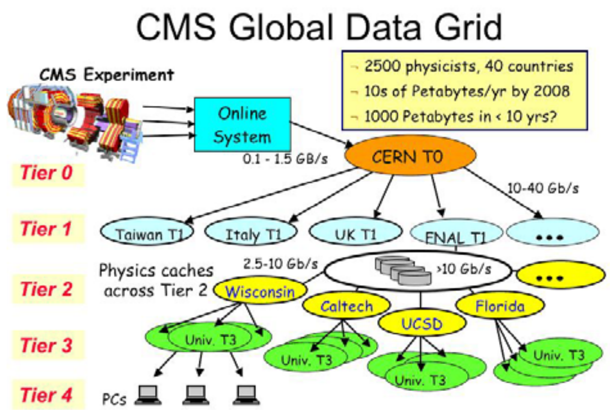

Figure 4: The LHC Data Grid Hierarchy Model

The architecture of the LHC Data Grid Hierarchy shown in Fig.(4) resulted from a number of studies and experiments, including GIOD (Bunn, 2001) and MONARC (Ghiselli, 2001). Raw event data emerging from the online system at around 100Mbytes/sec are stored at the CERN site (repository) and replicated, stored and thus shared at a number of Tier 1 sites around the world connected to CERN over one or more 10Gbit/sec WAN links. The raw data format consists of digitized signals from the many subdetectors that form the complete CMS. Thus each event in this raw format needs to be converted to a form useful for physics analysis. This involves reconstructing the particle tracks and energy deposits from the raw digital information.

The reconstruction program, which processes <u>every</u> event emerging from the online system, is CPU and I/O intensive, and requires access to databases containing the highly complex detector geometry and calibration constants essential to make sense of the data. The reconstruction program in fact runs constantly at the

CERN Tier0, and at the Tier1 centers, thus sharing the task in the first layer of the computational grid.

The reconstructed events are collected together in large files that together form event collections. The events are categorized by broad, topology-based physics “types”. Individuals and groups of physicists who are working on particular analysis problem access these event collections. For example, events that are categorized as being likely due to the decay of a Higgs particle into four muons form one large (and undoubtedly popular) collection. The collections are distributed to the Tier2 centers, which is where the lion’s share of the physics analysis is done.

The Tier2 centers typically comprise large clusters of compute and I/O server nodes, and are connected to the parent Tier1 by WAN links of at least 10Gbits/sec. Their primary purpose is to support data intensive analysis by university based physicists. In the USA, Caltech and UCSD house the first prototype of Tier2 centers, and there are now others at Wisconsin, University of Florida, Purdue and Nebraska.

At the next level down in the hierarchy are the Tier3 centers, conceptually small university departments hosting a few physicists, equipped with small clusters or a couple of decent servers capable of analysis on a modest scale. Finally, at the bottom of the hierarchy, but at the top of the event food chain, are the end user physicists, sitting in front of their Tier4 workstations, PCs, laptops, or PDAs. Such devices are capable of limited analysis, but are interfaced to the rest of the computational grid by a set of Web Services based tools that allow transparent access to all the grid resources.

This architecture is intended to provide a simple logical view of all objects and resources needed to perform physics analysis at the LHC. Studies identified several motivating goals of the architecture described above:

- Objects that cannot be recomputed must be stored somewhere.
- Physicists at any remote site must be able to query/retrieve any object.
- The requirements must be met efficiently and affordably.

There are several aspects of the LHC Data Grid Hierarchy that are immediately relevant for FMS applications in general and data exploitation frameworks (e.g., SDG), in particular:

- The underlying database is distributed over a very large network.
- Associated data set sizes are comparable.
- Operations at the user level must be transparent.
- Capabilities are required for significant, user-directed data manipulations.
- Security and data integrity are critical.

The expertise available within the HEP community on these issues could clearly aid development of similar capabilities for FMS. In particular, existing strategies for optimizations of the data access/manipulation procedures could perhaps be adapted to FMS applications.

The initial data-distribution Tiers in Fig.(4) are not immediately relevant for FMS, where primary data are generated from distributed assets rather than from a single, huge monolithic source. Nonetheless, the lessons in efficient data communications gleaned from the HEP Data Grid activities could, by themselves, have value for the different but equally intense data distribution problems within large-scale Discrete Event Simulations (DES), in general.

An additional point of interest has to do with the nature/schema of the underlying distributed database. For all the apparent complexity of Fig.(3), the reconstructed physics data are fairly simple - lists of momentum states for elementary particles within an event. The database schemas used for FMS, by contrast, are quite a bit more complex. Typically (e.g., in SDG), all entries in the relational tables are indexed by the full Federation Object Model (FOM) of the exercise. This is consistent with the general considerations and constraints of DCEE, as described in (Ceranowicz 2003). However, the simple expediency of using a common denominator FOM to index basic database components comes at a significant price in terms of database complexity and efficiency. These “nature of the data” complexities are even more significant for SWS.

The overall Data Grid model in Fig.(4) emerged after careful studies based on efficiency and effectiveness of overall operations - not component operations in isolation. It seems certain that similar considerations will be fruitful for FMS. At a minimum, flexible low-level database implications should be considered sooner rather than later as advanced information exploitation systems such as SDG are deployed.

IV. DISTRIBUTED COMPUTING

The overall network designs and data distribution models of Fig.(4) are simply the beginnings of the overall CMS data management system. Effective exploitation of the event data requires efficient computational models and procedures.

This section looks in more detail at three particular aspects of the evolving CMS/LHS model:

- Grid Computing
- Real-Time Analysis Capabilities
- User-Comfort Considerations

These were important, up-front factors in the eventual design and implementation of the overall HEP data management system. Similar considerations will undoubtedly drive effective realizations of the schematic FMS information management system of Fig.(1) – and even more so the ambitious objectives of SWS.

Grid Computing

The success of a distributed computing model hinges on the availability of excellent network connectivity between the computing resources and on rational allocation and use of these distributed resources. This was recognized early on in the development of the LHC Grid hierarchy.

At the outset, network bandwidth extrapolations indicated that there might be several bandwidth bottlenecks that would prevent the distribution of the LHC data to the participating institutes around the world. It was even considered that bulk freight of tapes might be needed. Happily, due to various commercial and technological factors, the availability and capacity of international high-speed networks has increased beyond expectations. Although by no means ubiquitous (especially to and within the developing countries), high-speed WAN links are already in place and planned to meet and exceed the required bandwidth for distributed LHC data processing and analysis in 2007 and beyond.

Indeed, the emerging network situation has revealed a new limitation in the LHC computing model: the ability of end systems and servers to cope with the capacities of the network pipes that interconnect them. Currently there are efforts underway (e.g. UltraLight[5]) to make sure that the networks are treated as an equally crucial component in the ensemble of computing resources that make up the computational grid, and to balance the network capabilities with the end systems, and in particular with the I/O storage subsystems.

The use of Grid technology by the LHC experiments is now well established in the context of dynamically deployed compute resources for large scale production tasks. Solutions are in place that utilize tried and trusted batch job infrastructure and tools, such as Condor[6], PBS[7], and LSF[8]. Indeed, many millions of simulated events have been generated using the resources of the worldwide Grids already available to the LHC experiments. Coupled with the batch infrastructure is a set of software, collectively called the "Virtual Data Toolkit" (VDT)[9], which together form a suite of compatible tools for Grid computing. The VDT is in wide use, and is a pre-requisite installation on all clusters that need to participate in the LHC Grid.

As is acknowledged in (Cerri, 2005), grid computing is essential in large-scale FMS systems such as SWS. The relevant "lessons learned" from HEP in this regard include the observation that the network itself is only a modest part of the overall grid computing package. Long term FMS objectives clearly require grid computing in the sense of significant computational horsepower that is transparently available.

However, grid computing implies much more than "mere" distributed CPUs linked by adequate network bandwidth. This much is already routinely done within a number of FMS applications, such as JSAF/JUO. The LHC development program has led to the (not surprising) conclusion that data organization and user-friendly computational services are the truly essential components if the available computational/network resources are to be exploited fully, particularly in the context of real-time analysis, as is next discussed.

Real-Time Analysis Capabilities

More recently, attention is turning to the use of Grids for real time analysis of the LHC data by end user physicists. This is a much more challenging problem, but it is ultimately the acid test of the utility of Grid computing for Particle Physics - and in FMS as well. The analysis of the LHC data is what will result in the physics discoveries. Similarly, the ultimate measures of

[5] UltraLight, http://ultralight.caltech.edu

[6] Condor, http://www.cs.wisc.edu/condor

[7] PBS (Portable Batch System), see http://hpc.sissa.it/pbs/pbs.htm

[8] LSF (Load Sharing Facility), see http://platform.com

[9] See http://vdt.cs.wisc.edu

effectiveness and value within FMS will come from real-time user analyses and gained insights within complex, dynamic environments.

Data analysis in particle physics typically involves isolating a set of events that meet certain selection criteria, and then deriving from those events a set of physical characteristics, or identifying a few unusual features. The event sets in question may range in size from a few dozen events for the smallest, to many millions of events for the largest, which corresponds to TeraByte scale file sizes.

Providing high-speed access to such large analysis data sets for individual users will measure the success of the LHC Grid. The complication is that either these data sets need to be created on-the-fly by a compute and I/O intensive algorithm or the much larger data set against which the selection is being made needs to be moved to a location with free compute cycles and storage space. Moreover, the demands on the distributed system are being made by a distributed (unruly?) community of physicists.

The grid computing task requests are typically chaotic, unpredictable and extremely hard to anticipate. This requires an interrupt-driven distributed system that is constantly in a state of reconfiguration as it tries to match the end user demands with available resources and prevailing conditions. The operational constraints for SWS (24x7 operations with occasional, extremely CPU-intensive excursions) further underscore the needs for careful high-level design within the metacomputing system. Existing FMS systems such as DARWARS[10] are almost certain to face similar, uneven resource demands during ongoing 24x7 operations.

This is actually a very important point in near-term implementations of initial FMS systems, such as the SDG model of (Yao 2005). A mature information management system for FMS must also cope with arbitrary, unscheduled user requests. Indeed, analysts' queries with the initial database implementation for JUO occasionally resulted in significant degradations of overall simulation performance.

The CMS experiment has investigated several technologies that could make such a system a reality. One very crucial component of such a system is a *sophisticated monitoring and control mechanism* that is constantly keeping tabs on the state of all significant resources within the system, together with all tasks executing or awaiting execution in the system. The monitoring and control system is able to reconfigure the resources dynamically so as to respond to unexpected or unusual conditions (such as, for example, the loss of network connectivity between sites, or the exhaustion of space in a storage subsystem). Indeed, the monitoring and control is a crucial component in the planning stage for a new task that enters the system, capable of providing strong hints on placement based on existing and historical trend analysis for similar (or identical) tasks in the system. The CMS group has developed such a monitoring and control system called MonALISA[11], which is included as part of the Virtual Data Toolkit mentioned above.

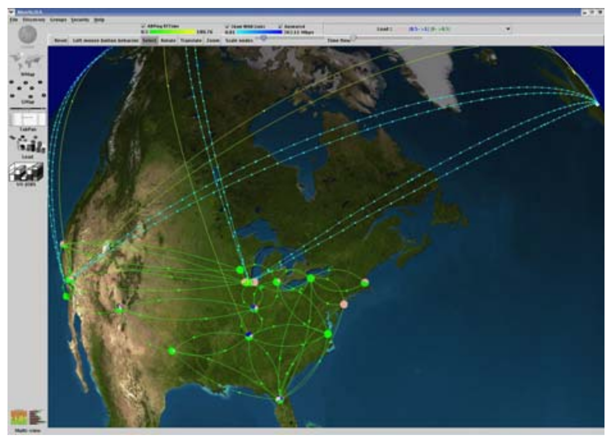

Figure 5: A MonALISA display.

Figure 5 shows a screenshot of the MonALISA agent-based monitoring tool. The distributed clusters engaged in a massively distributed simulation task are shown as colored circles. The lines indicate active network connections among the compute sites. Color codings indicate compute and network loads. This display is dynamic and (essentially) real-time.

The User Comfort Factor

Another important component of a grid system that supports analysis is a means whereby an end user can exploit the full capabilities of the global resources without having to jettison his/her favorite analysis tool. Particle physicists tend to be very faithful to a chosen analysis tool, in which events are processed, selections applied, histograms filled, paper-ready graphics are generated. They are not amenable to discarding this

[10] See http://www.darwars.com

[11] MonALISA http://monalisa.cacr.caltech.edu

very familiar, comfortable and richly featured analysis tool just to "fit in" with what Grid architects try to impose.

The Grid system thus must provide interfaces to the most popular of the analysis tools in use. In that way, the end user *retains the comfort and familiarity of her/his chosen analysis tool*. Beyond that, he/she can call out to the enormous resources of the Grid, when required, to expand the scope of the analysis. To realize this goal, the CMS team has developed a Web Services based framework called Clarens (van Lingen, 2005) that offers a single sign-on to the Grid, and subsequent access to all the Grid resources, all while never leaving the GUI of the most popular analysis tools being used in CMS.

Accommodation of familiar user tools is hardly a unique requirement or constraint of the HEP world. A number of existing and emerging interfaces exist within FMS, particularly in terms of expected displays for C3 components. The web-based paradigm used within the CMS Data Grid is another area with applicability to FMS overall system design. This issue will be extremely important for Sentient World Simulation. Indeed, SWS involves a merging of a number of existing sophisticated operational elements:

- Large-scale constructive simulations
- Emerging socio-economic simulation schemes
- Real-time, WWW-based information extraction and interpretation.

Each component has been developed somewhat in isolation, with large investments in "standard tools". Fusing these tools in a coherent manner will be a challenge.

V. OPERATIONS

An important operational consideration for the CMS Data Grid is the management of policy, authentication and authorization. Physicists working on CMS are affiliated with an institution, and involved in one or more branches of physics analysis. Some of the analysis tasks are carried out individually, and some tasks by groups. A physicist may work on one topic today, and a different topic tomorrow. The CMS Data Grid has to support all such activities. The CMS collaboration as a whole has a set of physics and computing goals, which are prioritized. For example, the collaboration has as very high priority the timely reconstruction of the raw data as it emerges from the detector. Reconstruction gives way somewhat to analysis, however, whenever major physics conferences occur.

The means by which policies, authentication and authorization is supported can be summarized in the term Virtual Organization (VO). A VO is a collection of physicists who share a common goal, for example, the study of the decay of the Higgs to two photons, or the re-processing of all raw data using a new particle tracking algorithm. Physicists can join and leave a VO at will, simply by authenticating themselves with the VO Membership System (VOMS). When a VO is joined, the physicist gains access to resources (e.g. datasets, or CPU allocations) particular to the VO. The same physicist can simultaneously be a member of more than one VO, perhaps in the morning working on the Higgs analysis and in the afternoon dedicating time to managing the reprocessing of raw data. The VO system takes care of matching authenticated users with access to resources, and applying policies (such as CPU or storage allocations) dynamically.

As truly large-scale FMS matures from specific experiments such as JUO/CMR towards an operational tool such as SWS, the scheduling issues will undoubtedly be even more serious than those faced within HEP. It is important that the constraints of eventual operations be considered early (and often) as the system suggested in, *e.g.*, the SWS vision paper is constructed. Note that a "Virtual Organization" for FMS will need to deal with a number of issues not common in HEP, particularly issues related to (multi-level) security. The lessons from HEP are that these eventual operational issues should be considered early in the design phase of any large-scale system.

VI. CONCLUDING OBSERVATIONS

The Sentient World Simulation vision paper (Cerri, 2005) includes a list of high-level requirements for a successful SWS program. These requirements are applicable to more general large-scale FMS, and include:

1. Continuously updated knowledge base.
2. Effective organizational structure.
3. Adequate computational and data grids.
4. Live, historical and synthetic data.
5. Software services transparent to users.
6. High bandwidth networks.
7. Secure, trustworthy environments.
8. State-of-the-art facilities.

It is our contention that the existing large-scale HEP programs (such as LHC) provide significant practical insights relevant to essentially all the elements in this requirements list – a list that is clearly applicable to other large-scale FMS applications.

In terms of effective utilization of high-performance computing and data management, there is clearly much value to be found in the existing models and procedures within the High Energy Physics community. Successful incorporation of HEP advances will, however, require some effort.

Many of the difficulties are likely to follow from two fundamental differences between the overall HEP and FMS data interpretation problems:

1. Data generation for HEP is largely divorced from Human In The Loop (HITL) effects, while HITL is essential to the very concept of FMS.

2. The elementary data of HEP are ultimately far simpler than the data needed to capture the "essence" of an FMS exercise.

These two differences will be particularly significant in adapting specific software products from the Data Grid effort. Nonetheless, the successful frameworks already developed for HEP are valuable - and adaptable.

It is probably best to explore multiple strategies for incorporating the insights and successes from large-computing HEP research into FMS activities. At one end, specific insights from the large data management system of Figure 4 have immediate applicability to the SDG. Explorations and exploitations of these commonalities are underway within the JESPP (Joint Experimentation, Scalable Parallel Processors) project, with particular attention given to the different character of FMS and HEP basic data.

A second aspect has to do with overall information system concepts. The LHC Grid in Figure 4 emerged after much discussion of both visionary goals and the actual usage model (meaning poor post-docs chained to Tier 4 workstations). Both considerations influenced the final architecture. An effective realization of Figure 1 cannot happen without comparable (indeed, more difficult) considerations of the objective, fielded information system.

The insights from HEP in this regard may ultimately be the most useful products of careful HEP⇔FMS studies. A strategy for pursuing this high-level synergy is greatly needed. Might we suggest that the time to collaborate is now!

APPENDIX: SENTIENT WORLD SIMULATION

Sentient World Simulation (SWS) is a newly proposed concept for a "continuously running, continuously updated mirror model of the real world that can be used to predict and evaluate future events and courses of action". Formal documentation of the SWS concept is still somewhat sparse[12], so this appendix provides a brief overview.

Figure 6 presents a very schematic architecture for the SWS system, from the overall systems perspective of this work. There are several essential components to the system:

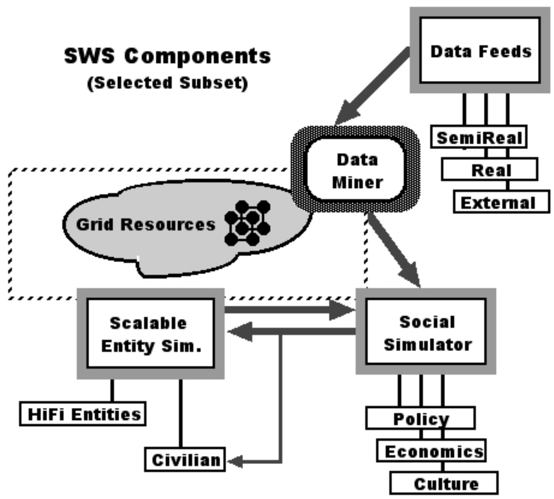

Figure 6: An Idealized SWS Schematic

- Scalable entity simulator – e.g., large scale distributed JSAF, including both high fidelity entities and a much larger population of lower fidelity civilian "background"
- A substantial political, military, social, economic, information and infrastructure (PMESII) model that, in essence, determines and drives consistent behaviors of the simulated civilian component.
- Real time data collection components, extracting relevant real-world data from various sources (newspapers, www, etc.).
- Data mining capabilities to extract significant "trends" from the real-world data feeds.
- Adequate, "on-demand" computing resources – e.g., substantial access to adequate computational grids.

[12] Some SWS documentation is collected in http://www.isi.edu/~ddavis/SWS

The SWS system is intended to provide a continuously running, 24x7 synthetic world. Its uses would include both concept explorations (e.g., for the active "entity simulator" end) and trend identifications within the "social" component.

At the present time, SWS is arguably simply a "grand concept". It is certainly not yet a defined program with well defined requirements. However, from our perspective, it is an extremely interesting "target system" for future FMS technology objectives. We believe that HEP experience in similarly very large information environments could, in fact, be useful in formulating next level, implementable steps towards a real SWS system. The strongest similarities between SWS and HEP/CMS systems are related to data management and knowledge discovery within very large, always operational information systems. HEP experience in distributed data management, on-demand resource allocation (grid computing) and data mining can provide significant insights to the development of similar functionalities as SWS evolves from a concept to a functioning system.

APPENDIX: ACRONYMS

ASCMSRC:	Aeronautical Systems Center Major Shared Resource Center
CACR:	Center for Advanced Computing Research
CERN:	*Conseil Européen pour la Recherche Nucléaire* (European Council for Nuclear Research)
CMR:	Counter Mortar/Rocket
CMS:	Compact Muon Solenoid
DCEE:	Distributed Continuous Experimentation Environment
DoD:	Department of Defense
FMS:	Forces Modeling and Simulation
GIOD:	Globally Interconnected Object Database
GriPhyN:	Grid Physics Network
HEP:	High Energy Physics
HITL:	Human In The Loop
JFCOM:	Joint Forces Command
JSAF:	Joint Semi-Automated Forces
JUO:	Joint Urban Operation
LHC:	Large Hadron Collider
LSF:	Load Sharing Facility
MonALISA:	Monitoring Agents using a Large Integrated Services Architecture
MONARC:	Models of Network Analysis at Regional Centers
MHPCC:	Maui High Performance Computing Center
PBS:	Portable Batch System
PPDG:	Particle Physics Data Grid
SDG:	Simulation Data Grid
SEAS:	Synthetic Environment for Analysis and Simulation
SPP:	Scalable Parallel Processor
SWS:	Sentient World Simulation
VO:	Virtual Organization

ACKNOWLEDGEMENTS

This material is based on research sponsored by the Air Force Research Laboratory under agreement number FA8750-05-2-0204. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation thereon. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the Air Force Research Laboratory or the U.S. Government. We are pleased to acknowledge several useful discussions with our colleagues at Caltech/CACR and USC/ISI during the writing of this manuscript.

REFERENCES

Barrett, B. and Gottschalk, T. (2004), Advanced Message Routing for Scalable Distributed Simulations, Proceedings of the 2004 I/ITSEC Conference, Orlando Florida.

Bunn J, Holtman K, Newman H, Wilkinson R (2001); *The GIOD Project – Globally Interconnected Object Databases* – Comp. Phys. Comm. 140 (2001) 162-171

Bunn, J, Holtman K, Newman H, Wilkinson R (2001) Globally Interconnected Object Databases (GIOD), Comp. Phys. Comm. 140 (2001) 162.

Bunn J and Newman H (2003); Data Intensive Grids for High Energy Physics, in Grid Computing, Making the Global Infrastructure a Reality, Berman, Fox and Hey (Ed.), Wiley, UK, 2003, ISBN 0-470-85319-0

Ceranowicz, A., Torpey, M., Hellfinstine, W., Evans, J. and Hines, J. (2002), Reflections on Building the Joint Experimentation Federation, Proceedings of the 2002 I/ITSEC Conference, Orlando Florida.

Ceranowicz,A., Dehncke, R., and Cerri, T. (2003), Moving Towards a Distributed Continuous Experimentation Environment, Proceedings of the 2003 I/ITSEC Conference, Orlando, Florida.

Cerri, T, and A. Chaturvedi, Sentient World Simulation, white paper version 1.0, 2005..

Chaturvedi, A. et al. (2005), Bridging Kinematic and Non-Kinematic Interactions Over Time and Space Continua, Proceedings of the 2005 I/ITSEC Conference, Orlando, Florida.

Ghiselli A. et al., (2001), Models of Network Analysys at Regional Centers (MONARC), Comp. Phys. Comm 140 (2001) 219.

Gottschalk, T. and Amburn, P. (2005), Extending The MeshRouter Framework for Distributed Simulations, Proceedings of the 2005 I/ITSEC Conference, Orlando, Florida.

Graebner, R., Rafuse, G., Miller, R. and Yao, K.-T. (2003) , The Road to Successful Joint Experimentation Starts at the Data Collection Trail, Part III, Proceedings of the 2003 I/ITSEC Conference, Orlando Florida.

Lucas R., and Davis, D., Joint Experimentation on Scalable Parallel Processors (JESPP), Proceedings of the 2003 I/ITSEC Conference, Orlando Florida.

Yao, K.-T., and Wagenbreth, G. (2005), SCALED Data Grid: Joint Experimentation Data Management and Analysis, Proceedings of the 2005 I/ITSEC Conference, Orlando, Florida.

F. van Lingen, J. Bunn, I. Legrand, H. Newman, C. Steenberg, M. Thomas, A. Anjum, T. Azim, (2005) “The Clarens Web Service Framework for Distributed Scientific Analysis in Grid Projects” , Submitted to the Workshop on Web and Grid Services for Scientific Data Analysis (WAGSSDA), June 14-17, 2005