Post-Secondary Education Network Security: AcceleratingDiscovery atan Experimental Facility

Prof. Haresh R. Parmar

Assistant Professor in
Computer Engineering Department
Silver Oak College of Engineering &
Technology, Ahmedabad.

Prof. Hemal U. Patel

Assistant Professor in
Computer Engineering Department
Silver Oak College of Engineering &
Technology, Ahmedabad.

Abstract

Advances in both sensor and computing technologies promise new approaches to discovery in materials science and engineering. For example, it appears possible to integrate theoretical modeling and experiment in new ways, test existing models with unprecedented rigor, and infer entirely new models from first principles. But, before these new approaches can become useful in practice, practitioners must be able to work with petabytes and petaflops as intuitively and interactively as tToday, irrespective of what career field a college graduate enters, personal computer literacy is agiven requirement. Personal computer security literacy is rapidly becoming as important as officeapplication software literacy for today's typical employee. Coping with technology security issues is notsomething that can be simply accomplished through personal experiences. Currently, research ofyoung adults and students indicates that 7 out of 10 frequently ignore IT policies, and 3 of 5 believethey are not responsible for protecting information and devices. In the past, fallout from poor IT habitswas buffered by the IT department's iron control over the infrastructure.hey do with gigabytes and gigaflops today.The Discovery Engines for Big Data project at Argonne National Laboratory is tackling key bottlenecks along the end-to-end discovery path, focusing in particular on opportunities at Argonne's Advanced Photon Source. Here, we describe results relating to data acquisition, management,

And analysis. For acquisition, we describe automated pipelines based on Globus services that link instruments, computations, and people for rapid and reliable data exchange. For management,

we describe digital asset management solutions that enable the capture, management, sharing, publication, and discovery of large quantities of complex and diverse data, along with associated metadata and programs. For analysis, we describe the use of 100K+ supercomputer cores to enable new research modalities based on near-real-time processing and feedback, and the use of Swift parallel scripting to facilitate authoring, understanding, and reuse of data generation, transformation, and analysis software.

1.1 Significance of the problem

The explosion of the Internet, though it has benefited society a lot, has brought along with it newerways to commit frauds, scams, robberies, and so forth. Network fraud has grown as the Internet hasbecome popular. As one CSI survey found [2], such crime peaked to record highs in 2001 to$3,149,000 per respondent, before everyone realized the importance of network security. Networkrelated fraud, though, has come down sharply to $269,000 per respondent in 2008 [2] due to variousmeasures being taken; however, it still causes substantial losses to various organizations. The CSIsurvey [2] also found an increase in unauthorized access of networks from 25% in 2007 to 29% in2008 despite a drop in all other types of incidents. The Internet Crime Compliant Center (IC3) in their2009 Internet Crime Report [3] found that the increase in complaints from 2008 to 2009 was 23%. TheIC3 report [3] indicates that of the top five categories of offenses reported to law enforcement during2009, non-delivered merchandise and/or payment occurred 19.9% of the time; identity theft, 14.1%;credit card fraud, 10.4%; online auction fraud, 10.3%; and computer fraud(destruction/damage/vandalismof property), 7.9%.

1.2 Student end user computer security concerns

Every person should care about computer security because an attacker can not only access thedocuments stored in the computer but also can use the computer to send forged messages andlaunch attacks on other computers. Arian [4] defines computer security as the process of preventingand/or detecting unauthorized use of your computer. With increasing software complexities it hasbecome difficult to completely secure computer systems against vulnerabilities. An attacker can usethese vulnerabilities to get into a computer and launch an attack. Unless taught, the typical end userhas limited awareness of how to protect themselves; they do not know how to use proper settings (or“least privileged user account”) for the software programs so that an attacker can not use them toaccess their computer.

Accelerating data movement

It is common practice today at facilities such as the APS for data collected during experimentsto be transferred to hard drives, carried home by the investigator, and only thenanalyzed—a process that may take weeks or months. This approach has many disadvantages.The investigator may subsequently discover that the data was taken incorrectly, inwhich case the experiment was wasted and perhaps cannot be repeated for many months(if at all). Beyond this, innovative methods in which online analysis results are used tosteer experiments towards optimal solutions are impossible using these practices.Particularly as data volumes and the complexity of the algorithms and software requiredto make best use of data (e.g., to extract information from noisy data) grow, newapproaches to computing are required in which: (1) data is delivered in real time, asit is collected, to storage systems large enough to hold the data for extended periodsand to computers powerful enough to perform a range of analyses; (2) analyses can beperformed with great rapidity to determine whether data is useful and to provide otherfeedback to investigators during an experiment; and (3) other analyses can be performedover time, for example to extract additional information and/or to compare with otherdata—with results that may provide new research insights and/or guide future experiments.Such new approaches, when combined with innovations in analysis methods, canboth allow for more efficient use of current facilities, permitting more experiments and

More users, and extend the capabilities of current and future experimental facilities.

The benefits of rapid data capture, analysis, and dissemination have been long understood. However, limited resources mean that the small teams that run experimentsat facilities such as the APS still lack the capabilities required to exploit fullythe large data sets now being produced To address these issues, we have developed

Methods, services, and tools both for managing big data and for interactive analysis ofsuch data.

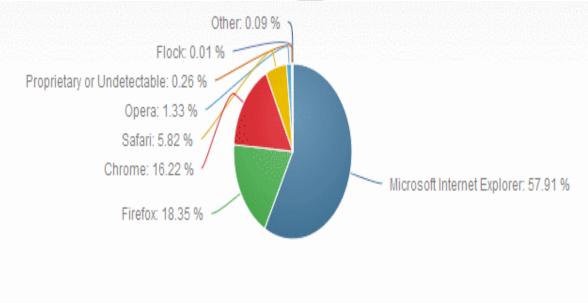
Leveraging Globus high performance data transfer, sharing, and synchronizationservices, we have developed an integrated suite of data management capabilities that areboth easy to use and reliable. Globus transfer supports reliable third-party data transferbetween remote endpoints. It is designed to scale to huge data sizes and includes sophisticatedfunctionality to manage transfers on behalf of users. For example, Globus managessecurity configurations, handles authentication with participating endpoints, tunessettings to optimize bandwidth, provides automatic fault recovery, guarantees integrityvia checksums, and notifies users of errors as they occur. Using Globus, researchers are able to automatically and manually move data as they are generated from acquisition tolarge storage systems and to transfer partial datasets to analysis resources from whichautomated analysis procedures can provide immediate feedback. Users can also use thissystem to share large datasets with collaborators or move data back to their home institutionor laptop without needing to physically transport hard drives.

We are applying these developments in the context of analyzing APS data collectedfor experiments in single crystal diffuse scattering, high energy diffraction microscopy,and X-ray micro tomography. This work has involved modifying several APS detectorcomputer stations to allow for real-time streaming of data to a remote computing platform.The network capacity at the APS has proved to be sufficient for these early experiments,after minor adjustments to detector configurations—for example, to improvedata buffering when network bottlenecks cause momentary delays in data offload. Theresulting systems can move raw image data to a high-capacity, redundant storage systemwithin seconds after acquisition.We have integrated these technologies with other systems components that are engagedin a typical data lifecycle, such as visualization codes, analysis scripts, and datamanagement capabilities. In the following, we briefly describe how this integration isperformed in this case of one existing visualization and analysis tool.

PROPOSED SOLUTION

Several regional focus groups consisting of major employers from New Bern, Greenville, andWilmington, North Carolina [9] disclosed that students needed more training with MS Officeproductivity

tools as well as having an overall awareness of technology and its appropriate use in the
Workplace. In addition, the authors of this paper attended many Tillman School of Business meetingswhereby student deficiencies or misconduct (e.g., not knowing how to use basic application software,the lack of knowledge in end-user computer security, using the internet to plagiarize, etc.) werediscussed. Further, the IT Department at the University of Mount Olive was engaged regularly inalerting faculty, staff, and students via email, about compromising passwords and phishing schemes.Just as business organizations are increasingly requiring their members to undergo annual or semiannualPC-based ethical and security awareness training, educational institutions may wish toconsider emulating this for their staff, faculty and students on the topic of personal computer end usersecurity best practices. The MIS program at the University of Mount Olive addressed the challenge oftechnology/business computer security literacy by implementing a new e-learning solution to augmenta traditional course on the topic of computer security. The e-learning solution consists of acustomized, self-paced, web-based end user digital security awareness tutorial.



It is common at light source beamlines for the stream of raw images (usually formatted as TIFF or Crystallographic Binary File: CBF) generated by a detector to be merged,
once collected on stable storage, to form a single Hierarchical Data Format version (HDF5) dataset structured according to the NeXus data format . The resultant NeXus raw data file, once in place on the remote server, can be immediately accessed by scientiststo perform checks on its quality and to determine experimental parameters suchas orientation. We enable this access through a Python-based graphical user interface(GUI), NeXpy. This GUI application runs on a researcher’s computer, either a beamlinecomputer or a personal laptop; we enable online access from NeXpy to the remotedataset(s) via the use of network operations.

Demographics

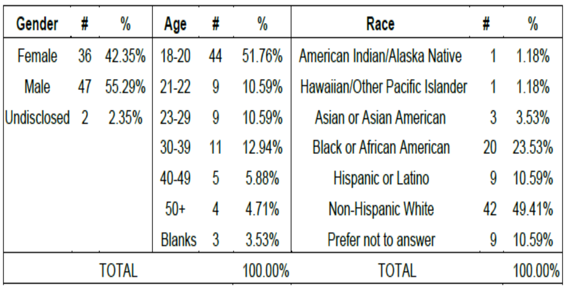
A voluntary exit survey was completed by students to determine their perceived efficacy of the onlinesecurity tutorial on computer security topics. 85 respondents participated in the survey study. Thesample was fairly evenly split across males and females. Furthermore, more than half of the sampleconsisted of traditional students in the 18-22 age bracket. Finally, roughly half of the samplerespondents identified as White whereas a quarter of the sample respondents identified as Black or

African American. The sample was representative of the larger population of typical college studentsin a four-year degree programs. Table 1 below provided more details of the sample demographics.

Multiscale data management

Experimental data management spans multiple timescales and operation modes: catalogingand checking incoming data, performing analysis and tracking progress, and long-termdata sharing and publication. Catalogs for data and metadataA single APS experiment session can produce many thousands of files, and reconstructionand analysis tasks invariably produce yet more files. Keeping track of this data and its location(s) is frequently expensive and error prone. To streamline this process, we adopta digital asset management approach [5,29] built on Globus Catalog, a cloud service formanaging user-defined catalogs. In this service, catalogs contain one or more datasets;datasets contain one or more members (files and directories accessible via Globus transfer);and metadata (typed key-value annotations, or tags) can be associated with bothdatasets and members. Catalogs are created and managed by users and can be used forany purpose (e.g., within a collaborative research project; for annotating and bundlingdata from a beamline; or for personal use). Within a catalog, users can create nameddatasets, associate data members with a dataset, and associate arbitrary metadata withdatasets and members. Once a catalog has been created and populated, owners can defineaccess permissions with respect to the users and groups that can read and write it.

Table 1: Demographic data for the sample



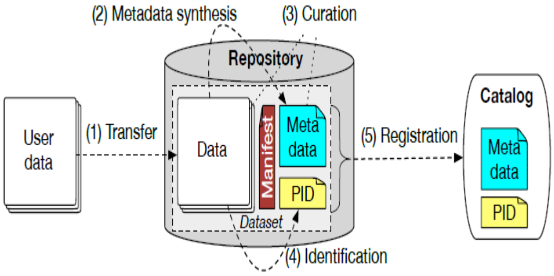
Gender	#	%	Age	#	%	Race	#	%
Female	36	42.35%	18-20	44	51.76%	American Indian/Alaska Native	1	1.18%
Male	47	55.29%	21-22	9	10.59%	Hawaiian/Other Pacific Islander	1	1.18%
Undisclosed	2	2.35%	23-29	9	10.59%	Asian or Asian American	3	3.53%
			30-39	11	12.94%	Black or African American	20	23.53%
			40-49	5	5.88%	Hispanic or Latino	9	10.59%
			50+	4	4.71%	Non-Hispanic White	42	49.41%
			Blanks	3	3.53%	Prefer not to answer	9	10.59%
TOTAL					100.00%	TOTAL		100.00%

Datasets allow users to organize heterogeneous and disparate data elements (files,directories, URLs, and even other datasets)into content-specific collections. We leveragethis capability to raise the level of abstraction by which users may express common datalifecycle operations, so that, for example, a researcher can “analyze dataset A” ratherthan (as is commonly the case today) “analyze the 1,500 files that were collected duringthe second experiment, which are distinguished from those obtained in other experimentsby their file names.” This abstraction allows dataset interactions (e.g., browsing, search,discovery, inspection of data integrity, analysis, movement) to be aligned with scientificinvestigation rather than physical storage methods.

Globus Catalog’s query interface allows users to discover, screen, and retrievedatasets based on a broad range of search criteria, from subject type to scientifically relevanttyped metadata, such as material composition and energy density among manyothers. The query interface also supports faceted search [43,35], an intuitive techniquefor exploring classified or categorized information. Faceted search is especially valuablewhen exploring large volumes of heterogeneous data, as it provides both an initial summaryof the data as well as a means to “drill down” into the data by applying increasinglyspecific filters, each resulting in a new summary of the data.

Data publication

Challenges relating to the publication of materials-related data and metadata,an essential task if we are to enable experimental and simulation data created for onepurpose to be reused in other contexts, for example via the knowledge base depicted inFigure 1. The importance of this task has been recognized by the US Materials GenomeInitiative, which includes as one of its four goals [25]: "Making digital data accessibleincluding combining data from experiment and computation into a searchable materialsdata infrastructure and encouraging researchers to make their data available to others."Historically, materials data exchange has focused on carefully and repeatedly validateddata such as standard reference data. Such data have been captured in computerizeddatabases since the 1970s. In addition to compilations of experimentaldata , there are extensive efforts to create repositories of computed properties,such as crystal structure parameters and formation enthalpies for binary alloys , themany data collected in the the Computational Materials Repository , MaterialsProject , Aflowlib.org ; and the NIST repositories . But, all focus on organizingand curating large numbers of derived materials properties, whether experimentalor computed. However, at present there are few options for publishing large file-basedmaterials datasets that may be required to reproduce these derived values. Data publication is a multi-stage process that encompasses the transfer of data to Persistent storage; assigning a persistent identifier (PID) that can be used to refer to itsubsequently without ambiguity and registering it in catalogs for discovery.



The publication process, showing the distinct transfer, metadata extraction, and curation, identification,and registration steps involved in publishing a dataset. A complete dataset comprises a set of files, associatedmetadata, a PID, and a manifest that allows us to easily verify dataset contents.

Datasets have an associated landing page—a web page from which the dataset can be accessed—that is referenced by a URL. Datasets are discoverable, based on metadata, via the data publication search interface. The PID and selected metadata may also be Published in external catalogs, such as DataCite, so that users can discover and thenresolve a PID to obtain a reference (URL) to the dataset's landing page. Landing pagescan also be indexed by web search engines such as Google, so that users can discoverdatasets by searching on public metadata, in the same way that they find web pages.

The publication service manages the entire data publication lifecycle from submission through curation, embargo, sharing, and finally access. Its self-service administrationinterface allow authorized administrators to create communities and collections withassociated policies. For example, administrators can specify which users can performactions on a collection (e.g., submission, curation, administration); define the submissionand curation workflows to be used for a collection; choose what storage repositoryis be used for storing a collection's data; and select what PID providers are to beused. The ability to specify publication workflows is particularly important, as it supportsmany different publication models, such as open access publishing and curation-basedapproval, and permits these models to be tailored to the requirements of different groupsand institutions.

Accelerating data analysis

The next step in the experimental process involves data modeling using a range of computationaltechniques, from advanced statistical analysis of data correlations to comprehensivesimulations using molecular dynamics and ab initio modeling. As noted above,the considerable computational cost of advanced analysis methods, plus the increasingsize of datasets, results in a considerable computational bottleneck at many APSbeamlines. To address such bottlenecks, we and our colleagues have explored methodsfor high-performance parallel execution of reconstruction and analysis methods used insuch fields as micro tomography, diffuse scattering, and high-energy diffraction microscopy. These methods can run both on high-performance clusters, with 100s to1000s of cores, and on Argonne's IBM BG/Q supercomputer, with 800,000 cores. Thechoice of platform depends on the computational cost of the problem and its responsetime requirements. We report here on some results obtained in diffuse scattering.

Diffuse scattering

Single crystal diffuse scattering is a powerful method of determining short-range order within crystalline solids. Comprehensive measurements of single crystal diffusescattering over a wide volume of reciprocal space can provide detailed insights into thenanoscale disorder that underlie technologically important materials properties, such asfast-ion conduction, relaxer ferroelectricity, colossal magnetoresistance, and unconventionalsuperconductivity.The experiments that generate the data analyzed here use a Dectris Pilatus series detector,which make it possible to measure diffuse scattering using continuous rotations ofthe sample without opening and closing the shutter between exposures, collecting complete360_ data sets in steps of 0.1_ in under ten minutes. Thus, we can obtain volumes ofreciprocal space coverage, about 20 to 30GB in size, with high angular resolution in allthree dimensions. The rapid data acquisition can be used either to improve statistical accuracythrough repeated rotations or to collect fine-grained parameterized data, e.g., as afunction of temperature. The Pilatus detectors have essentially zero background and highenough dynamic range (106) to allow the simultaneous measurements of Bragg peak anddiffuse scattering intensities.Detector images, which were collected at a frame rate of 10 Hz or 5 Hz dependingon the rotation speed, were triggered by a signal emitted by SPEC, the instrument controlprogram used to control the rotation motors. The images were then streamed usingGlobus to a data repository on the BG/Q supercomputer. In the case of the Pilatus measurements,a few frames were lost because the Globus transfers used too much bandwidthand prevented the detector from transferring images from its memory buffers. This is aproblem that can be solved by increasing detector memory.In two sets of experiments, we collected 25 TB and 14 TB, respectively, and measuredsix samples in each case with varying compositions, collecting data on each oneat 40 to 60 temperatures. The experiments therefore demonstrated that our goal of beingable to measure complete phase diagrams in less than a week is perfectly feasible.

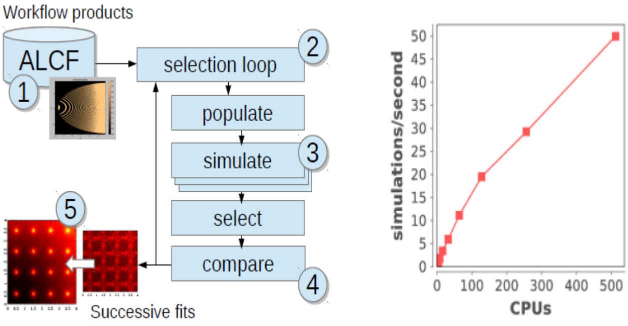
Use of Swift parallel scripting

One distinguishing feature of our work is our use of a high-level parallel scripting language,Swift, to implement the parallel structure of our applications. The ease withwhich Swift allows programmers to define large-scale parallel programs via the compositionof existing sequential procedures has proved highly beneficial; furthermore, theSwift/T implementation allows us to run on large parallel computers such as thoseat Argonne Leadership Computing Facility (ALCF). Swift/T produces an MPI programfrom the input script, and is scalable to the largest HPC systems.We use Swift in both the Nexus assembly workflow, to organize the large collectionof concurrent data manipulation tasks, and in our implementation of the diffuse scatteringreconstruction algorithm (see x4), to perform large-scale parallel computing. Inthe parallel version of the Crystal Coordinate Transformation Workflow (CCTW), weleverage Swift/T's ability to call directly to C++ methods and Script functions, whichallowed us to produce what is essentially an MPI version of CCTW. Finally, in our implementationof parallel evolutionary optimization for diffuse scattering (see x4.4), weleverage Swift/T's ability to call Python functions directly. These functions are wrappedaround the internal DISCUS FORTRAN functions through the use of the F2PY Fortran toPython interface generator, making for a hierarchical scripted programming modelaround pre-existing native code components.

Experiment-time data analysis

We next describe work that we have undertaken to provide visual data analysis results tousers while using the beamline. This processing pipeline provides the user with visual experimentalresults in reciprocal space and real space, and results from inverse simulationand Bragg peak analysis.. This data is transferred to ALCF resources for stable storage 2 visualizable real space NeXus file and produces inputs for further processing—inversesimulation-based modeling 8 and Bragg peak modeling 9 . Implemented as a Swiftscript, it runs automatically on a parallel cluster as data is ingested, and is capable ofusing the whole 100-node cluster, concurrently transforming one dataset per node.CCTW is the new transformation code developed for this project. It is a nearly-allnewC++ code that operates on NeXus or other HDF5 datasets. CCTW may be calledin an an automated manner as part of the pipeline. Additionally, the C++ interfaces areexposed to Swift, allowing the parallelization of CCTW itself—a feature that will becritical for real-time experiment calibration, etc., as the first visualization in a run must bedone quickly (in less than 10 minutes). As of early 2015, we have collected and processedabout 50 TB of data.and processing. The raw data is tagged in the Globus catalog 3 , along with pipeline outputsas they are produced. Then, multiple components operate on the data. If necessary,the detector background signal is subtracted from the data 4 . The raw image files aremerged into large NeXus files, which are visualizable in NeXpy5 . Then, the maximalpeak and other peaks are discovered in the data 6 . The data is transformed into realspace via CCTW 7 , which runs as a subcomputation.

Evolutionary optimization for simulation-based inverse modelingWe have also implemented a diffuse scattering evolutionary algorithm using DISCUSas the simulator and a Swift-based, concurrent evolutionary selection loop based on theDIFFEV code and algorithm. This algorithm starts with a proposed crystal structure,which it perturbs to create a population of potential structures. For each potentialstructure, the algorithm uses the DISCUS code to generate a simulated scattering image.That image is compared with the experimental image; if the two are “close enough,”the algorithm terminates; otherwise, it continues. This method has proved effective forestimating the structure of some disordered crystals.The process starts with thereal space experimental data produced by CCTW 1 . In each iteration of the selectionloop 2 , multiple approximations are created to form a population of potential crystalstructures. Each structure is run independently in DISCUS, providing a great deal ofavailable concurrency 3 .the Swift implementation can exploitthis concurrency and use up to 512 cores, enabling many more simulations to completeper second. The best fits are selected and compared to the experimental data 4 andvisualizations are produced 5 .Computationally, the DIFFEV algorithm is an exciting use case for the applicationof HPC for diffuse scattering. While we have run up to 512 concurrent simulationsin the evolutionary population, the method can benefit from at least 5000 concurrentsimulations. Furthermore, we have incorporated OpenMP support into a performancecriticalDISCUS method, allowing each simulation to use _20 threads. Thus, we can use_100,000 cores concurrently on a BG/Q or comparable supercomputer, such as thoselinked with the XSEDE network.



a) Algorithm schematic b) Scalability

Figure 6. Evolutionary inverse model based on DISCUS simulation

Other projects at Argonne and elsewhere have also demonstrated the value of highperformancecomputing as a discovery accelerator in photon sciences, for micro tomography[6,13], diffraction [22,36], high energy diffraction microscopy (HEDM) [21], grazingincidence small angle scattering (GISAXS) [9], and other imaging modalities.

References

[1] NIST Repositories of Materials Data and Models. http://nist.matdl.org. Web site. Accessed:2014-11-29.

[2] ALLEN, B., BRESNAHAN, J., CHILDERS, L., FOSTER, I., KANDASWAMY, G., KETTIMUTHU, R.,KORDAS, J., LINK, M., MARTIN, S., PICKETT, K., AND TUECKE, S. Software as a service for datascientists. Communications of the ACM 55, 2 (2012), 81–88.

[3] ANANTHAKRISHNAN, R., CHARD, K., FOSTER, I., AND TUECKE, S. Globus platform-as-a-servicefor collaborative science applications. Concurrency and Computation: Practice and Experience 27, 2(2015), 290–305.

[4] ARMSTRONG, T. G., WOZNIAK, J. M., WILDE, M., AND FOSTER, I. T. Compiler techniques formassively scalable implicit task parallelism. In Proc. SC (2014).

[5] AUSTERBERRY, D. Digital Asset Management. Taylor & Francis, 2012.

[6] BICER, T., G¨U RSOY, D., KETTIMUTHU, R., DE CARLO, F., AGRAWAL, G., AND FOSTER, I. Rapid tomographic image reconstruction via large-scale parallelization. In 21st International European Conferenceon Parallel and Distributed Computing (Euro-Par 2015), (Vienna, Austria, 2015).

[7] CHARD, K., LIDMAN, M., BRYAN, J., HOWE, T., MCCOLLAM, B., ANANTHAKRISHNAN, R.,TUECKE, S., AND FOSTER, I. Globus Nexus: Research identity, profile, and group management as aservice. In IEEE 10th International Conference on e-Science (Oct 2014), vol. 1, pp. 31–38.

[8] CHARD, K., PRUYNE, J., BLAISZIK, B., ANANTHAKRISHNAN, R., TUECKE, S., AND FOSTER, I.Globus data publication as a service: Lowering barriers to reproducible science. In 11th IEEE InternationalConference on eScience (Munich, Germany, 2015).

[9] CHOUROU, S. T., SARJE, A., LI, X. S., CHAN, E. R., AND HEXEMER, A. HipGISAXS: a high-performancecomputing code for simulating grazing-incidence X-ray scattering data. Journal of AppliedCrystallography 46, 6 (2013), 1781–1795.

[10] CURTAROLO, S., SETYAWAN, W., WANG, S., XUE, J., YANG, K., TAYLOR, R. H., NELSON, L. J.,HART, G. L. W., SANVITO, S., AND BUONGIORNO-NARDELLI, M. AFLOWLIB.ORG: A distributed materials properties repository from high-throughput ab initio calculations. Computational MaterialsScience 58 (2012), 227–235.

[11] DAGOTTO, E. Complexity in strongly correlated electronic systems. Science 309, 5732 (2005), 257–262.

[12] DEDERICHS, P. The theory of diffuse X-ray scattering and its application to the study of point defectsand their clusters. Journal of Physics F: Metal Physics 3, 2 (1973), 471.

[13] DESLIPPE, J., ESSIARI, A., PATTON, S. J., SAMAK, T., TULL, C. E., HEXEMER, A., KUMAR, D.,PARKINSON, D., AND STEWART, P. Workflow management for real-time analysis of light source experiments.In Proceedings of the 9th Workshop on Workflows in Support of Large-Scale Science (2014),IEEE Press, pp. 31–40.

[14] FISCHER, C. C., TIBBETTS, K. J., MORGAN, D., AND CEDER, G. Predicting crystal structure bymerging data

mining with quantum mechanics. Nature materials 5, 8 (2006), 641–646.

[15] GEHRKE, R., KOPMANN, A., WINTERSBERGER, E., AND BECKMANN, F. The high data rate processingand analysis initiative of the Helmholtz Association in Germany. Synchrotron Radiation News28, 2 (2015), 36–42.

[16] GUNTER, D., CHOLIA, S., JAIN, A., KOCHER, M., PERSSON, K., RAMAKRISHNAN, L., ONG, S. P.,AND CEDER, G. Community accessible data store of high-throughput calculations: Experiences fromthe Materials Project. In SC12: High Performance Computing, Networking, Storage and Analysis (Year), IEEE, pp. 1244–1251.

[17] INTERNATIONAL COUNCIL FOR SCIENCE : COMMITTEE ON DATA FOR SCIENCE AND TECHNOLOGY.Data citation standards and practices.

[18] Lowe, S. (2005), Responding to learner needs in distance education: providing academic andrelational support (PARS). In J. Levine, (Ed). Making distance education work: understanding learning and learners at a distance. Okemos, Michigan: Learner Asociates.net, LLC.

[19] Cisco. (2011). Cisco 2011 annual security report. Retrieved Dec 14, 2011, from:

[20] McLeish, S. (June 2009). Enterprise Plans For Productivity Tools: Holding Out For MicrosoftOffice 2010. Cambridge, MA, Forrester Research.

[21] Murray, J. (2003). What is “contemporary literacy?” *Multimedia Schools*, *10*(2), 15-18.

[22] Proserpio, Luigi. &Gioia, Dennis. (2007). Teaching the Virtual Generation. *Academy ofManagement Learning and Education*, 2007, Vol. 6, No. 1, 69-90.

[23] Mitchell, M., Sheard, J., & Markham, S. (2000). Student motivation and positive impressions ofcomputing subjects. *Proceedings of the Australasian Conference on Computing Education.*

[24] Ratliff, Victoria. (2009). Are College Students Prepared for a Technology-Rich LearningEnvironment? *Journal of Online Learning and Teaching*, Vol. 5, No. 4. Retrieved July 3, 2010

[25] Kaminski, Karen.,Seel, Pete, & Cullen, Kevin. (2003). Technology Literate Students? Resultsfrom a Survey. *Educause Quarterly*, Number 3, 2003, 34-40.

[26] MOC Focus Groups, 2010. Mount Olive College, Mount Olive, NC.

[27] Andersson, D., Reimers, K., Maxwell, S. (2013).Utilizing the Web to Enhance StudentEngagement and Achievement. INTED2013 (International Technology, Education andDevelopment Conference)Valencia, Spain

[28] Andersson, D. &Reimers, K. (2012). Post-Secondary End User Security: Addressing the EndUser Challenge. EDULEARN2012 (International Conference on Education and New, LearningTechnologies) Barcelona, Spain.

[29] Reimers, K., & Jordan, E. (2012). Innovative Education: Operational teamwork producessuccessful capstone experience. Institute for operations research management, 39.

Retrieved from: http://www.informs.org/ORMS-Today/Public-Articles/August-Volume-39-Number-4/Innovative-Education-Operational-teamwork-produces-successful-capstoneexperience.