RESEARCH ARTICLE | JULY 27 2016

Real-time data-intensive computing

Dilworth Y. Parkinson; Keith Beattie; Xian Chen; Joaquin Correa; Eli Dart; Benedikt J. Daurer; Jack R. Deslippe; Alexander Hexemer; Harinarayan Krishnan; Alastair A. MacDowell; Filipe R. N. C. Maia; Stefano Marchesini; Howard A. Padmore; Simon J. Patton; Talita Perciano; James A. Sethian; David Shapiro; Rune Stromsness; Nobumichi Tamura; Brian L. Tierney; Craig E. Tull; Daniela Ushizima

AIP Conf. Proc. 1741, 050001 (2016)
https://doi.org/10.1063/1.4952921

Real-Time Data-Intensive Computing

Dilworth Y. Parkinson[1,a], Keith Beattie[2], Xian Chen[1], Joaquin Correa[3], Eli Dart[4], Benedikt J. Daurer[6], Jack R. Deslippe[3], Alexander Hexemer[1], Harinarayan Krishnan[2], Alastair A. MacDowell[1], Filipe R.N.C. Maia[6], Stefano Marchesini[1,2], Howard A. Padmore[1], Simon J. Patton[2], Talita Perciano[2], James A. Sethian[2,5], David Shapiro[1], Rune Stromsness[2], Nobumichi Tamura[1], Brian L. Tierney[4], Craig E. Tull[2], Daniela Ushizima[2]

[1]*Advanced Light Source, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA*
[2]*Computational Research Division, Lawrence Berkeley National Laboratory Berkeley CA 94720*
[3]*National Energy Research Scientific Computing Center, Berkeley, CA 94720*
[4]*Energy Sciences Network, Berkeley, CA 94720*
[5]*Department of Mathematics, UC Berkeley, Berkeley CA 94720*
[6]*Uppsala University, Uppsala, Sweden*

[a]Corresponding author: dyparkinson@lbl.gov

Abstract. Today users visit synchrotrons as sources of understanding and discovery—not as sources of just light, and not as sources of data. To achieve this, the synchrotron facilities frequently provide not just light but often the entire end station and increasingly, advanced computational facilities that can reduce terabytes of data into a form that can reveal a new key insight. The Advanced Light Source (ALS) has partnered with high performance computing, fast networking, and applied mathematics groups to create a "super-facility", giving users simultaneous access to the experimental, computational, and algorithmic resources to make this possible. This combination forms an efficient closed loop, where data—despite its high rate and volume—is transferred and processed immediately and automatically on appropriate computing resources, and results are extracted, visualized, and presented to users or to the experimental control system, both to provide immediate insight and to guide decisions about subsequent experiments during beamtime. We will describe our work at the ALS ptychography, scattering, micro-diffraction, and micro-tomography beamlines.

MOTIVATION FOR REAL-TIME DATA INTENSIVE COMPUTING

Evolution of synchrotron user facilities

Throughout their history, synchrotrons user facilities have continuously expanded their capabilities. In the 1980s, many beamlines simply delivered X-rays to users. In the 1990s, facilities increasingly provided end stations and additional equipment to users. With improved storage rings, optics, and detectors, both the quality and rate of data collected have continuously improved since that time. Along with these improvements, since 2000, the trend has been to increasingly provide computing and analysis capabilities to users. An important motivation for this evolution is the desired expansion of the user base. A large fraction of users today would not be able to use a synchrotron that only provided X-rays, or X-rays plus end stations. Users bring their samples and their expertise in their own area of science, and they rely on the hardware, computing and software infrastructure built by the facilities, along with the expertise provided by synchrotron staff to take advantage of the infrastructure. Providing these capabilities has allowed synchrotrons to vastly expand their influence and use across many science domains. Increasingly it will be improvements in software, computing, and data analysis that will have the greatest impact on the number and variety of users those synchrotrons can serve.

Proceedings of the 12th International Conference on Synchrotron Radiation Instrumentation – SRI2015
AIP Conf. Proc. 1741, 050001-1–050001-6; doi: 10.1063/1.4952921
Published by AIP Publishing. 978-0-7354-1398-6/$30.00

It is important to remember that users visit synchrotrons as sources of understanding and discovery—not as sources of light, and not as sources of data. In many cases, a user may take X-rays for a number of hours, and may collect terabytes of data, but at the end of the experiment, what they really want is one key new insight into the problem they are investigating. Synchrotron facilities must keep that end goal in mind as they develop tools and capabilities to help users along that path.

Approaches to Computing and Relative Costs

The importance of computing is emphasized by the trends shown in Figure 1. Over the years, an increasing fraction of the total time spent by scientists on an experiment is spent on data management; meanwhile the relative amount of time spent on data collection continues to shrink (data compiled by Kevin Mader, 4Quant,[1]).

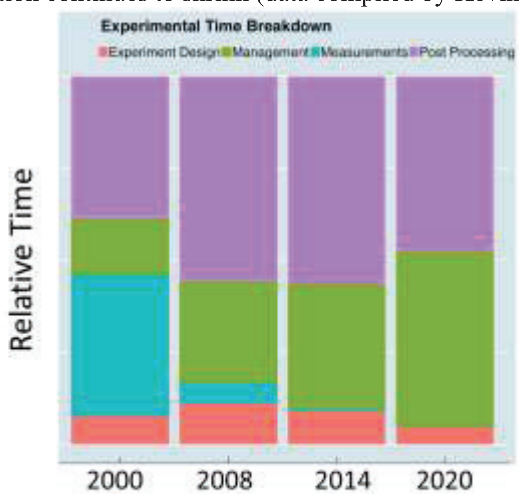

FIGURE 1. Trends in relative time spent on different parts of the scientific process: Experimental design, Data collection (measurement), Data management, and Post processing.[1]

Part of the reason for this is that improvements in computing at synchrotron facilities often lag behind other improvements. Despite significant increases in data rates, many users are still using what could be described as “small data” approaches for dealing with their data: copying data sets back and forth by hand and analyzing one data set at a time, often within a graphical user interface that requires a lot of mouse clicks or keyboard interactions for every data set. More users are moving to a “medium data” approach, which combines some scripting and automation, but often still based on software designed for a “small data” approach. However, the “medium data” approach does not scale well beyond a few tens of samples—data management is totally manual, and in many cases this approach stitches together multiple software packages, in an ad hoc way. If a change is made in one part of the workflow, everything after that point in the workflow must be repeated.

A “big data” approach would use a framework built specifically for large data sizes, built to run on large computing infrastructure. True “big data” approaches scale very well to large data sizes, but they do require more up-front development costs. Figure 2 shows the “cost per sample” and “total cost” for computing when using small, medium, and big data approaches, as a function of the number of samples to analyze. This data is based on experience especially in microtomography analysis projects (data compiled by Kevin Mader, 4Quant, [1]). Many users have 1-50 samples, so they can minimize their total cost using the small or the medium data approach. Thus most individual users cannot justify investing their time and resources in developing a Big Data approach. However, from the perspective of a user facility, if a big data framework that benefits a group of users can be developed, the overall computing costs for the group can be reduced dramatically.

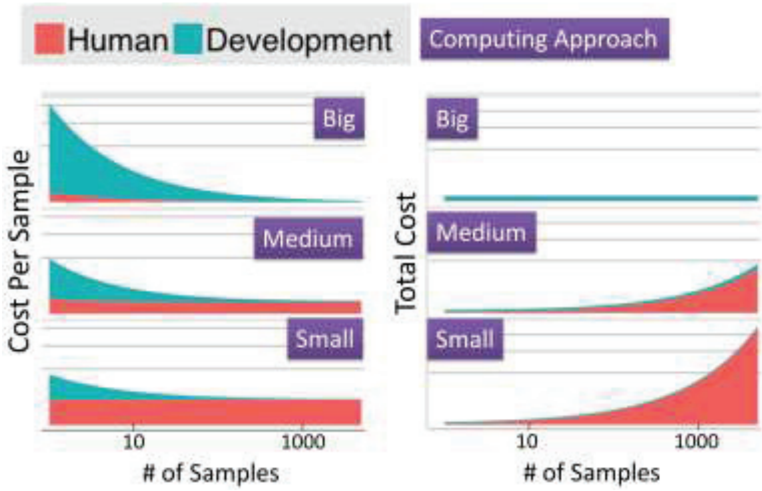

FIGURE 2. Trends in time/cost spent on different parts of the scientific data analysis process for "small", "medium" and "big" data approaches. [1]

Scenarios Requiring Real-Time and/or Data-Intensive Computing

Although most users eventually process their data—one benefit to providing real-time data processing capability is that users walk away at the end of their beamtime with that step taken care of . The alternative has been memorably described by Peter Zwart of Berkeley Lab:- "Not processing your data in real time is the first step in not processing your data at all". Over 20% of users at the Advanced Light Source are first-time users, and these users often don't have appropriate computers or software (or expertise) for performing the relevant data processing if it does not happen while they are at the facility. Of course, this is especially true for users who collect large amounts of data.

A stronger justification for real-time computing is for in situ or time-resolved experiments where feedback based on the data is required to control the experiment. Many users need to know when to change sample environment parameters (like temperature or pressure or flow rate, among many others), and rely on looking at results of real-time processing to make these decisions. In other cases, the acquisition software is coupled to the results of data processing to automate these adjustments. An extension of this case is the "digital twin,"[2] in which initial measurement are used to create a digital model. Simulations based on that model are then used to predict interesting time points or regions in the sample, and experimental parameters are adjusted based on this guidance. This can be useful when trying to measure rare and hard-to-catch time points or regions of interest. A final case in which data-intensive computing is increasingly used is when data quality is very low. Rather than just throw out samples that yield poor data, additional data is collected on these samples. This data then requires advanced mathematical and statistical techniques to sort through the over-determined problem and the ambiguous signal, to extract the relevant information.

THE ADVANCED LIGHT SOURCE AS PART OF A SUPER-FACILITY

Our experience at the Advanced Light Source (ALS) is that to build a true Super-Facility requires collaboration. We have worked closely with our local networking groups, LBLnet at Berkeley Lab and ESnet[3] to develop a sophisticated network infrastructure based on the Science DMZ[4] model (DMZ=demilitarized zone). In this model, security roadblocks between two trusted resources can be removed so as to improve the speed with which data can be transferred between them. The ESnet vision is that "scientific progress will be completely unconstrained by the physical location of instruments, people, computational resources, or data," and we have taken advantage of this.

The ALS has forged a close partnership with the National Energy Research Scientific Computing Center (NERSC, nersc.gov). This facility provides significant cutting-edge computing power, data storage, archiving facilities, software and expertise that make these resources readily available to us and to users. The ALS has developed a significant collaboration with the Center for Applied Mathematics for Energy Research (CAMERA, camera.lbl.gov), an integrated cross-disciplinary center at Berkeley Lab aimed at inventing, developing, and delivering new mathematics required to capitalize on experiments at user facilities. The ALS has been collaborating with the Berkeley Lab Computational Research Division to develop SPOT Suite which is a set of tools to provide ALS scientists access to best-of-breed data management, analysis, and simulation tools.[3,5] Users access SPOT through a web portal, where they can manage and share their experimental and simulation data, analyze experimental data, and view real-time results during beamtime. In each of the examples below, these building blocks are combined in different ways to deliver real-time and/or data intensive computing to users.

Example 1: Microtomography Data Management and Image Processing

At the microtomography beamline of the ALS, all data sets are automatically packaged with their metadata in HDF5 files and transferred over the 10G Ethernet Science DMZ network from the beamline to NERSC. Data and metadata are added to a MongoDB database. SPOT Suite, which manages this data transfer and warehousing, also detects relevant parameters in the metadata and then automatically launches the relevant preprocessing and fast analytic tomographic reconstruction.[3,5] Jobs are submitted to NERSC with a RabbitMQ worker approach to reduce the impact of queue wait times. After tomographic reconstruction, results are immediately added to the database, which is then viewable by users through a web portal. In addition to a standard interface in which digital slices through the 3D reconstructed volume can be viewed, a new viewer developed through a collaboration between SPOT and CAMERA allows viewing volume-renderings of the data. This viewer uses VisIt to render the data on NERSC, and passes the rendered image result to the browser. This allows users to get 3D volume renderings of data sets of any size (many gigabytes is a standard size of one reconstruction) without having to first download the data or have access to other rendering software. In addition, this viewer allows collaborative viewing, through which a user can generate a weblink to send to a collaborator so that they can view a data set together. Figure 3 shows a screenshot of this 3D viewer. Finally, the web interface allows users to submit a limited number of data sets for iterative tomographic reconstruction (MBIR), which is computationally expensive but provides superior results. Besides being able to give fast automatic feedback, another advantage to working with a high performance computing center is being able to give users seamless access to advanced algorithms which requires significant amounts of computer power.

The microtomography beamline has also been collaborating with CAMERA to deliver additional image processing and analysis tools to users. One of these tools in use now is called “F3D”, which provides a suite of filtering and image processing tools that are optimized for 3D volumes and take advantage of GPUs, and which run orders of magnitude faster than the unoptimized versions of the tools. These have been delivered as FIJI/ImageJ plugins, because users are familiar with this software and can easily add those filters to their workflows. A more recent capability CAMERA is developing is for fiber and crack analysis, which has been applied to ceramic matrix composites. [6,7]

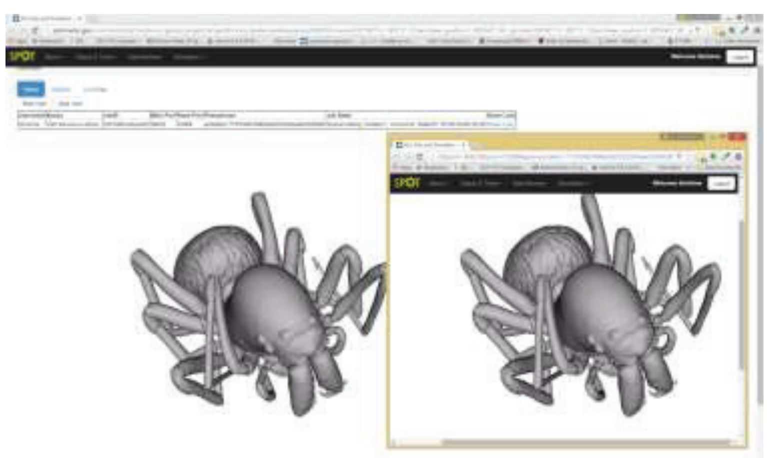

FIGURE 3. Screenshot of 3D web viewer; volume is rendered on a high performance computer, and the resulting image is passed to the web client. Screenshot courtesy H Krishnan.

Example 2: High Performance Scattering Simulations During Beamtime

HipGISAXS is a High Performance scattering simulation code developed by the ALS and CAMERA, and which is optimized to run on GPUs. Last year the ALS Small Angle Scattering (SAXS) beamline 7.3.3 performed a demonstration experiment in which they collected time-resolved data and took advantage of SPOT Suite and Globus tools, to launch and manage data preprocessing jobs at NERSC, and then HipGISAXS simulations on 8000 GPU nodes on the Titan system of the Oak Ridge Leadership Computing Facility at Oak Ridge National Laboratory. Results were presented to users through a web interface within minutes, allowing better use of beamtime.

Example 3: Streaming Data for Immediate Ptychography Reconstruction

Ptychographic reconstruction requires significant computing power and smart algorithms.[8–10] ALS and CAMERA have developed a streaming system, in which data is streamed from the acquisition system to a 16-GPU cluster and reconstructed using the ALS/CAMERA package SHARP[11]. This system has been sufficiently optimized that the ptychographically reconstructed image is updated with new information within seconds of a region being scanned. This allows users to know if changes or optimizations need to be made to the experiment, and thus make much better use of beamtime.

Lessons Learned

We have found that one size of computing does not fit all. Different beamlines produce different rates and volumes of data, they define "real-time" in different ways, and the average "user" at each beamline has a different skill set. We have been exploring a number of ways to improve computing, and most of them have both pros and cons. For example, using the computing resources of a high performance computing center like NERSC means that we can quickly run very ambitious and computationally intensive analysis and simulations; in addition, we do not have to manage the computer hardware on site at the ALS; finally, we can take advantage of their massive amounts of disk and storage space as we consider how to deal with data preservation and curation. However, using these facilities also means that we do not have complete control over the resources and are subject to the facility's schedule, downtimes and policies -(such as queue wait times or the need to schedule computing time.. We believe that the best solution for us will be a combination of computing resources physically close to the beamline and those located at computing facilities but connected through fast networks.

WORLDWIDE TRENDS

Facilities throughout the world are increasingly focused on providing new computing tools to their users. A recent issue of Synchrotron Radiation News[12] highlighted a few of these efforts, and we mention some work here. We will mention here only developments relating to tomography beamlines, although this represents only a small fraction of the significant work being done in this area. The TOMCAT beamline at the SLS is pushing the boundaries of data rates from detectors, with development of their giga-FROST detector which provides *continuous* acquisition at up to 8 GB/s. This enormous data rate requires significant developments in the data management and processing pipeline to keep up with this data collection rate[13] and they have created a spin-off called 4Quant to further develop this pipeline.[1] At the APS, TomoPy[14], a collaborative framework for analysis of synchrotron tomography data, has been developed; recently, researchers from the the iMinds Vision Lab at the University of Antwerp, who develop ASTRA, have worked together with the developers of TomoPy to provide tomographic reconstruction with GPU acceleration, greatly increasing the speed of reconstruction and the possibility for real-time feedback. Workers at the Karlruhe Institute of Technology have developed Ultra-fast X-ray Imaging (UFO),[15] which includes, among other things, accelerated tomographic reconstruction, high-throughput FPGA-based detector readout, and image-based control loops. The Helmholtz Centers have been working on the High Data Rate Processing and Analysis Initiative (HDRI), which has efforts in data management, real-time processing, and analysis, modeling, and simulation.[16]

Despite excellent work at many places, there is a significant amount of duplication of effort between facilities, and many of the efforts are still not taking a "big data" approach to data analysis. We hope that work at all facilities will be based on efforts to adhere to some standards, whether that be for file formats or in approaches to software design, and that software will be designed using composable components which can be easily shared. We hope to see more effort to collaborate and share software, and that facilities will invest in making software they do develop useful to the community at large. The synchrotron community has always achieved software sharing at some level, but now we have an opportunity to ramp up the sharing to advance the various synchrotron techniques at a worldwide level. As facilities continue to develop capabilities and hopefully develop compatible tools, entirely new scientific possibilities arise. For example, a dream of a number of researchers is for all data and metadata collected at facilities to be archived and then made truly public and searchable after a certain amount of time. This fits in with the increasing trend towards open access journal publications.

ACKNOWLEDGMENTS

We acknowledge the work of many at ALS, LBLnet, ESnet, NERSC, CAMERA, and SPOT Suite who were not included on the author list. We thank Susan James, Karen Fernsler, and the LBNL IT division for their help with the computer systems. The Advanced Light Source is supported by the Director, Office of Science, Office of Basic Energy Sciences, of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. This research used resources of the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

REFERENCES

1. Mader K. 4Quant Website. *http://4quant.com/*. Available at: http://4quant.com/.
2. Tuegel EJ, Ingraffea AR, Eason TG, Spottswood SM. Reengineering aircraft structural life prediction using a digital twin. *Int J Aerosp Eng*. 2011;2011. doi:10.1155/2011/154798.
3. Blair J, Canon RS, Deslippe J, et al. High performance data management and analysis for tomography. In: *SPIE Optical Engineering+ Applications*.; 2014:92121G–92121G.
4. Dart E, Rotman L, Tierney B, Hester M, Zurawski J. The science DMZ: A network design pattern for data-intensive science. *Sci Program*. 2014;22(2):173–185. doi:10.3233/SPR-140382.
5. Deslippe J, Essiari A, Patton SJ, et al. Workflow Management for Real-Time Analysis of Lightsource Experiments. *2014 9th Work Work Support Large-Scale Sci*. 2014:31–40. doi:10.1109/WORKS.2014.9.
6. Ushizima D, Perciano T, Krishnan H, Loring B. Structure Recognition from High Resolution Images of Ceramic Composites. In: *2014 IEEE International Conference on Big Data*.; 2014:683–691.
7. Ushizima D, Parkinson D, Nico P, et al. Statistical segmentation and porosity quantification of 3d x-ray microtomography. In: *SPIE Optical Engineering+ Applications*.; 2011:813502.
8. Marchesini S, Tu Y-C, Wu H-T. Alternating projection, ptychographic imaging and phase synchronization. *Appl Comput Harmon Anal*. 2015;(June). doi:10.1016/j.acha.2015.06.005.
9. Yang C, Qian J, Schirotzek a., Maia F, Marchesini S. Iterative Algorithms for Ptychographic Phase Retrieval. 2011;0:32. Available at: http://arxiv.org/abs/1105.5628.
10. Yu Y-S, Kim C, Shapiro D a., et al. Dependence on Crystal Size of the Nanoscale Chemical Phase Distribution and Fracture in $Li_x FePO_4$. *Nano Lett*. 2015:150618150444005. doi:10.1021/acs.nanolett.5b01314.
11. Sethian J. CAMERA website. Available at: http://camera.lbl.gov/software/sharp_camera_downloads.
12. Hexemer A, Parkinson D, Tull C. Information Technology/Large-Scale Data Handling. *Synchrotron Radiat News*. 2015;28(2):2–3. doi:10.1080/08940886.2015.1013412.
13. Mokso R, Marone F, Irvine S, et al. Advantages of phase retrieval for fast x-ray tomographic microscopy. *J Phys D Appl Phys*. 2013;46(49):494004. doi:10.1088/0022-3727/46/49/494004.
14. Gürsoy D, De Carlo F, Xiao X, Jacobsen C. TomoPy: a framework for the analysis of synchrotron tomographic data. *J Synchrotron Radiat*. 2014;21(5):1188–1193. doi:10.1107/S1600577514013939.
15. Vogelgesang M., Chilingaryan S., Rolo TDS., Kopmann A. UFO: A scalable GPU-based image processing framework for on-line monitoring. *Proc 14th IEEE Int Conf High Perform Comput Commun HPCC-2012 - 9th IEEE Int Conf Embed Softw Syst ICESS-2012*. 2012:824–829. doi:10.1109/HPCC.2012.116.
16. Centres H. *High Data Rate Processing and Analysis Initiative (HDRI)*.; 2014.