PRINCIPAL CONTACT:

M. Bauer
Department of Computer Science,
University of Western Ontario
bauer@uwo.ca

Instrumentation & Techniques

Development of Remote Control and Streaming Data Interpretation Capabilities on the VESPERS Beamline

D. Maxwell, D. Liu, E. Matias (1), N.S. McIntyre, M. Suominen Fuller, J. Qin and M. Bauer (2)
(1) Canadian Light Source, Inc.
(2) Faculty of Science, University of Western Ontario, London, ON, N6A 5B7

Introduction

Science Studio is web services software that is under joint development at the CLS and University of Western Ontario, with the prime objective of allowing full service access to beamline facilities by users located anywhere in the world. It is also intended to be an experiment management system that will provide permissions, scheduling, sample transfer and communication with floor personnel during a remote experiment. The initial prime application is on the VESPERS beamline where a high level form of software is being designed to control and monitor many motor and detector functions that are normally run under EPICS software. In the past year, significant progress has been made in providing remote users (mostly at UWO) to visualize microscopic areas on a sample, followed by the acquisition of XRF spectra and maps. In parallel, control of XRD functions has also advanced to the point that large blocks of data can be moved in one step to cell processing at SHARCNET at UWO.

In parallel with the above activities, a new project called ANISE (Active Network Interchange for Scientific Experimentation) has begun. The goal is to create a high speed processing network capable of carrying out near real-time analysis of scientific data arriving at high rates (up to10 terabytes per day) from advanced scientific projects such as synchrotrons. This will allow users to make high level judgments about the information from an experiment within minutes and to make changes to the experiment to maximize the scientific impact of such experiments. Previously, such data would only be analyzed days after the access was terminated. The successful implementation of this network will completely alter the way that synchrotrons are used – shifting from a present emphasis on data gathering and storage to one of rapid processing of the data and feed back of that information to enable more effective experimentation. This will result in greatly improved productivity and much more industrial relevance for synchrotrons.

Discussion

The project is highly innovative because new stream processing software from IBM (Infosphere Streams) will be used interactively with High Performance Computing (HPC) resources at SHARCNET for the first time to process and interpret high fluxes of data that are produced by recently developed CCD detectors for X-ray diffraction experiments. This software and its support is being donated by IBM Canada and IBM Research. The data will be streamed from synchrotrons in both Canada and the US to processing by the SHARCNET system at UWO. The network will operate within the Science Studio distributed platform, greatly enhancing its value to users worldwide.

ANISE enhances the concept of a Network-Enabled Platform (NEP) by making use of advanced detectors combined with high speed computational facilities in order to process the near real-time streaming data, quickly and efficiently. Previous attempts to analyze such data streams have been done on dedicated servers that have been slow and cumbersome. The platform being developed here would be expandable to process data coming from synchrotrons and comparable experimental sources around the world, in near real-time.

Canada will benefit from its leadership in the development of ANISE: Canadians would have early access to ANISE which will enhance our use of advanced synchrotron studies of industry critical materials and aid in better academic understanding of such materials. As well, the experience of developing this project will lead to a creative software capacity to provide such active high speed interpretive networks for other purposes such as medical imaging. Finally, by developing such an enhanced processing capability, Canadians will learn how to make the recent large investments in science more available to more working scientists, potential future scientists and to the public at large.

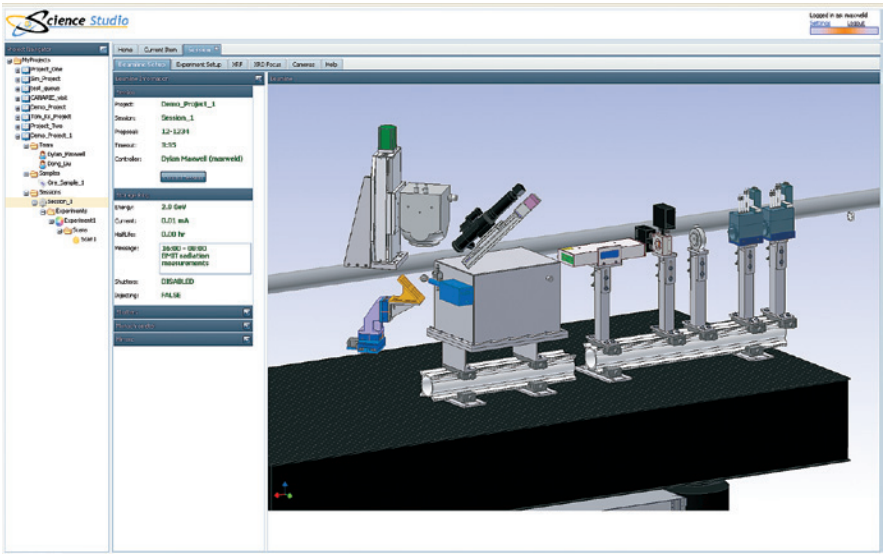
Science

Figure 1(a and b) shows two of the Science Studio screens presently used during a remote XRF experiment. Figure 1a is the page showing the present beamline conditions. In Figure 1b the access of data after the experiment, including an optical image of the mapped region and the raw data files that are available for download. All experimental parameters are stored in Common Data Format (CDF) files that may be accessed at any time from the Science Studio website (www.sciencestudioproject.com) where further information on the project may be obtained.

Conclusion

It is envisioned that Science Studio will provide a remote user with several brief sequential access sessions that can be arranged on relatively short notice. Data acquired in one session would be fully interpreted using ANISE and discussed among collaborators prior to the undertaking of a subsequent experiment.

a)



b)

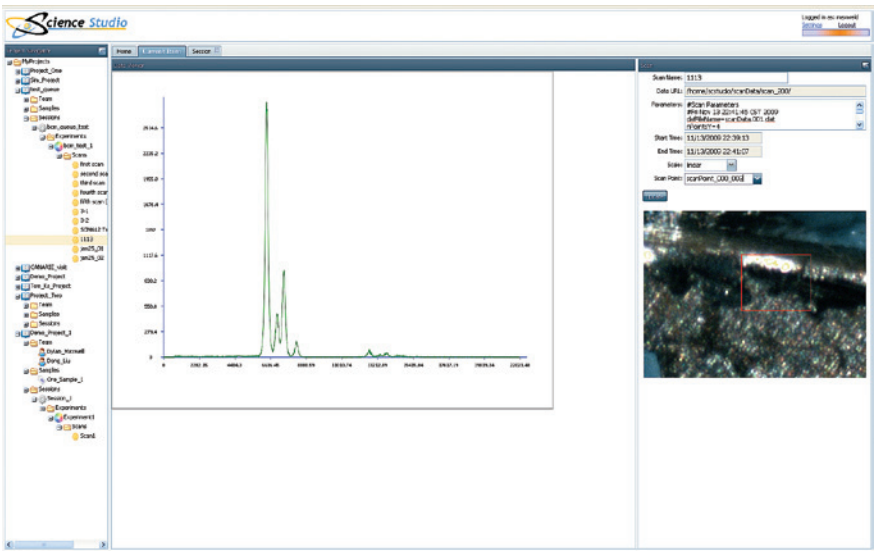


Figure 1: Two of the Science Studio screens presently used during a remote XRF experiment are presented here. (a) is the page showing the present beamline conditions (b) the access of data after the experiment, including an optical image of the mapped region and the raw data files that are available for download.