CANADIAN LIGHT SOURCE:
DATA MANAGEMENT

MBA 832

Management Information Systems

Professor Graham Links

Team A:
Tyler Case
Carl Nahachewsky
Tim Shirkey
W. Graham Stephens

March 28, 2004

List of Figures...in order following Text

Appendices… in order following Figures

List of Tables

Executive Summary

The Canadian Light Source (CLS) provides the opportunity for scientists to research the nature of atomics through analysis of photon radiation created by an electron beam. The electron beam nearly reaches the speed of light, and then transfers to the storage ring, where it can circulate for hours, even days. Powerful magnets are used to guide and refine the beam until the researcher is ready to begin analysis on the shed photons. CLS currently has operational capability of 12 beamlines, of which 7 will be functional by the end of 2004, as well as capacity to expand to 30 beamlines and 50 research stations. Data management at CLS has been an evolving strategy, building on acquired systems and upgrading the local area network as technology and affordability became available.

The data is created through the exposure of a sample to photon radiation, which is then used to create a series or set of data files. This procedure allows scientists to probe the state of atomic matter, produce detailed imagery and representations of atomic matter, and analyze the composition of matter. This can result in data that is structured relationally in tables, as flat files, or as image files. The incremental expansion of personal computers (PC's) has been most cost effective, replacing machines and upgrading components as needed.

There are currently over 200 nodes across the Local Area Network (LAN) at CLS with another 20 PC's dedicated to network administration and three front-end processors to the Internet and telephony, all entirely based on fibre optics. Current servers include Dell-EMC and Dell PowerEdge tower servers, providing LAN storage capacity at 5 Terabytes (TB). Many machines have either RedHat™ or Windows XP™ as operating systems, depending on the user's needs. This allows for network administrators to set and maintain password verifications and user access controls. Client created data is typically stored for three weeks where it is most usually transferred over a removable disc drive, capacity to 200 Gigabytes (GB), or a DVD disc, capacity to 4.7 GB. The slow moving transfer over Internet throughput is also another option, but the decision should be dependent on cost and quantity of data, with respect to urgency of need, which may give rise to a "sneaker-net" strategy. CLS plans to allocate 75% of beamline research time to academic institutions and project-based academic research, and 25% to industrial commercial application.

CLS will want to focus on their system architecture to ensure that it can meet the needs of 50 research stations, manage and schedule the batch processing, and provide adequate local storage for projects. CLS may also further want to consider offering a wider array of data and computational services, which could help the researcher compile data more efficiently, acquire known results quicker, draw next-step conclusions faster, and create new testable hypotheses'. These added-value services may provide the client with incentive to stay and plan further projects with CLS.

1.0 Introduction

Canadian Light Source (CLS) Inc. operates a third generation synchrotron facility, roughly the size of a football field and capital investment of $173.5 million, on the University of Saskatchewan campus. This state-of-the-art research facility provides a new tool for researchers to apply microscopic analysis to molecular activity and determine the atomic state-of-nature of any material. The light is created through a beamline, a series of interconnected magnets that propel the light through incremental condensers to create a beam of light the width of a human hair. The facility will offer services for academic and commercial researchers to utilize the beamline for capturing and compiling data for research analysis.

This paper examines the current framework of Information Systems (IS) at CLS, develops an understanding of the users or clients of CLS, and presents some future issues and needs that CLS may want to consider. The profile of IS will include the five needs for data management in the IS supply chain at CLS, namely data acquisition, data storage, data access, data manipulation and data transfer. Figure 1 illustrates the pathway of the supply and maintenance of data flow at CLS, as well as the users' specific needs, and limitations, during use of CLS services. This paper fully explores these needs and limitations and how CLS should address them to meet the requirements of their diverse client base.

2.0 Canadian Light Source Profile

A synchrotron creates intense, highly focused beams of light across the electromagnetic spectrum, including infra-red, visible light, ultra-violet light and x-rays. Bursts of high speed and high energy electrons are produced by an electron gun and accelerated through a linear accelerator to nearly the speed of light. The electrons are then sent into the circular booster, where their energy is further increased to 2.9 Giga-electron volts (GeV). When the electrons reach the desired energy level they are transferred to the storage ring where they continue to circulate for hours, or days, dependent on the speed and width researchers need for data acquisition.

This intense light is delivered in very short pulses, 10 to 100 picoseconds, and directed down beamlines. Since electrons normally travel in straight lines, magnets in the storage ring are used to force the electrons to bend around corners in the series of straight segments that make up the ring. As the electrons circle the ring, they shed energy in the form of photons. This photon radiation is then directed along the beamlines to scientific work stations for experiments.

CLS currently has 7 operating beamlines and plans to implement a total of 12 beamlines serving 15 research stations. This modular capacity will allow CLS to increase the number of beamlines to 30 with up to 50 research stations (see Appendix 1).

The data management process begins with a user set-up, where user designed software is utilized to set-up the collection parameters under which the experiment will occur. The user also manipulates the sample and prepares it for treatment from the beamline (see Figure 1).

Once the sample has been prepared, the data acquisition can occur. This is the critical part of the experiment where the sample is subjected to the focused, specified light. There are four primary types of analyses that can be created by synchrotron light:

- probe the physical structure of materials down to the level of atoms and molecules;
- analyse the chemical composition of materials;
- produce detailed images for medical and other purposes; and,
- create tiny three-dimensional structures out of silicon and polymers.

These types of analyses are critical to note as they form the structure of data that is created in the research experiment. Although specific data can vary by user, the basic structure of data produced by experiments types is usually includes the following:

- material structure – relational/flat database
- chemical composition – relational/flat database
- spectrographs – .jpeg, .bmp, .gif, other image files
- three dimensional structures – .jpeg, .bmp, .gif, other image files

2.1 CLS Hardware Profile

The computer infrastructure at the CLS can be referred to as a legacy system, in the sense that incremental expenditures have been spent on hardware as funding became available and many older dated technologies still form part of the network structure. Primarily, desktop or personal computers (PC) systems were replaced as needed or incrementally upgraded as processing power and capacity of machines offered on the market became more affordable. Through time, the numbers of PC's acquired were used to establish an internal local area network (LAN) topology, currently estimated at 200 work stations, or nodes. The central server of the network, or the network operating system, is essentially a sub-unit of aggregated PC's that act as servers to the network. CLS recently acquired a new Dell PowerEdge server for the network which provides additional storage of 1 terabyte (TB), adding it the other five Dell-EMC CX300 and CX500 servers onsite. Total current storage capacity is now 5 terabytes (TB) with the capacity-to-build up to 48 TB utilizing current network configurations, dependent on operating system software. Current needs however, do not require more than 1 TB. This network of PC's allows the IS staff to dedicate certain machines as front-end processors for communication, others as network servers and controllers, and still others as personal workstations and allows to create different levels of access through password verification log-on and folder access controls.

2.2 CLS Software Profile

As a result of this fragmented structure of PC's acquired through time and aggregated to create the LAN, there is no uniform application of operating systems among all machines. Most internal workstations execute a windows operating system, of which most have been upgraded to Windows XP. The need for clients to be able to set-up the data acquisition computers to cater to their needs requires that they be able to manipulate the operating system source code. As a result, CLS dedicates certain machines to utilize RedHat™, an open source Unix operating system. Other direct workstations utilize Windows XP and offer an array of office software and database compilation programs, such as MS Excel™, MS Access™ and SAS™ (see Appendix II). The open source code of UNIX™ allows for the customization of software to meet the autonomous needs of clients. CLS offers a standard software package with these Unix work stations that provides the user a variety of software suites for the monitoring, acquisition, configuration, compiling and organizing of acquired data. The advantage of creating a mixed software profile is the structure offers scalability for users and modularity for hardware in future expansion. License and software management is negotiated for the whole organization and premiums are paid based on the number of users to the system. As a result, CLS can offer a good variety of software choices to visiting clients. CLS does not offer analysis or computational processing power to its clients and data acquisition occurs primarily through batch processing, where data is created and processed in short periods as the experiments are executed. As such, more intense database management and manipulation software is not maintained or offered, such as Oracle™, SPSS™ or SQL Server™.

3.0 CLS Data Supply Chain

Figure 1 illustrates the pathway of data at CLS as it is requested by the researcher, created by the technology, managed by a variety of nodes, and then delivered to the client. As such, it represents the CLS data supply chain.

3.1 CLS Data Acquisition

Data acquisition systems for the synchrotron are composed of various hardware and software components. Beamlines are enabled by processing electronics and connecting to various types of interfaces in data acquisition in the front end, which include virtual machine environment (VME – an operating system of ICL mainframes) crates, personal computers and instruments bus adapters.[1] , Each beamline is provided with collection and manipulation systems consisting of a computer and a software package. The software package should include excellent technical support, operation within an industry standard operating system and analysis of high resolution data sets (i.e. precision <0.0001 cm-1 for peak finding, interpolation, transmittance). In addition, the software will allow for simultaneous viewing and manipulation of data by multiple users off-site after the collection of data.

The minimum requirement for system hardware includes a 2 GHz Pentium or equivalent processor chip, 512 Megabytes of RAM memory, a removable storage media consisting of 100 Megabytes Zip drive, CD-ROM drive with DVD+RW/+R, 24X-48X capabilities and a sealed internal hard disk drive of at least 120 Gigabytes storage. In addition, hardware components require an Ethernet interface capable of 100 Mbps operation, and 2 USB ports.

The brilliance and highly collimated X-ray beams in synchrotrons result in large amounts of quality diffraction data that are routinely collected from very small crystals in a short period of time. At the latest insertion-device end station, a complete data set can be collected in under ten minutes. Unfortunately, improvements in computational data collection methodologies do not improve the overall bottleneck issues related to data collection, namely the inefficient use of beam-time through crystal handling and alignment, and book-keeping.

3.2 CLS Data Storage

The data storage needs at CLS are very short term, apart from maintaining information needs internally for management staff. Client data acquisition will occur over very short time frames where data is created quickly over a 15 minute experiment and then analyzed, compiled and sequenced before proceeding with the following research step. This batch processing of data allows for extensive idle time of server resources. Further, the limitations of the network framework, with regard to computational or processing power, would not currently meet the needs of comprehensive data analysis. The costs for beamline time are also extensive so researchers have a great interest in maximizing their efficiencies and getting their data sets as fast as possible. Hence, the data storage needs at CLS will only need to meet these short term requirements, up to about three weeks. Electronic data interchange (EDI) secures the internal data during this timeframe where data is archived on a backup magnetic tape each night, providing the least expensive and most stable back-up system. For server storage, CLS maintains four Dell-EMC™ CX500 servers and one Dell-EMC™ CX700[2], most recently acquiring a Dell PowerEdge™ server making current storage capacity 5 TB (see Appendix II).

3.3 CLS Data Access

The main issue arising for CLS and its clients in relation to data access is to ensure that a controlled environment exists in order to reduce errors, interruptions of service, breaches of security and unauthorized access to CLS and client data. Specific policies and procedures must be incorporated into the implementation and design of CLS's IS to safeguard the IS, as well as internal and external access to CLS's data that is stored by CLS for its clients.

A combination of automated and manual measures should be implemented to protect and ensure that CLS's information systems perform their security functions according to management protocols. The methods, policies and organizational procedures that are utilized to maintain the safety of CLS's and client's assets and the adherence to specific CLS management standards should be contemplated in the scope of CLS's controls. General controls implemented by CLS and specific application controls applied from the functional business area of beamline experiments should be anticipated so that potential vulnerabilities and control issues are identified during this start-up phase.

General controls deal with the design, security and utilization of computer programs and the required security of data files in relation to the input and throughput of data to CLS's IT infrastructure. These general controls include physical hardware controls, software controls, computer operations controls, experimental data security controls, the controls over the systems implementation process and additionally, administrative controls which establish an overall control environment in the organization.

The application controls are specifically unique to each of CLS's computerized applications, such as payroll, personnel and more specifically, in this case to the retention and facilitation of experimental beamline data. Data security controls are designed and implemented to ensure that valuable experimental data files on disk or tape are not vulnerable to unauthorized access, while they are being used by an experimenter or when CLS is temporarily storing the data for use in experiments.

Data security controls require the input and review from end-users and CLS managers. CLS's information specialists should be responsible for certain aspects of these data security controls, such as making specific computer terminals available only to authorized users (internal and external clients). This can be accomplished through using system software and application software to create specific passwords, which users (internal or external) would require in order to access systems or specific client experimental data.

Specific data security controls, as application controls, generally include both manual and automated procedures that are responsible for ensuring that only authorized data are completely and accurately processed by the required application. Application controls are classified as input controls, processing controls and output controls. Input controls monitor data for accuracy and completeness as the data enters the system. Specific input controls are required for input authorization, data conversion, data editing and error handling. The processing controls ensure that the data is complete and accurate during any updating. Examples include computer matching, run control totals and programmed edit checks. Output controls are implemented to ensure that computer processing results are complete, accurate and properly distributed.[3]

3.3.1 Internet Security for CLS and its Clients

CLS anticipates using the Internet for data transfer to its clients in combination with "sneaker-net" as requested by the client. Internet access by clients to CLS or using the Internet as a potential transmission facility for client data raises vulnerability and Internet security issues which will require a security infrastructure for CLS.

CLS's architecture could include a Web client, a server and the CLS corporate information systems that are linked to its databases, which hold stored client data. The client and CLS are susceptible to computer viruses, loss of hardware, including machine and line taps. Internet communications are further subject to potential threats from Internet tapping, message alteration, sniffing and even potential theft or fraud. Servers are also susceptible to hacking, denial-of service attacks, computer viruses, line taps, theft and fraud. The data-bases themselves are also exposed to potential theft, alteration and copying of data.

A major security access issue for CLS is the "openness" and "closed" scope of its IS. In order to benefit from having clients send or receive all types of data the system must be open, however, the system must balance this openness with adequate security.

As CLS is linking to the Internet and transmitting information and data to clients, adequate and special security procedures and technologies must be implemented. The potential risks associated with linking to the Internet or having client access to stored client data or CLS operating beam line data can be minimized with the use of a firewall, to prevent unauthorized users from having access to CLS's private network. The firewall is generally located between internal LANs and the Internet. The firewall identifies and checks specific information, such as names, Internet Protocol (IP) addresses and applications of incoming information and potential client requests against access rules specifically programmed into the system. The firewall protects against unauthorized access or communication into the network or communication out of the network. This facilitates the organization in enforcing a security policy on Internet traffic and its network.

The two main types of firewalls are proxies and stateful inspection.[4] Generally, proxies are considered to be more secure of the two, however, they degrade network performance as a result of having to undertake a lot of work and as a result, can have a drain on system resources. Stateful inspection scans each packet of incoming data and determines its source and other destination addresses or service requests. While it does not consume the resources that a proxy system does, it is not as secure because some data passes through the firewall. A valuable firewall, is one that has been created by a very fine detailed compilation of information with respect the types of specific application and addresses that are allowed or rejected. A firewall however, also has limitations, as it can deter but not prevent complete network intrusion or penetration by outsiders.

Intrusion detection systems can add security in that they can feature full-time monitoring tools and be placed at the most vulnerable points of the network. There can be customization of the intrusion detection system to shut down a part of the network in the event that it detects unauthorized inquiries or traffic.[5]

3.4 CLS Data Manipulation

Before the inception of the CLS, the Saskatchewan Accelerator Laboratory was constructed on the U of S campus. This facility would eventually be incorporated into the design of the CLS. Computer processing was fundamental in creating and interpreting data but funding was scarce to meet the needs of IT expansion. As a result of these exogenous variables, an incremental expansion strategy arose. The small number of CPUs that were used upon the onset of the facility became the basis of a larger system of linked processing units that was admitting ugly, but functional. This original system was crude with a number of connected CPUs sitting on a table to collect data and eventually process it. Though the system was not particularly aesthetically pleasing, it met the needs of the users and provided a simple and cost effective way to incrementally expand the system. This processing system resembled a modular design as it was easy to expand and remove processing units.

The current capabilities of the CLS have remained similar to these origins. The system used is a form of batch processing. Though the CPUs are linked together, they are neither working on the same program nor are they able to process more than one instruction by breaking down a problem into smaller parts, thus not resembling parallel processing. Instead, these CPUs are a method of accumulating and storing data that can be processed when the desired amount of data is collected. This method is typical of older systems but also meets the uneven processing requirements of the CLS.

Processing requirements at the CLS vary greatly from user to user. Clients using the facility generally bring their own software that they expect to use while at the CLS. Actual processing of client needs is very uneven. Only a small portion of client time at the facility is spend utilizing computational capabilities. More research time is spent constructing experiments, discussing or interpreting results than actually collecting and processing data. Processing time does reach substantial peaks though for only shorter periods of time. Current data storage only uses approximately 1 TB while capabilities are as much as 5 TB. Thus, the legacy system that has emerged still meets the needs of the facility through batch processing. An effective analogy is of the eight hours daily spent researching at the CLS, seven are spent constructing and discussing data while only one hour is actually spent processing data. If substantial capital and time is spent expanding the system to reduce computing time from sixty minutes to thirty minutes, cost-value analysis does not equate. This does however pose some serious potential future issues.

3.5 CLS Data Transfer

The main issues facing CLS in data transfer are the current limitations of Internet throughput and security. External clients are limited in the amount of data transferred via the Internet and may need to rely on other methods such as direct downloading of data to company hardware or the use of “sneaker-net”.

To combat security issues, options include the use of an Electronic Data Interface (EDI) network run over an Extranet or encryption. The option of an Extranet may be a future consideration as it is expensive and CLS currently does not have the client base to fund this option. Encryption is the coding and mixing of messages to prevent the unauthorized access or analysis of the data being transmitted. There are several types of encryption with a current popular form being that of “public key” encryption, where the public key is maintained in a directory and the private key used by the recipient being held confidential.

Further security measures include digital certificates to verify data transfer authentication and message integrity. Since CLS has chosen to transfer data on the Internet to its clients, further options are to utilize security measures such as Secure Socket Layer (SSL) and/or Secure Hypertext Transport Protocol (S-HTTP). This allows the client and the server computers to manage encryption and decryption as they communicate with each other during a confidential and secure transmission.[6]

Clients are also amenable to the use of “sneaker-net” for the transmission of beamline experiment data. This is simply the downloading of an experiment’s results onto a DVD or hard drive to deliver the disc or drive to the client, probably through a contracted courier, such as

FED EX to take advantage of expeditious international delivery and mitigating liability through the carrier's insurance. In addition, delivery by "sneaker-net", will require compatibility between CLS's systems and the client's to ensure proper transmission of experimental data.

The issue of the "sneaker-net" and throughput for CLS is relevant because CLS's internal capacity of data transfer by fibre optics currently exceeds the capacity of copper line through the Internet. As a result, when the Internet technology for transfer catches up, Internet transfer will become more efficient to CLS and its clients to transfer large amounts of data. In the meantime, CLS's strategy is to utilize "sneaker-net" for large amounts of data, which can be transferred to the client by overnight or next day courier.

4.0 CLS Client Profile

The CLS is the only synchrotron with a focus on private-public partnerships and serving industrial researchers. The CLS plans to allow private firms to utilize 25% of the total beamline time. The majority synchrotrons around the world largely serve government and academia institutions, providing beamline time to industrial firms that amount to 10% of total usage.

4.1 Academic Profile User

The average academic user is more likely to have utilized other synchrotrons as a part of their fundamental research. Consequently, academic users are familiar with synchrotron use and associated components of data management. The academic users rapidly utilize the beamline and retreat to the laboratory for data analysis. In addition, academic users would require less hardware and software support from the CLS IT staff.

4.2 Industrial Profile User

Industrial users may regularly utilize the synchrotron as an important tool for innovative discoveries, while smaller firms may not understand the function of a synchrotron and how it can improve the firms' technology and competitive advantages. Industrial scientists and technicians that commonly use a synchrotron would exhibit similar characteristics as the academic users described above. Greater time and effort will be required to help new unfamiliar CLS customer's advance their knowledge of the intricate mechanisms of data management. CLS staff will likely be asked to provide opinions of the available relevant software packages for their new customers. In addition, the security concerns related to data storage and transfer are likely to be more stringent for industrial users than for academic users.

5.0 Future Needs and Issues

The IT system that has emerged at the CLS currently meets the needs of the seven operational beamlines in use. The facility has the capabilities for as many as 30 beamlines. If the CLS expands the number of beamlines and grows its client base to optimal usage, processing capabilities will need to rise commensurately. Not only will individual client needs and storage requirements rise but the system will need to accommodate multiple simultaneous projects which all need the bulk of the server processing power in order to execute the batch processing. This situation highlights the limitations of the batch processing infrastructure and the inability for it to meet the needs of a large and busy client base. It will be critical that system processing time is scheduled properly to ensure batch processing can occur at every research station on every beamline. At some point, if growth and feasibility are achieved, the system may have to be replaced with a more advanced database and network administration, such as Oracle.

Hardware and software compatibility are current concerns and as the facility continues to expand it may become paramount to the IT success of the CLS. Each client's autonomous needs require the IT staff to be proficient in many of the software applications. Standardizing requirements and employing user guidelines may ease the transition for both the CLS and its users. A testament to the critical importance of maintaining skilled and motivated IT staff that ensure network administration operates without flaws.

However, the fragmented set of PC's with no uniform operating system makes the implementation of network data sharing and software license sharing much more cumbersome. Some clients may want to execute software built for a Macintosh operating system, so CLS will want to ensure there are machines, perhaps even UNIX machines that are able to execute the compiled machine code, which can often be dependent on the operating system.

The current throughput of Internet capabilities limits the amount and size of data that may be transferred to clients, outside the fibre optic LAN at CLS. The firm will need to address a means of transferring large amounts of data, securely and in a timely fashion and how these Internet throughput limitations will affect the decisions their clients make.

Finally, CLS will want to consider the range of services that can be offered to the travelling researcher with regards to data analysis and manipulation. While software applications, such as ESRI and Oracle, are expensive to acquire and maintain, the benefits of enticing a research community to Saskatoon to use the beamline repeatedly, while adding to the local economy, is a welcome event.

Figure 1 – Data Management Supply Chain at CLS

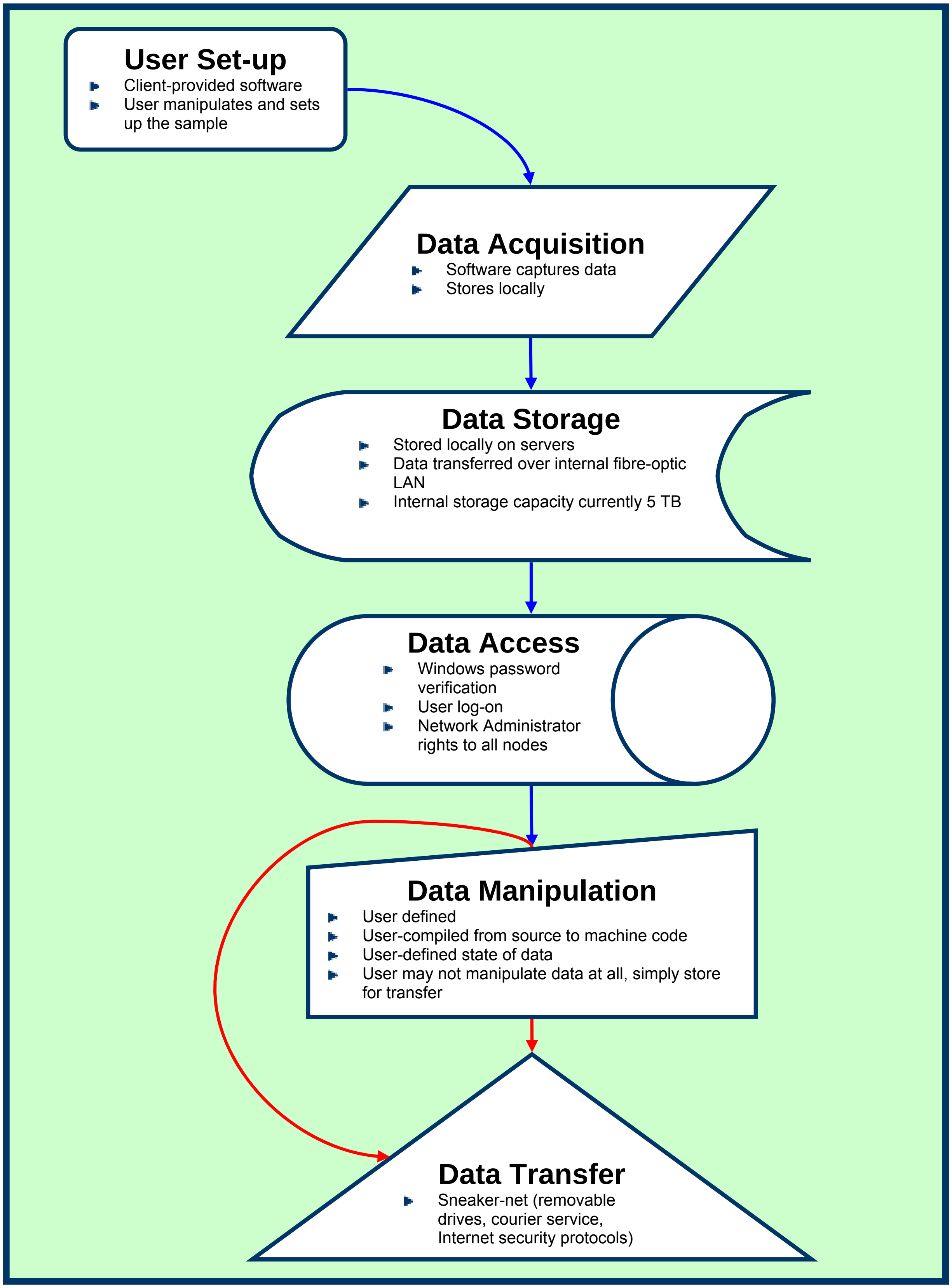
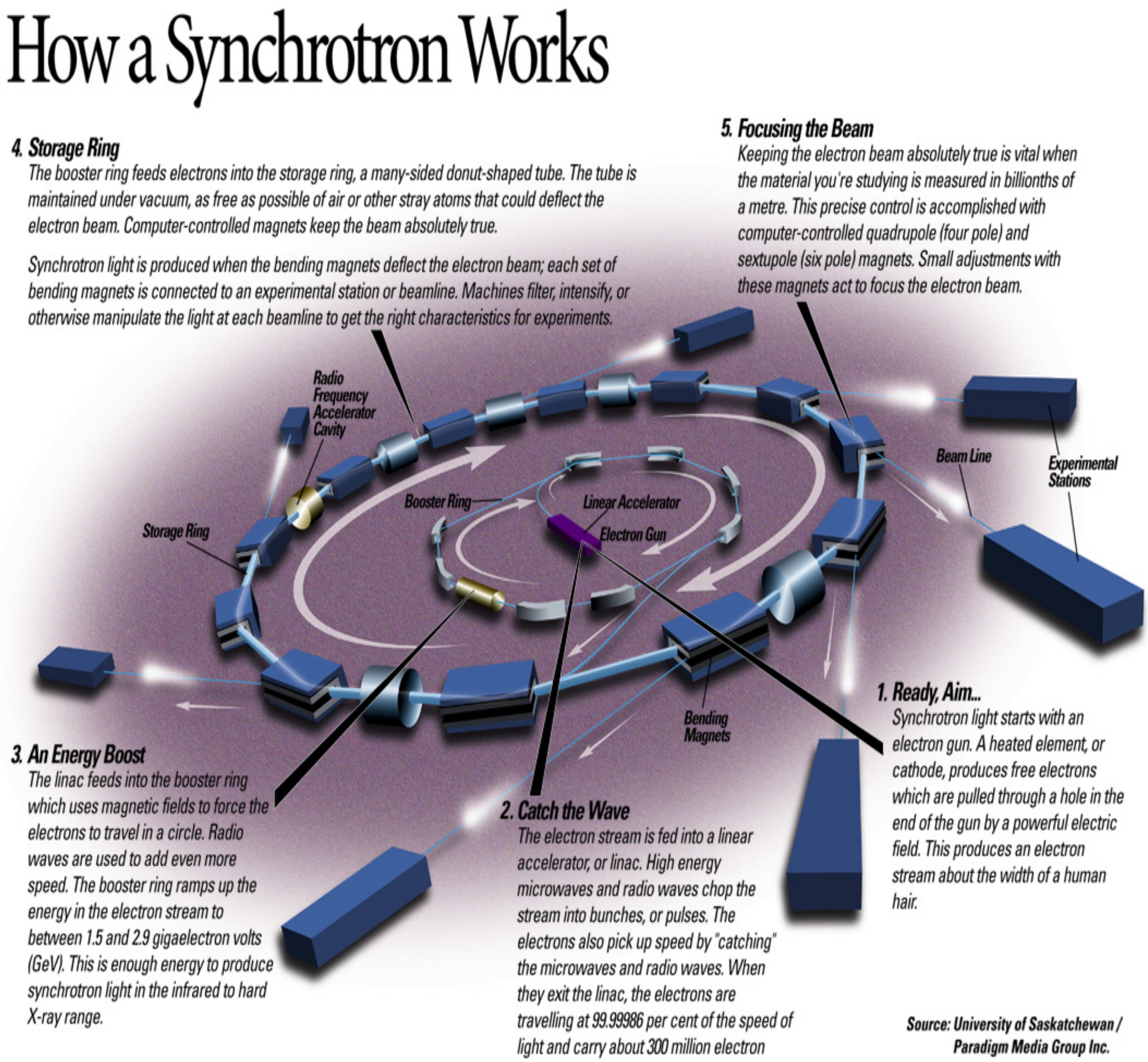


Figure 2 – The CLS Synchrotron



Appendix 1: Beamline Description and User Profiles[7]

Phase I

Canadian Light Source Inc. (CLSI) in collaboration with Canadian universities and agencies have designed and commissioned the following first six beamlines:

1. High Resolution Infrared Spectroscopy (2 beamlines: Far and Mid IR)

***Team Leader*:** A.R.W. McKellar, Steacie Institute for Molecular Sciences, National Research Council of Canada, Ottawa, ON.

Capabilities: High resolution spectroscopy, high pressure spectroscopy, surface spectroscopy.

Research Areas: Chemical composition and structure of polymers and sol-gels for manufacturers of bulk plastics, paints etc. In Addition, various life science applications, including medical and dental research, i.e. Alzheimer's, osteoporosis, bacterial identification, imaging and analysis of micro array slides, monitoring diffusion processes in biogels, biochemical analysis of samples in the 10 to 15 microns size, identification of cell types in pathology and forensic studies.

Applicable industries:

a. Innovative pharmaceutical companies in Canada, United States and international market and generic pharmaceutical companies
b. Biotechnology companies
c. Agricultural companies – agricultural chemicals, protein, carbohydrate, fat and fibre analysis in grain seeds, meat, milk, cheese and animal feed
d. Food and Beverage companies: food packaging, sugars, fibers acidity in juices, fermentation and microbial analysis in food products
e. Forensics and Criminology: heavy metal poisoning, pollution
f. Industrial Manufacturing: adhesives coatings, cosmetics, inks, dyes, soaps, organometallics, paints, polymers, pulp and paper, textiles
g. Petrochemical industry
h. Nuclear power industry
i. Semiconductor and electronics industry

2. Variable Line Spacing Plane Grating Monochromator (VLS-PGM)

***Team Leader*:** T.K. Sham, University of Western Ontario.

***Capabilities*:** The Variable Line Spacing Plane Grating Monochromator (on an insertion device source at CLS) will be a facility beamline. It will serve the needs of the CLS users in VUV spectroscopy and soft X-rays (5 to 240 eV) surface analysis. With its rapidly turn-round sample-transfer and manipulation potential, VLS-PGM will offer enhanced and desired capability for surface analytical applications. The beamline will be very competitive in this energy region for L-edge absorption analysis of thin films containing Al, Si, S, P and Cl, K-edge analysis of B, and variable photon energy valence band and shallow core level photoemission studies.

Research Areas: It will be used primarily by CLS as well as CLS users to engage in applied research and industrial applications in soil science, fertilizer chemistry, materials surface chemistry, coating for artificial joints, corrosion, geological chemistry, catalysis and biochemistry.

Applicable industries:

a. Fertilizer companies
b. Petrochemical companies
c. Coal mining companies
d. Power companies
e. Advanced metallurgical companies
f. Environmental companies

3. High Resolution Spherical Grating Monochromator (SGM)

Team Leader: T.K. Sham, University of Western Ontario.

Capabilities: igh resolution spectroscopy, high pressure spectroscopy, surface spectroscopy.

Research Areas: Chemical composition and structure of polymers and sol-gels for manufacturers of bulk plastics, paints etc. In Addition, various life science applications, including medical and dental research, i.e. Alzheimer's, bacterial identification, osteoporosis, imaging and analysis of micro array slides, monitoring diffusion processes in biogels, biochemical analysis of samples in the 10 to 15 microns size, identification of cell types in pathology and forensic studies.

Applicable industries:

a. Industrial Manufacturing
b. Plastics manufacturing
c. Life science companies
d. Biotechnology companies
e. Medical diagnostic companies

4. Soft X-ray Spectromicroscopy (SM)

Team Leader: Adam Hitchcock, McMaster University.

Capabilities: Scanning transmission X-ray microscopy (STXM): Imaging, microscopy; photoelectron emission microscopy (PEEM).

Research Areas: Soft materials (biology, environmental, polymer), electronic structures of advanced structures including semiconductors, high temperature superconductors; coating of artificial limbs, anti-wear coatings, next generation of optical luminescence materials such as porous silicone; ceramics; magnetic structures; lubricants; geochemistry; and oil refining.

Applicable industries:

a. Semiconductor industry
b. Petrochemical industry
c. Chemical industry

5. Macromolecular Protein Crystallography

Team Leader: Louis Delbaere, University of Saskatchewan.

Capabilities: This beamline is a hard X-ray DCM with a spectral range of 6.5 to 18 keV that will provide high resolution, multi-anomalous diffraction crystallography

Research Areas: Detailed atomic-scale images of molecules like viral and bacterial proteins, which will provide insight into novel drug design.

Applicable industries:

a. Pharmaceutical companies
b. Agriculture chemical companies
c. Plant and animal biotechnology companies

6. X-ray Absorption Spectroscopy (XAFS) with Microprobe Capabilities

Team Leader: De-Tong Jiang, CLSI, University of Saskatchewan.

Capabilities: Imaging, microscopy, micro-probe, micro-spectroscopy, XAFS, reflectivity.

Research Areas: XAFS application is widespread in almost all areas of science: physics, chemistry, materials, biology, environmental science, engineering and industrial research.

Phase II

Phase II partnership proposals are under negotiation for the development of additional beamlines. March 8th, 2004, the Canada Foundation for Innovation announced funding of $18 million for the development of five new beamlines. The beamlines are expected to be fully operational in 2005-06.

Appendix I: Table 1 – Phase II Beamline Planning[8]

Beamline	Team Leader	Organization	Description
Biomedical Imaging and Therapy (BMIT)	Dean Chapman	University of Saskatchewan	Advanced imaging for medicine offering unprecedented detail and high precision radiation therapies for cancer
Soft X-ray for Micro-characterization of Material	T.K. Sham	University of Western Ontario	Determine materials structures to nanometer-scale. Applications include environment, electronics and medicine
Very Sensitive Elemental and Structural Probe Employing Radiation from a Synchrotron (VESPERS)	Stewart McIntyre	University of Western Ontario	Determine trace elements and crystal structure in micro-samples. Applicable to mineral ores and metals
Resonant Elastic and Inelastic Soft X-ray Scattering	George Sawatzky	University of British Colombia	Atomic-scale microscopy with applications in environmental science and advanced material development
High Throughput Macromolecular Crystallography	Natalie Strynadka	University of British Colombia	Detailed, atomic-scale images of molecules including viral and bacterial proteins.

Phase III

Phase III beamline investment models are currently being explored. Selection and funding of new phase III beamlines is not expected until 2007-08.

Appendix II
Notes: Mon. March 22/04

Skeeter Abel-Smith
Manager, Information Systems
Canadian Light Source Inc.

- 8 total staff
- about 200 computers
- EDS consulted- practical implementation
- CLS only stores data for 3 weeks then the responsibility of client (may increase as throughput increases). Current storage rate far below capacity. Maybe no need for large database.
- Clients do not utilize throughput, only 10-30 gigs.
- Researchers generally want to be left alone, they have their own systems.
- Customer is king but not always right, but you must grin and bear it.
- Client usually brings own software: i.e. Photographs of material & graph; Tissue Sample "slices" (spectrograph?). Needs depend on what software client brings.
- Most problems are human interactions that are not related to IT. i.e. decisions, cleaning, change beam line etc. If IT capabilities are increased, still does not address human problems, so why spend money??? There is a lot of repetition so a lot of material to discard.
- Incremental Strategy. Current technology staying ahead of requirements
- Most CLS software is customized in-house
- Long history with legacy system with no money to expand. Current problem of updating and scaling outdated material (modular system). Technology was outdated when installed years ago now archaic.
- Everyone is connected to network, which is tied to University pipeline.
- Subnets and V-lines
- Ret Hat Linux and Microsoft (XP) Operating Systems.
- Incremental Strategy. Current technology staying ahead of requirements
- Problems with connecting and sharing info via Internet (slow, unsafe etc.). How do you reach files when you are geographically dispersed? Directory.
- Maybe easier to actually send a hard drive via FedEx than try to use network to send large files- sneaker-net
- Timing issues for e-mailing to CLS because of virus scanning.
- <u>Computational</u> – Linux, More like batch processing not parallel.
  Nodes – PC's that create network which is easy to add and subtract.
- <u>Data</u> – EMC storage (server) from Dell
  EDIC – tape backup is the cheapest/fastest way to take a daily snapshot
  Using only 1 terabyte with capacity to 5 and scalability to 48. (≈$8000 per)
  PowerEdge upgrade
- Problems with clients storing "junk" (i.e. MP3s) on computer, may have to introduce storage maximums to circumvent problem.
- External transfer rate – enhance modular system, Internet to catch up.
- Basic infrastructure is plug in a go, but how to say how in laymen's terms .

References

[1] http://www.hwi.buffalo.edu/ACA/ACA99/abstracts/text/W0274.html

[2] Dell-EMC Server Storage -
http://www1.us.dell.com/content/topics/segtopic.aspx/brand/dellemc?c=us&l=en&s=gen

[3] Laudon and Laudon p. 459

[4] Laudon and Laudon p. 462

[5] Laudon and Laudon p. 462

[6] Laudon and Laudon p. 464

[7] http://www.cls.usask.ca/beamlines/beamlines.php

[8] http://www.cls.usask.ca/media/CFI.php