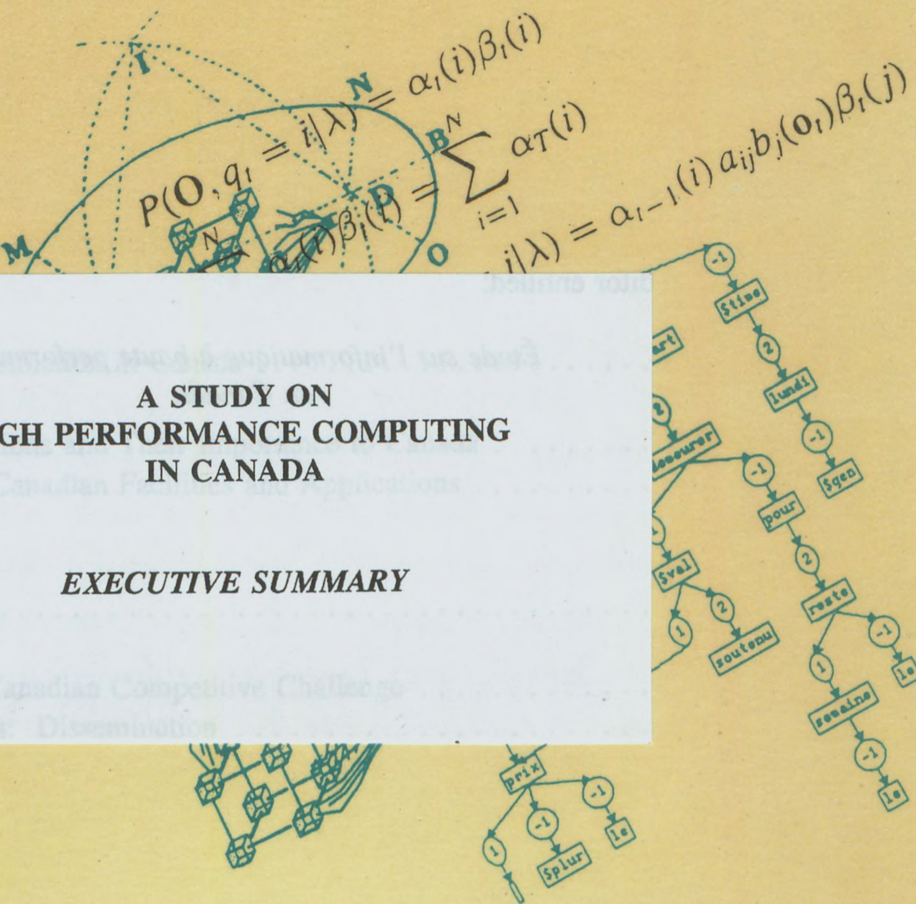
Industrie Canada Industry Canada



A STUDY ON HIGH PERFORMANCE COMPUTING IN CANADA

EXECUTIVE SUMMARY

CITI

Centre d'innovation
en technologies de l'information

Centre for Information
Technology Innovation

Canada

Industry Canada

A STUDY ON HIGH PERFORMANCE COMPUTING IN CANADA

EXECUTIVE SUMMARY

Centre for Information Technology Innovation (CITI)

Laval
February, 1995

A french version of this document is also available to the same editor entitled:

Étude sur l'informatique à haute performance au Canada

This report was prepared in connection with work carried out by the Centre for Information Technology Innovation (CITI) of Industry Canada.

Catalogue no. Co28-1/120-1995E
ISBN 0-662-23130-9

TABLE OF CONTENTS

Page

Introduction to the Canadian Study on High Performance Computing 1

1 Review of the HPC Situation in Canada 3

1.1 HPC Applications and Their Importance to Canada 3
1.2 Summary of Canadian Facilities and Applications 6

2 Canadian Issues 9

2.1 Address the Canadian Competitive Challenge 9
2.2 Primary Focus: Dissemination 10
2.3 Software: A Primary Consideration 12
2.4 Need for Canadian Facilities 13
2.5 Leverage via Existing Assets 14
2.6 Need for a Long-Term Commitment 16

3 CITI HPC Study — Final Recommendations 19

3.1 Promoting a Distinctive Canadian HPC Network 19
3.2 Building greater Canadian Awareness of HPC 23
3.3 Expanding Canadian Software Opportunities 25

Conclusions 29

Appendix

Appendix A HPC Technology Summary 33
Appendix B Comparisons of the HPC Situation in Canada with that in the U.S., Europe and Japan 35
Appendix C List of Advisory Committee Members for the CITI study of HPC in Canada 41

INTRODUCTION TO THE CANADIAN STUDY ON HIGH PERFORMANCE COMPUTING

In July 1994, the Centre for Information Technology Innovation (CITI), an applied research laboratory of Industry Canada, commissioned a study to examine the state of high performance computing (HPC) in Canada and to develop recommendations on a distinctive Canadian direction.

HPC is a term applied to the world's most advanced computing and communications technology, utilized by governments, academia and industry to solve the nation's most challenging engineering, scientific and information-management problems. Section 1 which follows contains examples of applications where HPC provides innovative approaches to improved industrial competitiveness, support of leading-edge research and provision of information that could guide government policy makers towards improved quality of life and a cleaner, safer environment.

This study was motivated by the realization of CITI directors that HPC provides a major stimulus to industrial competitiveness in the G-7 nations that have embraced the technology. Since there has been little apparent HPC activity in Canada, this study was intended to report on the actual situation and to develop recommendations for the dissemination of technology to the industries that could benefit most. Also, it was expected that detailed comparisons of the Canadian HPC environment to that of the U.S., Europe and Japan, where applications are far more advanced, would help identify the important opportunities open to Canada.

The Study began in July 1994 with a detailed examination of: the technology in place or planned; the usage and application of existing facilities; the benefits to industry, government and the academic community; the sources of financing; the nature of collaborative efforts; the strengths and weaknesses of current policies and programs; and, lastly, the identification of requirements, trends and success factors. The process included a review of the substantial amount of available Canadian and international documentation, complemented by interview with over 150 interested individuals across Canada.

A detailed preliminary report, documenting all the discoveries from the study, was issued on October 15, 1994. The report was distributed for further study to an advisory committee of 18 major stakeholders, representing a cross section of sectorial interest groups from all regions of Canada. The committee met for a full day on January 19, 1995, to discuss the relevant issues and develop a consensus of opinion on the recommendations for a Canadian strategy. These issues and recommendations are summarized in this report. For complete study details, the reader should obtain a copy of the companion document *A Study on High Performance Computing in Canada — Final Report*, issued in March 1995 (approximately 200 pages).

This project was conducted by Ron Crossan of the Montreal firm R.H. Crossan Consulting, in association with Bruce Attfield of iai Consulting Ltd. in Ottawa. Other technical experts and specialists were used occasionally to complete the collection and analysis of the required information. An advisory committee of 18 stakeholders, representing all regions of Canada, reviewed the study and subsequently met on January 19, 1995, to identify the issues and develop the recommendations that are summarized in this report.

1 REVIEW OF THE HPC SITUATION IN CANADA

1.1 HPC Applications and Their Importance to Canada

High performance computing (HPC) is generally considered a field of activity that introduces *computational science* as a third methodology used in science, complementing traditional theory and experimentation. By exploiting the capabilities of a special class of powerful computer systems, mathematical models are developed to examine physical phenomena that are too difficult or too costly to explore by means of traditional methodologies. The computing platforms that provide the necessary levels of performance include supercomputers, massively parallel processors and clusters of high performance workstations. Brief explanations of the various hardware architectures are provided in Appendix A, while this section examines the applications that run on these systems.

Academic researchers use these HPC techniques to examine the physics of very small-scale phenomena, for example the molecular reactions that reveal interesting properties of new materials, genetic structures or drug interactions with biological organisms; and of very large-scale phenomena, for example, problems in astrophysics that are so vast that they are impossible to explore by any other means.

The knowledge and skills developed in academic research are then passed on to industry, which does applied research *to make the organizations more competitive, through better product or process designs, with shortened market delivery times.*

For some industrial applications, such as in the design of aircraft, space vehicles, engines, ships and cars, HPC *simulation* is already an inescapable necessity, being economically and environmentally more effective than extensive physical testing. For example, auto manufacturers crash cars in computer simulations, as a lower cost alternative to the destructive testing of very expensive physical prototypes. Computational science is reaching the point where better information can be obtained from simulation than from real experiments. Using the crash-test example again, simulation allows "replays" in slow motion, and multiple retries with modifications to a multitude of design parameters. Now that HPC is well established and considered essential in very large-scale industries, the technology is beginning to trickle down to the next tier (e.g. auto-parts suppliers, where products may be optimized for maximum strength and durability, minimum weight, minimum manufacturing costs, etc.). It is these lower-tier applications that are most relevant to Canadian industry. It is important to realize that once any industry, in any country, discovers more cost-effective ways to design and manufacture products by applying HPC technology, they set a new world benchmark against which all other companies must compete.

As a second example, for years petroleum companies have modelled the subsurface from *seismic surveys* to minimize the number of "dry holes" that are drilled. New 3-D seismic interpretations of massive data sets enable a trained geophysicist to pinpoint the location of hydrocarbons far more precisely than ever before. However, these accurate mappings of the underground require the running of models on very large supercomputers, especially for marine surveys which require processing billions of seismic recordings. What is the payoff? To the *petroleum company,* avoiding one dry well could typically save $10 million to $50 million, which would enable the company to recoup any investment in HPC many times over. Imagine floating the multi-billion-dollar Hibernia rig to its final destination, only to discover that there is not enough oil in the reserves to cover the cost! For the Canadian *seismic processing company,* access to very large HPC facilities is mandatory to compete in the world market for the lucrative marine jobs, where the competitors all have access to supercomputers (e.g. in Houston or London).

A third industrial example is *rational drug design* — developing new drugs based on understanding a given disease and potential drug interactions at the molecular level. Bringing a single drug to market typically takes ten years at a cost of tens of millions of dollars. Molecular modelling on HPC systems is being used by the world's leading pharmaceutical companies to potentially eliminate years of design and discovery time at the front end of the product-development process. This speed-up not only adds millions of dollars to the profits of the developers, thereby stimulating the economy and creating jobs, but it will result in earlier relief of symptoms and hopefully cures for life-threatening diseases such as AIDS and cancer.

Optimization is another application for high performance computers, and it is becoming increasingly important in the operations of large industries and service companies. An example of a very challenging optimization activity involves the search for the most efficient design of an entire aircraft. This is of the utmost importance for economic designs of fuel-efficient aircraft, both civil and military. Aerospace companies in Canada are using some HPC methods; however there is a considerable lag behind the capabilities of U.S. manufactures with respect to high payback optimization. In many areas of business, the allocation of resources such as people, raw materials, capital or time can be among the most challenging problems. The difference between good solutions to these problems and a near-optimal or optimal solution may mean the saving of huge sums of money. The main application areas for optimization include the design of complex structures (e.g. aircraft), investment and portfolio management, production planning and scheduling, distribution planning, vehicle or aircraft routing and crew scheduling.

Data mining is a new class of HPC application for the extraction of information by means of complex query criteria from massive databases. Until recently, large database systems were the exclusive domain of large, expensive mainframe computer systems. However, the HPC community is now demonstrating that massively parallel processor (MPP) systems, by performing complex database operations on vast quantities of data in parallel, can potentially out-perform mainframes by orders of magnitude. The opportunities for increased cost-effectiveness are immense in large-scale applications for both massive corporate and government databases. An excellent example would be "micro-marketing" systems, which are already enabling financial institutions, telecommunications companies, airlines and retail chains to analyze consumer preferences and establish evolving trends in their businesses. A U.S. expert predicts that 10,000 information servers will eventually be installed on the information superhighway for multimedia applications in health care, education and entertainment. There are enormous opportunities in Canada to develop applications for these new information services.

A government *quality of life* example is the use of HPC applications to improve policy decisions regarding environmental protection. Environmentalists, by modelling the atmosphere and oceans, are predicting the effects of pollutants and noxious gases years before experimental results provide the data. This information enables government policy makers to take correct remedial action before the situation is out of control. The two most important environmental concerns that require international cooperative efforts are the expanding ozone depletion and the greenhouse effects of excessive carbon dioxide emissions. To comprehend these complex problems, there is no alternative to running HPC simulations on massive amounts of data over a large number of time steps — necessitating access to the world's largest supercomputers.

In summary, what HPC activities have in common is that they are at the cutting edge of information technology (IT). Hence, HPC is important to the IT industry in general because it is at the frontier of advanced information technology. This factor alone makes HPC essential to any industrialized nation. In addition, the advanced technology that is being developed in the international HPC programs is providing the information-highway designers with the required base technology. However, HPC is most important because of the extraordinary opportunities it creates and facilitates in other industries and in other areas of science and technology. The use of HPC, assisted by advanced human interaction including 3-D visualization, dramatically enhances the creative power of the human mind.

1.2 Summary of Canadian Facilities and Applications

The study confirmed that by international standards, there are still very few HPC facilities and users in Canada. *In the list of the top 200 installations in the world, Canada has only one entry* — the supercomputer used for weather forecasting at Environment Canada. Moreover, *Canada has only one publicly accessible supercomputer facility*, at the HPC High Performance Computing Centre in Calgary.

In Canadian universities there are a few small facilities scattered across the country, normally dedicated to a specific research project. For very large-scale computations, academic researchers may use the Calgary centre, or if they are collaborating with a U.S. university, they may be able to gain access to one of the U.S. government-funded National Science Foundation (NSF) centres. From the mid-1980s until 1991, there were supercomputers installed at the University of Toronto and the University of Calgary, but both facilities closed because there was no funding available either for continued operations or for technology renewal.

In industry there are even fewer installations and very few users. Workstations are prolific, but the large-scale problems, which could provide a major boost to competitiveness, are not being addressed. There are a few small HPC facilities at oil companies and seismic companies in Alberta and at aerospace companies in Quebec. A few other manufacturing companies across Canada are experimenting with HPC models on high-end workstations; once the models are tested and proven, there may be a need, at least in some cases, to have access to higher-performance systems. A few forward-thinking organizations are beginning to examine the value of adopting database parallel HPC machines for non-traditional applications such as consumer marketing analysis and other decision-support applications. It is expected that parallel HPC servers may also be required for full-scale multimedia information-highway services over the next few years.

It is evident from this study that there is considerable interest in HPC at the universities, but inadequate access to facilities. In industry there is a serious lack of awareness, which is the major impediment to investment in programs to adopt HPC methods and technology. The following section examines the important issues that help to explain the state of HPC in Canada and form the basis for appropriate recommendations for action.

2 CANADIAN ISSUES

During the HPC study and the subsequent advisory-committee consultations, a number of issues were raised regarding the acceptance and adoption of HPC technology and methods in Canada. At the advisory committee meeting on January 19, 1995, the issues were summarized as follows:

2.1 Address the Canadian Competitive Challenge

Canada needs to identify a set of "grand challenge" applications that would have a dramatic impact on *improving the world-wide competitiveness* of Canada's strategic industries. Case studies should be developed from industry success stories both within and outside Canada, in traditional areas such as the aerospace, automotive, pharmaceutical, medical, chemical, petroleum, materials and micro-electronics industries, and also in emerging non-traditional areas. For any applications where one may have to look outside Canada, the relevance to the Canadian business environment must be made explicit.

For example, whereas pharmaceutical R&D in Canada is reasonably significant, the focus is primarily on clinical trials rather than rational drug design, where HPC modelling techniques could save millions of dollars by reducing the overall time and cost involved in the drug-design cycle. If Canada had the HPC infrastructure, including well-trained experts on HPC molecular modelling, would Canadian firms be able to secure the world mandate to conduct such research in this country?

As a second example, it is understood that *integrated product development and optimization* in manufacturing does have a major impact on industry competitiveness. Despite the high costs involved in competing with major multinational companies in the implementation of these processes, there is a need to discover collaborative ways to share assets and resources in order to introduce this technology to Canadian companies.

The study has demonstrated that commercial applications of HPC technology, e.g. for managing and analyzing massive enterprise databases, are growing at a much faster rate than traditional scientific applications. One member of the committee predicted that soon there will be more parallel computers on Wall Street and Bay Street than in laboratories. Furthermore, there seems to be an enormous need emerging for powerful servers for multimedia applications in health care, education, and digital libraries on the information-highway. In Canada, where scientific R&D accounts for a small percentage of the GDP by international standards, there may be far more opportunities in these business and social applications than in traditional engineering and scientific applications.

Of the utmost importance is the need to address the special concerns of industry users. While the issues and concerns will vary to some extent by industry and by region, the following seem to be shared by all: the high cost of migrating from legacy systems, high communication tariffs over long distances, need for security, lack of suitable software tools and lack of trained experts. These concerns could be alleviated by a compelling set of demonstrations of positive economic impact.

2.2 Primary Focus: Dissemination

The HPC study confirmed that a major impediment to acceptance of HPC in Canada is an insufficient knowledge and appreciation of the economic benefits of HPC by business and government leaders. *Awareness must be extended* far beyond the few individual users and the members of the HPC advisory committee.

As described in 2.1 above, there is an urgent need to document successes both in Canada and abroad, with very specific explanations of the economic benefits. The next challenge is to disseminate this information as widely as possible, with the objective of building support for a national HPC direction.

The examination of successes in the U.S. and Europe amply demonstrates that universities and government laboratories will be willing participants in early knowledge creation and development of HPC methods, algorithms and tools, as well as pre-competitive versions of industrial applications. There is a need to diffuse this knowledge and expertise as rapidly as possible to industry by involving industrial participants in the early stages of the HPC development process. In the U.S. and Europe, HPC research centres normally focus their attention on the local industry requirements and provide awareness training to executives, technical training to industry scientists and engineers, development assistance for software prototypes and demonstrations, and ongoing support to any production users. These collaborative research and development centres may be public or private, but the vast majority are funded by government, with industrial clients providing a small percentage of the funding (typically no more than 10 to 15%). Corporations that become large users of HPC technology may eventually acquire their own equipment, or they will become paying customers of a public or private shared facility.

As HPC activity increases, there will be a greater need for *multi-disciplinary academic programs* in the universities that provide computational science and modelling training to students of specific science or engineering disciplines. HPC systems require very specialized knowledge of novel architectures such as parallel and vector processors, optimization techniques, visualization methods, networking, and new programming environments. More importantly, students must learn how to develop new algorithms and analytic techniques to take full advantage of the newest computational tools. So far, these complex techniques have been acquired ad hoc, on-the-job or have been self-taught. Consequently, it is very time-consuming and costly for industrial users to migrate applications to HPC technology. In summary, there is an industry need to train a new breed of HPC-literate engineers, scientists and businesspeople.

The time requirement for dissemination must not be underestimated and program horizons must be long enough to give industry the time necessary to become proficient in the use of HPC. For example, one of the most successful industry outreach HPC programs is at the NCSA Supercomputer Centre in Illinois. The director of industry partnership programs explained that the "proof-of-concept" phase typically takes about 18 months, and it is only after this initial training and demonstration phase that a corporation will make a serious commitment to develop new HPC applications.

2.3 Software: A Primary Consideration

The HPC advisory committee generally agreed that *HPC software significantly lags behind* the hardware technology. There is already a diverse choice of very high performance computer systems on the market, with very impressive "theoretical peak" performance ratings, but a real-world software application may be capable of performing only at a small fraction of this peak (often in the 3 to 5% range for new parallel computers).

The software challenge should be regarded from two distinct aspects. First and foremost, there is a need for much better systems software, which consists of programming environments, languages and a plethora of tools for fine-tuning, optimizing, trouble-shooting and, in general, automating the process of getting applications to run efficiently enough to make the HPC equipment highly attractive to the user. Corporations that develop HPC application software today are concerned that it takes far too long to develop an application with adequate performance — a problem that they attribute to insufficient HPC programming tools. Software-application development companies which may distribute their product globally, on a multitude of different HPC platforms, have the additional aggravation of dealing with incompatible architectures from different vendors. To counteract this problem, there are international research efforts to develop HPC standards and also to develop "portable" environments. The goal is to provide software developers with a programming environment that will allow for software written on one platform to be installed subsequently on virtually any other platform without modification.

The second software challenge is to develop actual application software, using the tools described above. Naturally, development will be less costly and products will emerge at a much accelerated pace once adequate development tools reach the market.

In Canada, there are approximately half a dozen small companies developing HPC-related system software and about the same number developing application software. The Canadian market for products will always be extremely small, so a software company will survive only if it manages to sell most of its products outside Canada (typically over 90% of software revenues come from export markets).

Canadians have proven in other software fields that they can compete with the best in the world. There does seem to be a very significant opportunity, especially for parallel computers, where the potential is enormous and the present impediment is a serious lack of software. As parallel systems are expected to find very important applications in *commercial database applications and multimedia information-highway servers*, there will be a multi-billion-dollar opportunity for new software over the next few years. Since this is still a new, emerging opportunity, no country has yet developed such a lead that others will never catch up (as is the case with hardware). Therefore, the time to act is now.

2.4 Need for Canadian Facilities

The point has been stressed in this report that the first priority in establishing an HPC presence in Canada is to develop the intellectual expertise. Until there is a heightened awareness of benefits and a pool of development experts, the demand for large-scale facilities will be minimal. However, *the Canadian community will need some facilities for training, pilot projects, demonstrations and production facilities for early adopters of HPC technology.*

Larger companies may eventually be able to justify setting up their own facilities, but even in the U.S. this is rare. Companies of any size can justify the acquisition of workstations, but since access to large-scale HPC facilities is normally required only sporadically, it makes sense to share the cost of the facilities among several organizations. Even the largest corporations will want to share the cost of facilities during the lengthy application-development phase, when access requirements are ad hoc and occasional. Such sharing arrangements have been difficult in the past because of technical issues such as inadequate network bandwidth, insufficient security and incompatibility with the local in-house computing environment. These problems can be resolved with today's technology, as described in the detailed HPC Study Report yet while the technical problems can be solved, Canada still has one impediment to shared facilities: long distance rates that are several times higher than those for comparable services in the U.S. This disadvantage has motivated some users to connect to facilities in the U.S. to avoid the excessive rates charged for a connection to a centre remotely located in Canada.

There will always be a debate among HPC purists regarding the type of technology best suited to a given facility and there will always be differences of opinion regarding preferred vendors. However, there seems to be unanimous agreement that *there should be a diversity of facilities in Canada and that these facilities should be shared across the Canadian CANARIE network.* This sharing of computer and intellectual resources, which minimizes the duplication of facilities and makes all resources available in a virtual machine room, is being called a *metacentre*. Each user on the network should have access to any of the diverse facilities, regardless of location, and ideally software should be developed that automatically determines the optimum facility for processing each user task at any point in time, according to the most suitable architecture and the greatest availability of computing resources at that instant.

2.5 Leverage via Existing Assets

With limited sources of funds available, supporters of a Canadian metacentre must examine existing resources and find ways to share these assets before committing to new expenditures. In this way, previous investments in facilities provide leverage. There are a few obvious opportunities to be explored at the outset.

For example, the *supercomputer facility at Environment Canada's Meteorological Centre* in Dorval, Quebec, is the largest HPC facility in Canada. Presently the centre is dedicated to weather forecasting and atmospheric research. Previous attempts to make the facility available to other users failed for two major reasons: first, there was no marketing or customer service, and secondly, external users were concerned that the weather processing took priority over any other work. Assuming that Environment Canada is interested in providing external access to excess capacity, perhaps both these factors could be overcome if this facility were connected as one of many nodes on the metacentre network, which by design will include national support resources and alternate processing nodes.

As another example, the *HPC High Performance Computing Centre in Calgary* has invested in a secure industry-focused infrastructure. There are opportunities to extend access to the Centre's facilities to other users or organizations across Canada.

What must be added to the metacentre is access to one or more medium-sized parallel machines. Again, existing assets should be leveraged by examining the possibility of working with the present owners of small parallel machines to place scaled-up versions on the metacentre network. Candidates for providing equipment access are research centres, such as CRIM, and universities.

Despite limited government funds, Canada must examine existing programs to find synergy between existing initiatives and HPC. Several opportunities have been suggested. First, within funding agencies such as NSERC and the Medical Research Council (MRC), it is critical to identify some mechanism to provide HPC *infrastructure support to university researchers.* Secondly, there are already HPC funding programs in certain provinces, e.g. Ontario (program under development) and Alberta; therefore, there needs to be *as much collaboration as possible between the two levels of government and between provinces.* Thirdly, there is a short-term opportunity for the Canadian HPC community to *submit HPC-related development proposals to CANARIE* for its next phase of funding support. In other countries there is a very close connection between high performance computing and high performance networking; in fact, in the U.S. the combined program, called HPCC, receives $1 billion per year of federal government funding. Initial indications are that CANARIE would be receptive to proposals that develop important HPC activity over the CANARIE network. Lastly, it will be important to determine how HPC programs can participate in other established pre-competitive collaborations in Canada.

It will not be economically viable to install the world's largest machines anywhere in Canada, so there may be occasions when it would be desirable to have access to a very large U.S. facility. Discussions have already commenced with the U.S. National Science Foundation regarding the linkage of the existing U.S. metacentre to a Canadian counterpart; such an arrangement would provide for occasional access to each other's facilities.

2.6 Need for a Long-Term Commitment

Another barrier to the acceptance of HPC in Canada has been the concern that facilities do not have a sufficient life span to warrant the investment, by potential users, in training and application development. However, this in fact becomes a self-defeating attitude, as centres do not survive if they do not have an adequate customer base, and users do not commit themselves to using a centre if they do not have adequate assurances that the centre will be viable over the long term.

The term "viable" deserves more comment. It is not sufficient to guarantee that a centre will simply exist in the long term — it is equally important that a commitment be made to a continuous technology-renewal program. Technology changes dramatically every two or three years, even at the high end of HPC. Therefore, a centre must keep its facilities reasonably current or customers will drop off.

In the early days of supercomputing, it was exceedingly expensive to renew technology, as each major upgrade involved a total replacement at a cost of several million dollars. In the future, with the trend towards scalable parallel systems, the upgrades will be in small increments at a relatively low cost; even upgrading processors will be much less expensive, as low-cost commodity chips will normally be replaced as new generations reach the market. While we are moving in this direction, there is still some way to go; therefore, in the next few years it will still be necessary to prepare business plans that accommodate upgrades and technology shifts. Potential users must understand the long-term commitments that are involved and, where possible, become long-term partners, sharing the risks and opportunities. For any HPC project to succeed, it will be essential to have well-articulated needs and commensurate long-term financial support.

3 CITI HPC STUDY — FINAL RECOMMENDATIONS

3.1 Promoting a Distinctive Canadian HPC Network

The review of Canadian high performance computing over the past eight months has concluded with a consensus amongst all participants that recommendations must *address the issue of the nation's international competitiveness.* There has been progress in HPC, but whatever has been done is inadequate, leaving Canadian research and development falling further and further behind our trading partners. Canadians enjoy a world-renowned standard of living, but maintaining that standard will depend on our continuing ability to compete in world markets.

The advisory committee was clear in recognizing that no one individual or organization can singlehandedly lead in establishing HPC as a major supporting technology. On the other hand, it was agreed that a "champion" would be very beneficial as a coordinator and a facilitator in moving initiatives along.

Collaboration will be key among all, participants, including government at both the federal and provincial levels, as well as universities and industrial users. Collaboration is also essential between current participants in HPC, others who are planning to expand their involvement, and the providers of networks, thereby ensuring cost-effective access. PRECARN, an industry-led consortium for pre-competitive research, should be examined from the perspective of possible HPC content in new initiatives that are currently being solicited from its members.

It was also concluded that a Canadian presence in HPC is essential, but it is not necessary to have the largest, newest most modern facilities, providing that there is access to these and to a wide variety of technologies when required, and at a reasonable cost. Where reciprocal needs can be identified, *links to large U.S. and European HPC centres* can be established and used for our largest HPC problems. Attention to the potential loss of Canadian expertise and risks of industrial security must nevertheless be considered.

Most importantly, the Canadian HPC network must also include the *expert intellectual resources* necessary to ensure that Canadians become effective users of HPC tools. Developing a pool of qualified computational scientists, accessible to industry and sensitive to their industrial needs, is a much higher priority issue than establishing new facilities. While we must train engineers and scientists on how to apply HPC to their own problems, we also require software experts and computational scientists who can provide an interface between the users and the complex technology, to ensure that optimal performance is obtained from a particular technology. Support resources must be service-oriented and situated both locally and throughout the network.

There are pockets of HPC expertise in universities and research laboratories scattered throughout Canada. Any Canadian HPC initiative must identify the resources that are currently available and establish a networking procedure that will put end-users, application developers, computational scientists and HPC software experts in communication with each other for the collaborative development of new strategic applications.

In our universities, there must be more *cross-disciplinary training* between computer science, engineering and physical sciences, so that we begin to produce students who can apply HPC methods to industrial problems. It was agreed that at this time our universities are producing excellent computer scientists who learn hardware architectures and software engineering, but we also must teach our scientists and engineers computational science and modelling methods so that experts in each discipline may develop effective HPC algorithms, models and programs.

The committee participants in the review also fully concerned with respect to Canada's present economic position. Any new initiative must be based on *leveraging current assets* in HPC and high performance networking (HPN), and redirecting funds from existing programs, not on new financial support from any specific source. This restriction also eliminates initiatives that had been suggested with respect to encouraging growth in HPC use through tax-incentive programs.

Industry and government collaboration in the CANARIE network initiative has provided Canadians with the opportunity to access new advanced services over a high-speed electronic network, from coast to coast. CANARIE's recently announced Technology and Applications Development Program (TAD) provides the HPC community with outstanding opportunities to submit proposals to provide matching funds for applications that demonstrate the economic and technical viability of a Canadian metacentre.

RECOMMENDATION

It is recommended that a Canadian metacentre network, partnering existing HPC centres, be established in collaboration with the federal government — sponsored CANARIE initiative, to provide HPC services for all regions of Canada. Such a network should provide access to a variety of support resources, software and HPC technologies.

Recommended Action

February 1995

- Identification of projects suitable for submission to CANARIE for the establishment of a Canadian metacentre network; identification of partners.

March 1995

- Preparation of preliminary CANARIE proposals, including project definitions, partnership support and funding commitments.
- Submission of CANARIE proposals.

April 1995	•	Development of an inventory of available HPC experts, including computational scientists and software experts; discussion of ways in which these experts could collaborate in the metacentre.
	•	Development of an inventory of current assets for potential sharing in the metacentre.
May 1995	•	Commencement of dialogue with other organizations that may be able to participate in a metacentre — e.g. PRECARN, CRIM, NRC, NSERC, MRC, ORTECH, HPC Calgary, universities.
June 1995	•	For March 1995 CANARIE proposals that make the short list, preparation and submission of final proposals.
August 1995	•	Commencement of implementation of any approved CANARIE projects.
	•	Commencement of implementation of any other projects that may be able to proceed without CANARIE funding.
September 1995	•	First facility and support nodes available for access through CANARIE.
October 1995	•	Second facility and support nodes available.
December 1995	•	U.S. (and possibly European) metacentre links operational.
January 1996	•	Metacentre Phase I fully operational.

3.2 Building greater Canadian Awareness of HPC

To increase the awareness of HPC, it will be necessary to more effectively *disseminate information* to large numbers of potentially interested individuals in industry, universities and government.

In the interest of "market pull", Canadian industry will require specific examples of the financial and competitiveness benefits to be gained from the use of HPC. As well, knowledge of the potential national economic benefits can encourage both partnerships and government support. Once the demand is established, "education push" will occur, leading to the development of courses and programs in universities and other educational institutions.

An expanded awareness of new opportunities to develop applications, especially health-care, social and business or commercial applications, will encourage industry to move toward solutions that involve the implementation of HPC resources.

The enhanced accessibility and availability of the HPC technologies and support resources recommended above will in themselves help to disseminate HPC information and allow for a user community to have the awareness necessary to be weaned from their workstations to more effective HPC systems as problem demands grow. The considerable time this process takes has already been demonstrated in the U.S. and in Europe.

Science, engineering and information technology associations across the nation have successfully developed information-distribution and awareness programs. HPC awareness can be served through *alliances with these associations*. Super*Can, the Canadian Association for High Performance Computing and Communications, should be considered a focal point for the HPC communications strategy. For example, this organization already distributes newsletters and is planning an annual magazine on Canadian HPC success stories. For Super*Can to be effective in a broadened awareness program, it would need to expand its membership, which currently consists mostly of university researchers. Cooperative efforts with national organizations such as ITAC (Information Technology Association of Canada) and CATA (Canadian Advanced Technology Association) should be explored. Regional and local economic development authorities may also be potential partners to promote the use of HPC facilities to stimulate their economies.

RECOMMENDATION

It is recommended that the current CITI program be continued in order to foster better understanding of the economic benefit and opportunity, with broader diffusion of information throughout government, universities and industry.

Recommended Action

March 1995

- Identification and verification of a minimum of 10 national and international case studies for detailed economic, societal and/or scientific analysis.

June 1995

- Analysis and documentation of a minimum of 5 exceptional cases.
- Introduction of HPC to a minimum of 8 national associations and/or organizations through letters and visits to national executives.

December 1995

- Major presentations at conferences or major articles published in journals of a minimum of 4 national associations/organizations.
- Identification of a minimum of 6 media opportunities associated with national events involving industry, academia and government, where an HPC message is appropriate; participation therein, as appropriate.

- Publication of a minimum of 20 HPC-related "introductory articles of interest" in professional journals, magazines and national newspapers.

- Presentation of a minimum of 10 seminars on HPC topics to interested groups in government, education and industry.

3.3 Expanding Canadian Software Opportunities

Software is a key element in the successful implementation of HPC and offers expanding opportunities for products on world markets for Canadian developers. Canadian researchers involved in HPC software development activities are held in high esteem by their international peers. In software areas somewhat related to HPC, Canada has had some amazing successes in companies such as CAE, ISG Technologies, Alias Research and Softimage. There is an excellent potential for leveraging the significant Canadian talent to a major international advantage.

The experience of successful Canadian HPC users shows that there is little risk associated with HPC hardware. Engineers continue to extract more and more capability from the technology. The risk lies with the ability of the software to use current and emerging hardware technology effectively. Software engineering has not kept pace with hardware developments.

It is agreed that *software development should be a primary consideration* for Canadian HPC, with concentration in both the development of applications in promising new environments and the creation of support tools to improve the productivity of developers. There is also a very large demand for software tools that will assist users in migrating their millions of lines of legacy code to new HPC architectures, and to then optimize the execution of these codes on the designated HPC hardware. Participation in the *development and incorporation of international HPC standards is critical.* This participation will lead to an opportunity to create tools that allow for application software to be developed in a manner that ensures portability between a diversity of HPC platforms.

It is also agreed that software development companies in Canada will identify where they can most successfully compete with their software in the international marketplace, without the intervention of government. However, canadian software developers tend to have difficulty in the sales effort and in the provision of user support services associated with these enormous markets; therefore some *assistance in international marketing* would benefit many of these companies.

The application of HPC in commercial enterprises is occurring very quickly, with international software opportunities estimated in billions of dollars per annum. This rate of development suggests an urgency for Canadian software developers to become involved in *demonstrations and pilot projects* within the Canadian metacentre defined above. Projects should include the major data applications, including multimedia services, essential for the operation of government and business, where massive volumes of essential data are made available to a widely spread user community for a broad number of applications. Governments can encourage local software developers by becoming model users or early adopters of new HPC software that may be applicable to government applications. Perhaps the new data warehousing and information-highway server applications would be an excellent target for government support through partnerships and early adoption.

RECOMMENDATION

It is recommended that the structure of a Canadian HPC network provide expanded software-development awareness and opportunities for Canadian industry.

Recommended Action

March 1995 • Inclusion of at least one software development, unique to a Canadian metacentre, in CANARIE partnered projects.

June 1995	•	Identification, in collaboration with provincial and federal governments, of at least 3 information-highway server applications that could be developed by Canadian software producers.
September 1995	•	Presentation to HPC software developers of the available government opportunities, and encouragement of collaborations to implement applications where there is mutual interest between the user and the developers.
December 1995	•	Identification of at least two other software development opportunities, of national and international significance, for Canadian development.

CONCLUSIONS

The importance of HPC technology for the advancement of research and the enhancement of industrial competitiveness has been highlighted by this study. A clear and succinct summary of the conclusions of this report are as follows:

A) Applications and software development should be the primary focus of Canada's efforts in HPC.

B) Use of public funding to augment Canada's HPC infrastructure should focus on the creation of a network-accessible *metacentre* that leverages existing investments.

C) Long-term development of Canada's academic and industrial capability in HPC should focus on education, outreach and application demonstrations.

APPENDIX

APPENDIX A

HPC Technology Summary

In the early days of HPC, large, challenging, scientific problems required access to large, expensive supercomputers that used a technology that utilized *vector processors* to speed up execution.

HPC today encompasses a wide range of technology that allows for a broad spectrum of applications. First, at the low end, desktop *workstations* have advanced so rapidly that they are now capable of solving some of the problems that previously required access to costly supercomputers. For all applications, workstations are now the "window" for access to an ever-expanding range of high performance computers. By means of a desktop workstation, a user will prepare the work and then "visualize" the end results on the same workstation, regardless of where the problem is directed for high performance processing or information access. Secondly, since most workstations generally remain idle a good part of the time, techniques have been developed to network multiple workstations in "clusters" so that a single large problem can be partitioned into smaller segments that run concurrently across multiple machines, taking advantage of the idle time in the network. Thirdly, there has been the emergence of *massively parallel processors* (MPP), which give the promise of orders-of-magnitude price-performance improvements over the traditional vector supercomputers for many applications. Lastly, there are also systems generally referred to as *multiprocessor servers*, which provide yet another way of combining workstation elements into a single system to solve larger problems. For the next several years, these different technologies and approaches will be complementary, as at this stage, no one technology can adequately solve all classes of problems.

One of the important conclusions of this study is that Canada should work towards the establishment of a *metacentre* which involves the networking of HPC facilities and human resources so that a "virtual machine room" is made available to users throughout the country, with access to technical support on the network regardless of the physical location of the technical-support staff. Lastly, HPC must include the software and the networks that make the HPC hardware accessible and easy to use.

APPENDIX B
Comparisons of the HPC Situation in Canada with that in the U.S., Europe and Japan

1. Comparison of HPC in Canada and the U.S.

There are several factors that have led the U.S. to its dominant position in both the development and use of HPC systems. Initially, in the 1970s, it was understood that supercomputers were necessary tools for maintaining U.S. dominance in strategic weapons — especially nuclear — and in space programs. There were no equivalent stimuli in Canada to encourage the adoption of supercomputers.

As a result of the military support of R&D in supercomputing, the universities began to understand the benefits of HPC to basic research. By the late 1980s, the National Science Foundation (NSF) Supercomputer Centres were approved and university access to HPC became widespread as NSFNet unfolded. Canada started down this same path with provincial government support of supercomputer centres in Ontario and Alberta, both in the late 1980s. However, while the U.S. centres flourished with continuous federal government support of technology renewal, the two Canadian centres ceased operations in 1991 because of an absence of ongoing provincial funding and a complete lack of federal support. Both centres were technically obsolete when they shut down for lack of funding, while the U.S. had meanwhile opened dozens of state and national centres with annual programs of technology renewal.

The eventual establishment of the national *High Performance Computing and Communications (HPCC) program* has placed the U.S. in the forefront of all current HPC technology developments and is starting to pay enormous dividends in terms of improved industrial competitiveness. This federal HPCC initiative currently provides approximately *$1 billion annually* for a variety of HPC programs. The combination of these federal funds, with what is probably an equal level of funding from state governments, represents over $6.00 per capita in annual HPC funding, as compared to 25 cents per capita in Canada.

In the U.S., it was only after the establishment of university HPC centres that *industrial partnership programs* began to emerge for the diffusion of HPC technology to the private sector. After several years of education, training and industry outreach programs, some U.S. companies developed important R&D partnerships with state and national HPC centres, while others eventually developed sufficient justification to install their own supercomputers. Canada has not graduated to this stage of technology transfer from university to industry, and at present there is no infrastructure support to enable the charting of such a course.

Several important lessons can be learned by studying the U.S. experience. First, an ongoing *long-term commitment to funding and technology renewal* is essential for building and sustaining a national HPC program. Industry will not consider investing in the development of HPC applications until an infrastructure is in place, including facilities, networks for easy access, and qualified experts for training and development assistance. Secondly, as individual organizations, especially small and medium-sized enterprises, cannot justify setting up their own dedicated HPC facilities, it is *important to establish shared centres*; even very large organizations begin building knowledge and experience at such centres. These centres must offer a diversity of equipment (for example parallel and vector processors) and application services. Thirdly, *awareness programs, training, education and industry outreach programs are critical to success.* Lastly, one can observe a rapid development of HPC technologies to support commercial applications and multimedia applications.

2. Comparison of HPC in Canada and Europe

There are more parallels between Europe and Canada than there are between the U.S. and Canada. The major reason for this similarity is that Europe comprises many countries, each on average being closer in size and economic development to Canada than to the U.S. These countries, until the recent *High Performance Computing and Networking (HPCN) Program* of the EC, have operated almost autonomously with respect to the development and implementation of HPC technology. There has been no major military imperative similar to that which gave the U.S. its head start. The other similarity with Canada is that Europe has never succeeded in developing a globally competitive computer-manufacturing industry, even though a few modest attempt have been made. Europe has clearly made a conscious decision to withdraw from any serious attempts to manufacture HPC hardware, while *concentrating all its efforts on software R&D.*

The major difference between Europe and Canada is that the EC now has a full-fledged HPCN program which will pump *$1 billion ECU's per year* in development of high performance computing and communications applications, again representing 10 to 20 times the government funding support in Canada, on a per capita basis. By bringing the EC countries together in a collaborative effort, Europe can compete on the same level as the U.S. and Japan, with each individual country receiving the benefits.

3. Comparison of HPC in Canada and Japan

Of all three major economic units (U.S., Europe and Japan), a comparison of Canada with Japan probably yields the fewest parallels or similarities, and Japan is probably the most difficult model to emulate. These difficulties are the result of major cultural differences, along with very different business practices and work ethics.

The first major difference is that Japan decided several years back that information technology was to be a major strategic target industry for development. This has resulted in Japan's becoming the only serious competitor to the U.S. for HPC semiconductor and hardware-systems manufacturing. Countries that manufacture supercomputers tend to establish programs for their application in domestic markets much faster than countries that only import the technology. Consequently, Japan and the U.S. have been the most successful countries in the world at developing industrial HPC applications.

Secondly, since the Second World War, Japan has been building high-tech industries with enormous investments in R&D (an R&D-to-GDP ratio of 2.87%, compared to 1.5% for Canada). Computers, and specifically HPC, have been long recognized as *a critical tool to support new product design and development.* Japanese researchers applied HPC to industrial R&D far more aggressively and far sooner than the U.S. (currently *65% of supercomputers in Japan are used by industry*, compared to 50% in the U.S.).

Lastly, Japan has found a way for HPC techniques to be applied to R&D in small to medium-sized industries. In all of these ways, the Japanese situation is most unlike the situation in Canada, even with respect to applications of the technology. While Canada has no aspirations to design or develop HPC hardware, perhaps we have much to learn from the Japanese about the application of the technology to industry.

4. Opportunities for Canadian — U.S. Collaboration

During this study, there have been discussions with the National Science Foundation directors responsible for the U.S. multi-billion-dollar HPCC initiative. The NSF was very interested in further exploring the possibilities and ramifications of Canadian — U.S. collaboration in HPC. NSF representatives stated that there is already a well-developed collaboration of university computer-science departments in the two countries. They can visualize a more extensive relationship whereby *Canada would become the first international extension of the U.S. Metacentre Program.*

In terms of sharing computer resources, some Canadian university researchers have already made their own arrangements to access U.S. HPC facilities. However, these are very informal arrangements, subject to termination at any time. Also, there are few or no support services offered in these ad hoc arrangements, nor are there any provisions for high-speed network access. A formal agreement with the NSF may open the possibility for Canadian researchers, at least in the universities, to have access, with some support services, to the high-end equipment that is not available in Canada. U.S. researchers, on the other hand, apparently have access to all the computer facilities they require; however, there appears to be an interest in gaining access to the Japanese supercomputers in Canada, as U.S. policy prohibits acquisition of Japanese supercomputers by government and universities.

In terms of U.S. collaboration on sharing resources, it is not likely that there could be any arrangement that satisfies industry. For example, U.S. industry has no incentive to allow foreign companies to access U.S. advanced technology which is currently providing a competitive edge to U.S. companies.

APPENDIX C

List of Advisory Committee Members for the CITI study of HPC in Canada

Dr. Andrew Bjerring
President
CANARIE Inc.
410 Laurier Ave. West, Suite 1120
Ottawa, Ontario K1P 6H5

Mr. David Burden
Senior Director, Information Services
Air Canada
ZIP-045, P.O. Box 9000
Saint-Laurent, Quebec H4Y 1C2

Dr. Arthur Carty
President
National Research Council
Montreal Road
Ottawa, Ontario K1A 0R6

Mr. Paul Davis
President & CEO
HPC High Performance Computing Centre
3408, 400-3rd Avenue S.W.
Calgary, Alberta T2P 4H2

Dr. Kirk Dawson
Executive Director
Canadian Institute for Climate Studies
650 Hastings Avenue, Suite 765
Penticton, British Columbia V2A 7N1

Mr. Gordon W. Gow
Assistant Deputy Minister
Ontario Government, MEDT
Technology and Communications
56 Wellesley Street W.
Toronto, Ontario M7A 2E4

Dr. Inge L. Hansson
President
Ashurst Technology Centre, Inc.
2486 Dunwin Drive
Mississauga, Ontario L5L 1J9

Dr. Kevin Keough
Vice-President, Research
Memorial University of Newfoundland
St. John's, Newfoundland A1C 5S7

Dr. Cooper Langford
Vice-President, Research
University of Calgary
2500 University Drive
Calgary, Alberta T2N 1N4

Mr. Andrew Macdonald
Chief Informatics Officer
Information Management, Systems & Technology
Treasury Board Secretariat
Ottawa, Ontario K1A 0R5

Ms. Elizabeth Pearce
President, Super*Can
51 Sandringham Drive
Toronto, Ontario M5M 3G4

Dr. Denis Poussart
Vice-President
CRIM
1801 McGill College Avenue
Montreal, Quebec H3X 2H9

Mr. Hugh Stanfield
President & CEO
Pulsonic Corporation
510, 400-3rd Avenue S.W.
Calgary, Alberta T2P 4H2

Mr. John Taylor
Vice-President Engineering
Bombardier Inc., Canadair
P.O. Box 6087
Station Centre Ville
Montreal, Quebec H3C 3G9

Ms. Harriet Velazquez
Vice-President, Delivery Programs
Canadian Imperial Bank of Commerce
901 King Street West, 7th Floor
Toronto, Ontario M5V 3H5

Dr. Donald Weaver
Dept. of Chemistry and Neurology
Queen's University
Kingston, Ontario K7L 3N6

DATE DUE
DATE DE RETOUR

1575, boulevard Chomedey, Laval (Québec), H7V 2X2 Téléphone: (514) 973-5700 Télécopieur: (514) 973-5757

1575, Chomedey Boulevard, Laval, Quebec, H7V 2X2 Telephone: (514) 973-5700 Facsimile: (514) 973-5757