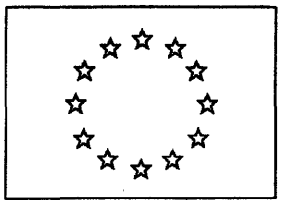
III/6072/94/EN.



Information Technology

High-Performance Computing and Networking

Report of the HPCN Software & Systems Industrial Working Group

" HPCN enables us to improve the way we work"

April 1994

Report of the

HPCN Software & Systems

Industrial Working Group

April 1994

GROUP MEMBERS

Mr. T. Bemmerl	Intel
Mrs. P. Blanchot	SEMA
Mr. J. Cholley	Telmat
Mr. M. de Lange	ACE
Mr. R. Fortier	Steria
Mr. R. Geisen	Parsytec
Mr. W. Gentzsch	Genias
Mr. P. Isambert	Matra Cap Systèmes
Mr. W. Mayr-Knoch	Siemens Nixdorf
Mr. M. Platt	IBM
Mr. A. Lichnewsky	ACRI
Mr. C. Skelton	ICL
Mr. K. Solchenbach	Pallas

EXECUTIVE SUMMARY

The HPCN Software and Systems Industrial Working Group was set up in January 1994 to identify future needs in the demand and supply of high-performance computing and networking (HPCN) systems and to recommend the actions to be taken at the Community level.

Numeric and commercial applications have already started to take advantage of HPCN technologies. Numeric exploitation of advanced parallel processing has had faster take-up because users are generally technically well-informed engineers and scientists. Commercial users, who have built information management and processing into their mission critical business systems, must be more cautious about moving to new technologies. A recent survey has shown, however, that 80% of high-performance computers are used in business applications. Both these markets are equally strategic for Europe. Until recently HPCN technology has been too expensive for widespread use in embedded systems, however the new HPCN technologies bring affordable power and equal emphasis should now be given to this market. This report, therefore, examines the industry needs and the priority underlying technologies of the three application areas: numeric (or more specifically simulation and design), information management and processing and embedded systems.

HPCN covers a broad range of high-performance computing platforms, including parallel machines, super computers, workstations and clusters, in networked, embedded or standalone configurations. These systems are fundamental components of HPCN. European vendors are currently in a strong position to compete world-wide but investment must continue to maintain this position. The Framework IV HPCN workprogramme should support emerging technologies and innovative applications thereby increasing the know-how and IPR of European industries.

The systems-level and application-level performance and price/performance perceived by users is impacted by many factors other than raw processing power, including the handling of data storage, access methods and interpretation, and the management of systems.

Users are very wary of being locked into proprietary systems and are increasingly demanding open systems and conformance to industry standards. There are many examples related to the HPCN area, including POSIX, UNIX, Fortran 90, ANSI C, C++, SQL, MPI message passing, HPF Fortran, and ATM. It is considered very important that European projects build on the appropriate standards.

There is much good HPCN research in Europe. The priority now is to exploit this research in revenue earning products. Product life is reducing and is now typically 2 to 3 years. Vendors must exploit new technology (particularly chips and disks) as it comes on stream, otherwise they will not remain competitive. Priority should be given to projects where European industry can take or maintain a lead and which show industrial-strength results in a short time frame, say two years.

The IV Framework HPCN workprogramme needs to focus on applying funding to high-leverage and high-potential areas. Criteria for such areas are the ability to develop European know-how, IPR and engineering organisations.

The exploitation of HPCN technology by European users is still insufficient and a priority action should be awareness: from industry relevant demonstrations, to training and marketing.

CONTENTS

1. MANDATE OF THE GROUP

The group comprised 13 representatives from IT Hardware, Software and Services supply industry.

In the context of the IT programme in the 4th Framework Programme for research and technological development, the group's mandate was:

- gather expert's advise to identify future trends and needs in the demand and supply of HPCN systems and to specify what kind of support would be required at the Community level.

- deliver a document which will directly contribute to the elaboration of the workprogramme document for next call.

2 SCOPE OF THE DOCUMENT

The priority for the group has been ***industry need.*** The availability of affordable high performance computing and networking, creates an opportunity to profoundly impact the way engineering, manufacturing and business is done.

The industry needs for HPCN systems are discussed under ***three stream headings:***
- Simulation and Design: section 3.
- Information Management and Processing: section 4.
- Embedded Systems: section 5.

The ***priority technologies*** underlying these streams are detailed within these sections. However, many of the technologies are applicable to more than one stream and these are included in section 6: Generic Technologies.

The uptake of new HPCN technologies is partly limited by lack of awareness and the availability of suitably trained people. Recommendations for an Awareness Programme and for ***Technology Transfer*** and Training are therefore made in section 7.

3 SIMULATION AND DESIGN

3.1 Industry Needs

An example of industry's reliance on HPCN can be found in the March 1994 annual report of a major European aerospace company, which quotes the following.

'The Research Centre is studying the capacity of the human brain to handle multiple tasks and how those tasks can be redesigned to match the characteristics of pilots, drivers and other operators. The enormous expense of practical demonstrations of new systems requires that they should use computer power to examine behaviour without building hardware. The Centre maintains its commitment to mathematical modelling for aerodynamics, electromagnetic fields, structural behaviour and many other areas of new technology. Doing this effectively requires very powerful computational facilities and the company continues to invest in advanced "super-computing" as a cost effective tool and an alternative to expensive test facilities and programmes.

For the same reason the company has substantial programmes in virtual reality and simulation which allows the exploration of these concepts safely and effectively.

In all this work the concern is not for the technology itself but finding ways of exploiting advanced technology to create business benefit.'

Simulation and Design is a major European strength. The competitiveness of substantial parts of the European industry depends on the use of simulation and design techniques. Key industries in this category relevant for the success of the European economy are the automotive, aerospace, chemical, pharmaceutical, environmental, electrical and power supply, machinery and mechanical industry. However, this only can be an incomplete list and design and simulation techniques are starting to be relevant in business and service industries too. A decrease of lead time in manufacturing, decrease of development time and a general decrease of time to market are the primary benefits of the before mentioned industries through simulation and design. However, the lack of applications is a major reason why parallel computers are not more widely used by industrial users. Among these applications and computational techniques are Computational Fluid Dynamics, Molecular Modelling, Seismic Processing, Reservoir Modelling, Quantum Chemistry, Chip Simulation, Computational Physics, Imaging, Crash Analysis and Structural Analysis. It is crucial, therefore to encourage software vendors and in-house code developers to port their applications to parallel systems. Only successful interdisciplinary projects with sufficient industrial relevance will guarantee that HPCN systems dominate the design labs of the automobile, aircraft, chemical, pharmaceutical, power supply, environmental, machinery and mechanical industry.

An evolutionary improvement of existing HPC technology needs to be followed which provides a performance increase compared to today's systems by an order of magnitude. With a focus on improvements in software technology a top down development methodology needs to be followed starting with the industrial application requirements. Therefrom industrial strength, stable and mature system software technology needs to be developed for general purpose, multi-user HPC systems. Industry quality programming models and program development environments are required to provide the underlying software and hardware technology for a successful migration of a huge application base onto HPC platforms. These programming aids include implementation of emerging programming models (message passing, shared virtual memory, data parallel), parallelisation tools, libraries, language compilers, debuggers, profilers, performance analysis tools, visualisation tools, system administration tools, etc. Some are already available and should be more widely used. Examples are the message passing libraries as PARMACS and the newly emerging MPI standard. Relevant languages include Fortran 77 and 90, HPF, C and C++.

Sequential and parallel versions of the programming environments are required to support single-node oriented development and system oriented development on HPC platforms. Standard operating system technology (UNIX, POSIX, Windows/NT compliant) needs to be enhanced with increased scalability, availability and reliability to be suitable for the evolving HPC systems used in simulation and design applications.

Some significant scalable OS work for real massively parallel systems is in progress at the moment and, as can be learned from most MPP manufacturers, the existing and available kernel technologies do not fully meet MPP requirements. The implementations of such scalable operating systems are following new OS paradigms while still supporting current standards and interfaces, such as POSIX.

3.2 Simulation and Design Technology Priorities

3.2.1 Applications

A straight forward porting of some industrial applications is covered by the existing Europort activity. These, however, mainly focus on the core kernels (e.g. solvers) and do not include the adaptation of the pre- and post-processing phases (model building and interpretation of results). All of the critical phases of a complex code, starting for example from the CAD process up to the animated processing are likely to benefit from parallel processing.

HPCN has the potential to improve not only the efficiency of industry but also its impact on the environment. Examples are energy consumption, efficiency of power plants, environmental protection and industrial plant design. Proposals for the development of new applications could include, for instance, a combination of methods in molecular design, creation of more advanced techniques for designing and improving environmentally and economically relevant industrial processes and improvement of methods to optimise metal working and metal forming processes.

3.2.2 Tools

The availability of suitable tools has a big impact on the introduction of HPCN systems in design and simulation applications. Tools important for other areas as well are included in Section 6. Here we concentrate on priority tools specific to design and simulation.

Application driven development and tools
Tools to understand and characterise the workloads which industrial applications put on HPCN systems at the processor, node and system level are key for application migration. System and application characterisation through benchmark kernels is a well known technique to be used in such tools. These tools will provide information on the influence of architecture and implementation parameters as communication/computation ratios, computation/I-O ratios, message flows and memory access patterns on the parallelisation strategy.
For the application parallelisation process itself, automatic mesh decomposition tools for static meshes (structured and unstructured) as well as for dynamic meshes (e.g. with mesh refinements) are required. Tools for the integration of mesh generation and decomposition in CAD standard environments in addition to the integration of standard post processing tools and mesh decomposition tools in a distributed post processing environment are needed.

Tools for the simplification of parallel programming
Use, evaluation, characterisation and qualification of existing parallel programming tools and environments on the basis of features, programming model dependency, architectural dependency and intrusiveness need to be the first step. On the basis of these real-world evaluations enhancements and extensions to existing industrial strength parallel programming tools as debuggers, performance tuning and visualisation tools, profiling tools and porting tools are to be developed.
Creation of a tools infrastructure to simplify the adaptation of parallel programming tools to evolving architectures, programming models and languages is a high industrial requirement. This will lead to debuggers, performance tools, application porting tools for new architectures and evolving programming models as fast message passing, shared virtual memory, data parallel models and combinations of these. Improvements of the scalability of instrumentation as well as display techniques in debuggers, performance tuning and visualisation tools up to thousands of processing nodes are the final step.

3.2.3 Platforms

Many of the requirements for platforms (hardware plus operating system) are similar for simulation/design, information management and processing and embedded applications. There are some differences in emphasis and this is expected to be reflected in project proposals. As an example, where commercial MPP systems measure parallelism in 100s of processing elements, simulation and design requires 1,000s of processing elements. Another example is data storage, where an information decision support application will want parallel access to vast amounts of data, but a simulation/design application may put more emphasis on the computing, networking and visualisation.

Scalability and Programming Models
Efficient implementation of scalable and easy to use programming models up to thousands of processing elements. Combination of message passing, shared memory and data parallel programming models as well as their integration into UNIX or POSIX compliant standard operating systems.

Availability, Reliability
Extension and enhancement of monolithic, micro kernel and server based HPC operating systems with features for increased availability and reliability. Fault-detection, redundant use, replacement and software reconfiguration during run-time are features to be added to today's operating systems of HPCN systems.

Network Integration, I/O
Operating system extensions to support improvement of multi-user capabilities and the integration of parallel systems into heterogeneous, distributed computing environments. Features to be worked on are resource management, multi-user scheduling, network queuing and accounting. Improvements of parallel I/O capabilities of HPC systems are a prerequisite for evolving applications. In particular software extensions in operating systems for support of parallel file systems and new network interfaces are required.

4 INFORMATION MANAGEMENT AND PROCESSING

4.1 Industry needs

Many businesses are now critically dependent on their IT systems and no longer have credible manual fall back systems. These businesses include finance, retail, manufacturing, public utilities, government administration, travel and the media. Many of these businesses are very conservative and place higher priority on the ease of migration of an existing mission critical application to new technology systems and other factors, such as security, than on raw performance. Businesses have grown to expect near 100% levels of availability and data integrity, ease of operation and professional support services. This level of service has taken 30 years to develop in mainframes but is required at day one with the new HPCN technologies.

Information highways based on fibre technology are promising to open up new areas of information management and processing. Domestic users are expected to welcome the increased choice of entertainment on demand, for example video, but on the back of this numerous other services, such as retail, are expected to follow. This will lead to enormous demand for networking bandwidth, multi-media data storage, sophisticated access methods and computation power.

High performance in information management and processing was achieved in the past by mainframes, with relatively dumb terminals. More recently PCs and mid-range machines, often UNIX based, have become a major influence. The industry has now established clear roles for these components in client-server architectures, where PC and workstation clients are networked to a variety of servers. The front end clients often have very sophisticated graphical user interfaces (GUIs) and application codes. This is leading to the requirement for HPCN technologies throughout the range from large scale servers to PCs and workstations.

Frequently, in commercial applications, very large numbers of clients are requiring access to a range of services. These services are mounted on servers, which can include a variety of databases and applications. Increasingly these servers are becoming specialised. For example massively parallel processing (MPP) machines are particularly suitable for database management servers but powerful single processor or shared memory (SMP) machines may be better for certain batch operations.

HPCN systems may need access to up to terabytes of multi-media data. This is leading to new pressures on data storage, decision support data-mining techniques, and on the administration and security of very large stores.

An important aspect of client server architectures is the networking. Emphasis should be on exploiting emerging new technology standards. Examples are FDDI and Fibre Channel in Local Area Networks (LANs) and ATM Broad-Band in Wide Area Networks (WANs). Priority should be given to the efficient integration of HPCs with LANs and WANs.

4.2 Information Management and Processing Technologies Priorities

4.2.1 Applications

The first priority is for demonstrators that convince users that HPCN is viable for them. Proposals will be welcome from a wide spectrum of domains, including finance, retail, manufacturing, public utilities, government administration, travel and the media. Demonstrator projects should be led by the users who will be expected to bring to the project their real world applications. Examples range from the migration of existing workloads, such as high performance financial transaction processing, to new applications enabled by the emerging HPCN technologies, such as large scale interactive broad band applications, typified by video on demand for the domestic market, and fraud detection in finance.

4.2.2 Tools

The availability of suitable tools has a big impact on the introduction of HPCN systems. Tools important for information management and processing but also for the other areas of simulation and embedded systems are included in section 6. Priority tools specific to information management and processing are as follows.

Application Migration
Many applications are written in COBOL with embedded databases. In many cases it is not today cost-effective to move them to a parallel machine, but there are cases where this is necessary. For example, it is first necessary to split the application into its client server components, then migrate the database onto the parallel server. This may also involve converting the database from a flat file, or Codasyl DB, to a Relational DB. Another solution might be to use a parallelising COBOL compiler.

Various tools are required for the porting of these applications, including data migration, modelling of workloads, system sizing and performance assessment.

Tools and techniques are required for the (semi)automatic parallelisation of an application, through detection and exploitation of inter-node as well as intra-node parallelism.

New Applications
Information management and processing applications can generally be classed as Transaction Processing (TP), Management Information (MIS) or Batch. Tools are required for splitting up, assessing, re-distributing and migrating the various components. The different components may end up on different servers. Tools are required to help build, debug, measure and tune new applications. These tools should preferably abstract from the precise details of the architecture so that the applications are not tied to a particular system, or even a parallel system.

Database management systems (DBMSs)
Organisations are identifying the requirement for extracting useful information from terabyte sized data bases. Generally they will look to systems like ADABAS, Sybase, Informix, Ingres and Oracle etc.. There are complementary tools where European suppliers can compete. These will include DBMSs and access methods for specific domains, such as retail, finance, travel etc., and extensions for complex queries, high rate transaction processing, multi-media and parallel processing. Tools are also required here for performance predictions, to analyse and tune, data administration, security etc., associated with large scale databases running in parallel.

Systems Management
Large scale commercial systems are expected to run virtually non-stop with little operator attention. Systems management covers facilities to hide the complexity of HPCNs and services. Facilities are required to optimise configurations, balance the utilisation of critical components, and report on the status of all components and flag problems, often over remote links. Many of these facilities exist for mainframe computers but there are currently no large scale market leaders in the area of Systems Management for MPP systems. Europe has a good research base and European companies are in a good position to exploit this market.

The priority areas where new, additional or enhanced Systems Management Tool developments are required are as follows.
Operations. Users want a 1000 processor system, each with its copy of the operating system, to behave as a single system. A large parallel server will probably have many different databases on it: a development DB, one for MIS and another for On Line Transaction Processing (OLTP).
Capacity. It is necessary to display system resources and their utilisation and to model the effect of changes.
Configuration. The user wants the ability to move or allocate additional resources to balance the system. This must be possible without interference to the live system.
Problem. In the unlikely event of a disc or processing element failure, it is necessary to automatically re configure, perhaps bringing spare discs on line and replication the data again. To achieve this, reports from all components must be collected and displayed,
Tele-diagnosis. Difficult problems may require attention of skilled specialists, often from remote sites. Users now expect systems to offer remote access by these specialists.
Resource accounting. It is often necessary to be able to allocate computing costs to individual users and these costs depend on the utilisation of critical components. Accounting packages are required for this.

4.2.3 Platforms

The operating system is a critical component in HPCN. Improvements in performance and functionality are urgently required and conformance to open standards is mandatory.

European software suppliers are in a strong position to influence the emergence of world wide HPCN operating system industry standards. Proprietary OSs will continue to be used for traditional mainframe computers, but UNIX and Windows (NT) are emerging as the industry standards for servers and clients. Work is in progress, supported by the OUVERTURE project, to integrate the Chorus micro kernel with UNIX SVR4, and within EPOCH to produce a parallel version of this: this work is recognised as world class.

In Europe, some very advanced work is being done in the area of operating systems that are scalable to massively parallel systems (1000-up processors). This needs support if Europe wants to be a player in the IT industry. Significant European OS work includes Chorus, Amoeba, PARIX, Idris, PAROS, peace etc.

The underlying OS technology development priorities for information management and processing systems are as follows.

Client server and distributed architectures.
Client-server architectures give an organisation considerable flexibility in the allocation of computing resources. Graphical User Interfaces (GUIs) and front end user tools are best placed on the desk. Other global services, such as mail, databases, data-mining, etc., are best placed in servers. This places special requirements on certain components of the OS, such as communications, within and external to the MPP system. All OS projects should include analysis of the overall systems architecture for the user domain.

Multi-user environments.
An important requirement for information management and processing is the ability of up to 1000s of users to access server resources. OS developments are required to manage these multi-processes.

Availability.
Availability includes reliability, resilience and recovery. OS developments to detect, isolate and replace failing hardware and software components are required.

Integrity and Security
The data integrity and data security of current MPPs are poor compared with mainframes. Priority work is required on both of these if European MPPs are to take mainframe like business.

Single System Image
Each user (user space) will access many processing elements. Likewise an individual file space will access multiple discs. There are current developments (e.g. in OUVERTURE) on small numbers of loosely coupled systems but future MPP systems will require these to scale to 1000s, tightly coupled.

Integration with high performance networks, including broad band
There is risk of bottlenecks between the high performance internal network of an MPP, and the external broad band highways. In some cases, such as video on demand, very wide bandwidth access to very large data stores are required, with little processing of the data. This leads to special OS developments, for example in the area of light weight communications software.

File systems for large multi-media data storage
Further developments are required on very large, very high performance file stores which support an atomic transaction model, are secure and reliable. For example, for performance, it is necessary to stripe data across multiple discs, yet maintain a single file store, and for this to scale up to 1000s of discs.

File systems on today's MPPs are crude compared to their equivalents on mainframe systems.

File systems are the subject of much international effort and European suppliers should work, and add value, to emerging industry standards.

Distributed lock and kernel resource management
Multiple users can access common data and kernel resources. Global distributed lock managers are necessary to synchronise updates. The lock manager can generate considerable inter-processor traffic in a large scale MPP system and much more work is required in this area. Again this is another topic where European suppliers are in a strong position to influence industry standards that are likely to emerge.

Processor interconnect networks.
These remain one of the most critical areas of the parallel machine. Different topologies are optimum for different applications: a complex query might best be executed through a tree network whilst high transaction processing rates might prefer a switched network. It is expected that the next generation of network will have much higher bandwidth and more intelligence.

Data storage
There has been much attention given to high speed networks for processing elements but little attention to the equally important area of networks for accessing terabytes of data on 1000s of discs. Fundamental work is required on the architecture of high performance I/O, need for storage devices and for high speed client/ server networking.

Migration of critical functionality from software to hardware
Considerable functionality is being put into the OS, but at the expense of overall performance. A careful analysis of performance critical resources, and which of these should be mapped into hardware, is required.

5 EMBEDDED SYSTEMS

5.1 Industry Needs

A new promising market segment for HPCN and parallel computing is emerging in the area of embedded computing, namely embedded HPC systems.

Embedded computers address systems, where a computer is an integral part of a machine, such as a car or a manufacturing system. Albeit not well known in the HPCN community the market for embedded systems is significant: it is estimated that there are three times as many embedded systems in the world today as the ca. 140 million personal computers.

Until recently the cost of HPC systems made their use in this field uneconomic. But this has changed with the dramatically improved cost/performance of HPC systems: Gigaflops are becoming a commodity. Today one raw Gigaflop is already available for 20.000 ECU. Consequently, embedded HPC products are for the first time feasible.

Three visions of the applications of embedded HPC systems might clarify the technical as well as economical potential of this technology:

The intelligent car: in several years from now you will want cars which are equipped with an active safety system which prevents accidents rather than preventing damage after an accident. Cars will have a security system to alert a (possibly sleeping) driver of hazardous situations and assist him in driving the car under varying street and weather conditions. These cars will need advanced computer vision and pattern recognition technology to perceive the road and the traffic. Researchers in this area have estimated the required computing power for such a security system as more than 20 Gigaflops. This computing power should fit into a small cabinet of less than 5 litre volume and should cost, considering the expected mass market, not more than 5000 ECUs.

The seamless integration of printed and hand-written documents into the digital world: contrary to early expectations the advent of the computer has not reduced the amount of paper documents, instead it has multiplied it. Paper documents are produced daily in the form of computer printouts, photocopies, faxes, letters, and forms. The existence of this non electronic communication is not going to cease. The problem is how to integrate these documents into the digital world of the computer and the upcoming world of the digital media. Truly intelligent optical character recognition, OCR, is the technology to solve this problem. The need for computing resources in this area is immense, and can only be fulfilled by MPP system. This can be demonstrated by the inadequacy of current word recognition technology, which is only able to read the writing of one person and that only after a long learning phase. Since there is a need for OCR in every office or maybe even on every cash desk, the market potential for OCR is very promising.

Fully automatic garbage sorting systems: in Germany it is obligatory that all the private packaging rubbish be recycled. Since the necessary separation of the rubbish cannot be done by the consumer, it has to be realised at central collecting points in every town. There it should of course be done fully automatic by garbage sorting machine. The separation of the rubbish is a demanding task for such a machine: how, for example, to distinguish a tin can from a yoghurt cup which may be totally deformed? Sophisticated pattern recognition software is needed which integrates different information about the objects, namely shape, colour, shades, and texture. Since the trend to recycle is getting stronger, there is a pressure to have means for an efficient separation technology.

Several European industrial projects have shown the beginnings of a realisation of these visions:

* Several organisations including a major European car manufacturer are involved in projects relating to self driving vehicles. Currently, experimental automobiles that use computer vision technology implemented on parallel machines to perceive the road are being tested. They have even been used to transport people without a driver on unblocked highways at speeds up to 160 kilometres per hour.

* Hand-written text recognition on credit card slips and letters in sorting machines is a technology which is increasingly and successfully used by European postal services and credit card companies. Parallel processing is used in all of these projects in conjunction with neural network technology. This enables the recognition of up to 25,000 hand-written slips per hour with a recognition rate of more than 90 per cent.

* A test plant to sort rubbish by means of neural networks and a parallel computer is being realised. In this concept, a camera detects and identifies individual objects moving across a conveyor belt and gives a command to which group it does belong. The recognition rate is already more than 70 per cent and an improvement in the recognition software and computing performance will make it ready for industrial exploitation.

Further examples of products which share these features are: satellite imaging systems, automotive toll systems, quality inspection systems, and speech recognition.

In all these diverse applications the core of each system is an HPC device which does pattern recognition and interpretation, referred to as a ***cognitive system*** in this text. The pattern to recognise might be acoustical, optical or formed by other signals.

Cognitive systems form the core of a wide ranging industrial market which has the potential, according to industrial analysts such as the Gartner Group, to exceed the two traditional HPCN areas of simulation and information management and processing in volume.

To draw a conclusion, cognitive machines are characterised by three key features:

* The integration of HPC in industrial and consumer products,

* The employment of pattern recognition technology, and, most important,

* A volume market.

Europe is excellently prepared to meet the challenge of cognitive machines and their applications:

There is a strong parallel processing know-how in Europe, and a large number of university graduates and engineers educated in this area. This know-how applies to parallel systems as well as to parallel applications. It has been triggered mainly by the wide distribution of Transputer technology in Europe. This expertise can contribute productively to the innovations required for cognitive machines.

Pattern recognition and interpretation technology has reached a mature state ready for commercial exploitation. Since these are new applications, they can be realised on parallel machines from the start. The considerable effort to parallelise existing sequential code can be spared.

Europe has traditionally excelled in the area of machine engineering, but has lost some of its lead recently. Together with the technology of cognitive systems this industry has the chance to regain its competitiveness.

5.2 Embedded Systems Technology Priorities

5.2.1 Applications

The first priority is to find applications which have a wide potential industrial impact.

The following list gives some examples of applications for cognitive machines. These examples are representative for an even larger class of applications and only give an overview of possible fields where cognitive systems can be used effectively. Most of the applications lend themselves to the use of advanced neural networks and for MPP systems, but also more traditional pattern recognition methods might be employed.

The different applications cited below are classified according to the environment where they are most likely to be used.

a.) Applications with impact mostly in industrial and commercial environments. These applications are mainly concerned with control automation.
- Quality inspections systems
- Hand-Written Form Reading (postal sorting systems and credit card slip reading)
- (Garbage) Sorting Systems
- Stock Exchange Prediction Systems
- Speech Recognition

b.) Applications with impact mostly in administrative environments.
- Security Systems
- Satellite Earth Observation Data Recognition
- Radar Signal Analysing
- Communication Network Management Systems
- Speech Recognition
- Automotive toll systems

c.) Applications with impact mostly in end-user environment.
- Medical Diagnostic Systems
- Self driving cars

All the above examples require expertise in at least three domains:
1. The application field to be represented by an industrial partner.
2. The specific algorithmic knowledge for pattern recognition. Partners for this domain can be found in research or small software companies which have begun to exploit the research results in pattern recognition.
3. The appropriate embedded MPP system.

5.2.2 Tools

Priority tools specific to embedded systems are as follows

Simulation tools for embedded applications
A system designer needs early feed-back on his proposed solution. He knows from the beginning all timing parameters and I/O requirements of his application: With this information he is able to design a system and describe the performance requirements, time behaviour and communication requirements. This information has to be fed into a simulator to get a behavioural simulation of his specification. With this rapid prototyping the system functionality can be reliably simulated and the time behaviour estimated.

Debugging and Monitoring Tools
A debugger to check the implementation against specification. A monitor is required to optimise the performance of the application on the embedded system.

Compiler for a concurrent object oriented language
An advanced concurrent object oriented language would allow for advanced design methods in developing embedded HPC application. An increase in productivity and the economical reusage of software could thereby be achieved.

Standard Libraries
Standard numerical libraries fitted to the requirements of pattern recognition software do ease their implementation and port, and ensuring highest performance on a platform.

5.2.3 Hardware Platforms

Cognitive machines share most of the requirements for the platforms for simulation and information management and processing. The following points regarding the hardware and the operating system for cognitive machines need special emphasis.

Reliability and Fault Tolerance
An embedded HPC system is mostly used as a production system for dedicated applications. Reliability and fault tolerance are of utmost importance in such an environment.

Advanced Interface Technology (Parallel I/O and Network Integration)
To integrate a HPC system into a cognitive system, specialised interfaces especially for parallel I/O must be developed to combine supercomputing power with different sensor/actor systems and databases.

6. GENERIC TECHNOLOGIES

In the above sections the special technology requirements for the three industry need streams have been highlighted. There are many HPCN technologies that are common to all three and these are included in this section.

6.1 System Level Designs

HPCN can have a major impact on an industry. To fully exploit the benefits, a total system view must be taken to build a complete solution for the need of the real user. This encompasses life cycle and competitive issues. The solutions are usually built on a combination of elements: hardware, software, services, etc.. The system value being much larger than the sum of its parts.

Examples of complex systems requiring system level design include command and control, the stock market and fraud detection. Such designs would lead to an analysis of HPCN architectures and a rationale for priorities for the underlying technologies for the particular domains.

Criteria for support to innovative designs of systems and systems technology in the HPCN area include the market and domain applicability, and the potential for new capability brought to IT industry, application industry and application users.

6.2 The Management of Heterogeneous Systems

Many computers of different architecture, with different features, and suitable for very different applications need to be interconnected via local and/or wide area networks, to form a large heterogeneous computing resource, and provide the user with the optimal system best suited for his application. However, today these heterogeneous systems are difficult to manage, to administrate and to exploit efficiently. Therefore there is a growing need for intelligent and user-friendly software which allows for a transparent and easy use of heterogeneous environments of computers, including workstation clusters and massively parallel systems. The following important features should be addressed within such environments:

- Load distribution and balancing
- Portability
- Performance
- Robustness
- Security
- Interoperability
- Integration of different techniques, e.g. databases real time+ artificial intelligence or parallel computing+serial-job computing+visualisation
- Tool/model for the management of large local and wide area heterogeneous networks of computers.

6.3 Neural Networks

Neural networks is an example of an emerging important new class of HPCN related application. Neural networks gain importance by solving problems in areas that could not be solved classically. Examples are quality control, robot control speech and visual recognition, forecasting and optimisation. The learning process of neural networks can be complex and therefore needs high performance systems and software. The potential of neural networks in many industrial areas is large and promising and their deployment will lead to large cost reductions in industrial processes and improvement of quality of work.

Therefore priority should be given to:
- Real life application of neural networks
- Improvement of neural network engines leading to high performance and cost effective solutions
- Combination of experimental methods and neural networks.

6.4 Tools

Complexity

One of the revolutionary aspects of the performance jump we will see with HPCN (factor of 1000) is in the possibility to combine in one application methods and technologies that had to operate disjointly until now. For example: an application combining a simulation model for prediction with a real-time process control function and multi-media interaction to the user, all in real-time.

This combining of different disciplines in S/W into one application very much complicates the engineering, and particularly the diversity and increase in size of such applications will induce a need for much more powerful engineering and S/W management tooling.

The ability to master this exponential growth of the complexity of applications will be a prime key to success in building real HPCN applications. A good understanding of this engineering challenge as well as having the appropriate tooling available in time should be stimulated.

Compiler Technology

In the coming decade we will see a lot of supplementary proprietary hardware in MPP systems in combination with standard processors. Examples are: special process to process communication channels across processor nodes, vector units interfaced to the processors, hardware support for virtual shared memory, hardware support for distributed locking, etc.

Each of these hardware devices will largely benefit from efficient support by compilers that are geared to implement and combine in a hybrid fashion such hardware devices with efficient code for the general purpose processors. When additionally support is required for high-speed parallel I/O, transparently compiled-in from regular programming languages that are being parallelised for MPP architectures, it is essential that elaborate compiler (generator) technology be put in place to flexibly accommodate such requirements.

Tools specific to the industry needs streams have been outlined above. Generic tools necessary to assist in the greater uptake of HPCN systems include:

- Engineering support tools for configuration management of large composite applications
- Programme debugging
- Modelling the behaviour of workloads
- Sizing and performance assessment
- Compilers for parallel languages and 4GL like environments
- Compilers for hybrid architectures
- Parallelising compilers for imperative languages such as C, C++ and F90
- 4GL like environments for these languages
- Debugging, also of parallel programs
- Performance improvements of sequentially compiled code to support enhanced processor features like super scalar scheduling, out-of-order execution, latency hiding and memory hierarchy management.
- Tools and techniques for (semi)automatic application parallelisation through detection and exploitation of inter-node as well as intra-node parallelism.
- Improvements in the use of compiled code through availability of easy to use tools and efficient libraries.

6.5 Platforms

Platforms include the operating system and the hardware.

Europe is in a strong position to lead on operating systems features and in particular on micro kernel developments, with several commercial products. For these to be accepted internationally, more work must be done on performance, integrity, resilience and integration with standards such as UNIX, POSIX and Windows NT. Increasing use of micro kernel technology is taking place world wide. A major advantage is its modularity, which makes possible, for example, module replication for resilience and selection of components for performance critical functions.

Priority operating system technologies are as follows:
- scalability
- Micro kernel technology
- Parallelisation
- Message passing
- Virtual storage
- Input/output
- Load balancing
- Integration with high performance networks, including broad band

Priority parallel hardware technologies are:
- Scalability
- Processor interconnect networks,
- Integration of internal with external networks,
- Data storage,
- Migration of critical functionality from software to hardware.

7. AWARENESS, TECHNOLOGY TRANSFER AND TRAINING

A major barrier for industry acceptance of HPCN is the lack of user awareness and of experienced personnel.

7.1 Awareness and Skills

Priority should be given to application integration projects, drawing together user applications and HPCN technologies. These should be seen as reference systems for wide dissemination of the advantages to industry of HPCN and for the output of 'aware' individuals. More precisely, a massive campaign is proposed for awareness and training through the developments of many new applications which are close to solutions for product and which require a non-obvious high performance computing component. The scope of a project in this campaign would be typically a one year product development involving few partners, the proposal for which being similar to a business plan. Marketing actions to promote the campaign and its results would increase the probability of success.

Support for users to evaluate the new HPCN technologies is needed. The PCI Parallel Computing Initiative, which is stimulating the dissemination of HPCN techniques in industry in Spain and Italy is starting this process. It is recommended that this should be extended to other countries in such a way that user industries, in particular SMEs, could trial their applications on parallel machines before they make any significant purchase investment.

Support for dissemination activities and their organisation is a necessary step to get more people aware of the technology advances. Organising the coordination of networks, conferences, students competitions/awards, publishing articles are examples of such activities.

The technology does not sell itself and European Union support for marketing the benefits of HPCN is needed. Grants should be awarded for marketing (to enter the market), for marketing studies and for software vendors willing to put their tools into the public domain.

7.2 Student Education and Training

The pervasive character of Information Technology is stimulating new approaches to education and training of the current and future workforce, in order to make them efficiently participate in a most demanding production environment. Focusing on the highest priority groups of higher education students and industrial managers and technical personnel, several actions are already on-going at Community level, in order to make them come closer to today's realities of computing. Yet, these actions for their larger part reflect conventional sequential computing and ignore emerging types of advanced architectures and software, relevant to HPCN.

More than 1 million graduates in science and technology is entering the work force each year in the EU area alone.

Currently, only an estimated 5% of students in science and technology are being trained to an adequate level of competence in advanced computing and networking techniques, which could be compared or assimilated to an extent to HPCN facilities and methods. Sophisticated client-server architecture systems are entering gradually the higher education environment, offering the students and academics therein, some powerful new features for increased performance. Yet, this rate is rather slow and there are major organisational and educational requirements which need to be set out in greater detail in order to motivate the greater use of HPCN within the population which has been pinpointed. It is proposed to tackle the problem at three distinct levels:

- increasing the number of students exposed to HPCN techniques and applications;
- transferring their skills from the academic environment towards industry and the wider economy;
- enhancing the receptiveness and take up within industry and the managerial environment.

Action is required to raise the number of students in science and technology in Europe acquiring HPCN skills from the current estimated 5% up to 40% in the five years up to 1998, and to reach 75% within a decade.

HPCN awareness must begin in secondary schools. Teachers have to be offered computer science re-training and refresher courses, together with the organisation of special activities (local exhibitions, presentations, events).

At higher education level, computer science courses have to become the focus for promulgating HPCN expertise and promoting specially designed modules with HPCN content. Courses, teaching materials and information dissemination can be organised on a European scale for teaching staff. Software tools and other results from ESPRIT projects would be made available for inclusion in course material.

Departments such as chemistry, biology and virology, materials science, physics, mechanics and electrical engineering can significantly benefit from applying HPCN skills and techniques to real situations. It is hard to find traditionally trained computer scientists who know enough about engineering and science to fully comprehend computational applications in these area. Community should particularly support these departments to make inter-disciplinary use of HPCN and teaching this necessary double competence.

A valuable support from the CE for science and engineering departments should be by jointly funding HPC systems and equipment (both hardware and/or software), by complementing and upgrading existing equipment, and by covering the expenses relating to the maintenance of such equipment and the training of staff needed to run the courses and the equipment. In practice proposals for such actions should demonstrate a clear focus upon industrial and technological application areas, and show evidence of willingness to share the experience gained in problem solving using HPCN in an organised way, on a case-by-case basis, with other university departments and the related industrial sector.

7.3 Managerial Education

The main thrust of HPCN training in the managerial and industrial context will consist of raising the awareness of engineers and managers as to the applicability of HPCN skills and techniques in their specific areas of responsibility. This is all the more pressing in that an awareness gap has been identified, and this can be traced back to the fact that there are currently too few engineers and managers who have the requisite experience and background in HPCN. These are specific short term activities, in order to accelerate the take-up of awareness. Demand will be created which may be satisfied, in the medium term, by the actions undertaken in education.

Feasibility studies and one-day information sessions of professional standard are considered key instruments to raise short term awareness of industrial managers. Professional and umbrella organisations representing user industries wishing to organise such courses and implement feasibility studies are viewed as key actors for the success of managerial education.

In the medium term, it is proposed to run courses of longer duration and workshops building on the information sessions and the conclusions of the feasibility studies.