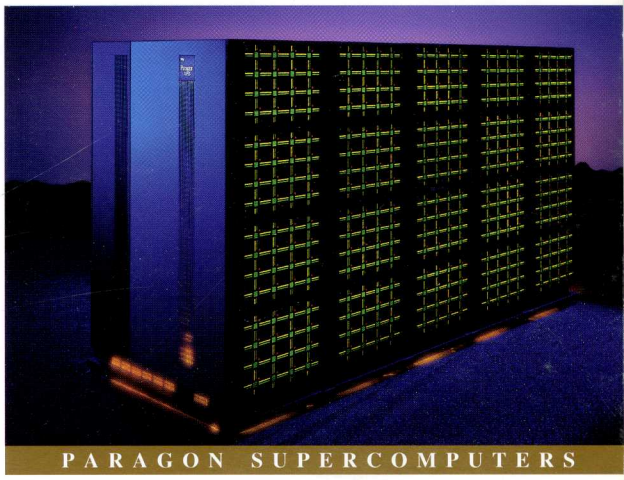
Intel Paragon™ Supercomputers

Paragon™ Supercomputers set a new standard for performance, programmability, and practicality in high-performance computing.



Designed to solve today's most challenging computational problems, Intel Paragon™ Supercomputers set new standards for sustained performance and usability in scalable, high-performance computing systems. Paragon XP/S supercomputers provide the overall capacity, balanced performance and ease-of-use needed to solve Grand Challenge computing problems, while offering the system services and administrative tools to manage the demanding environment of even the largest supercomputer centers. Paragon XP/E systems are entry-level supercomputers aimed at small institutions needing low-cost scalable systems for teaching and R&D, as well as consortium members needing local development platforms for larger, remote Paragon machines.

High Performance

- Configurations of 0.6 to 36 GFLOPS, 0.35 to 21 KMIPS
- Hardware node-to-node communications at up to 175 MB/sec, full duplex per channel
- Aggregate hardware interconnect bisection bandwidth to 5.6 GB/sec
- Scalable main memory to 77 GB†
- Scalable internal disk storage to 272 GB

Production Ready

- Industry-standard OSF*/1 operating system
- Multiple HiPPI, Ethernet, FDDI connections
- Simultaneous batch and interactive operation
- Resource control, accounting and scheduling tools
- Designed for high availability
- Air-cooled operation

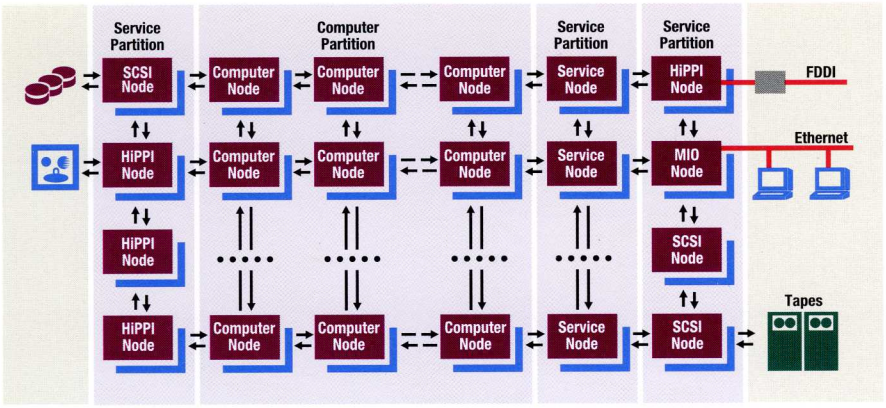
Enhanced Programmability

- Fortran-77, C, Ada
- High Performance Fortran††, C++††
- ParAide interactive, graphical development environment
- Performance-optimized node libraries and parallel equation solvers
- X Window System*, DGL*, OpenGL* and OSF/Motif* graphics
- Compatibility with Intel iPSC®/860 family

System Overview

The Paragon supercomputer is a distributed-memory multicomputer based on Intel's teraFLOPS architecture, which has been developed over the course of four generations of parallel supercomputers and is implemented using Intel's advanced microprocessor and

intel®



The Paragon™ Supercomputer. *Versatile nodes can support computation (compute nodes), input/output and network connectivity (I/O nodes), or interactive use (service nodes). The high-bandwidth interconnect provides node-to-node message passing at 175 MB/sec full duplex and allows the programmer to ignore the physical location of the nodes.*

semiconductor process technology. A modular implementation allows for continuous technology insertion as new generations of microprocessors and advances in messaging technology become available.

The system can accommodate more than a thousand heterogeneous nodes connected in a two-dimensional rectangular mesh. Nodes communicate by passing messages over a high-speed internal interconnect network. A general-purpose MIMD (Multiple Instruction, Multiple Data) architecture supports a choice of programming styles and paradigms, including Single Program Multiple Data (SPMD) and Single Instruction Multiple Data (SIMD).

System-Wide Balance for Maximum Delivered Performance

Architectural balance is the key to sustained multicomputer performance: the speed and memory capacity of individual nodes must be matched by the system's interconnect scheme, mass storage facilities, network connections, operating system software and graphics devices. To deliver maximum sustained performance, every aspect of the Paragon supercomputer is scalable and remains in balance as the system size increases.

Within each node, CPU speed is matched by on-chip cache performance and high-bandwidth memory units. Aggregate interconnect bandwidth scales with the number of nodes. The bandwidth and latency of internal communication interfaces are carefully matched to node memory bandwidth and execution speeds.

Internal storage capacity scales to 77 GB of dynamic RAM and more than a quarter terabyte of high-speed internal disk capacity. Aggregate I/O performance can be increased by adding multiple SCSI, HiPPI or other channels. I/O interfaces are custom engineered to the needs of specific industry-standard bus and channel interfaces. The operating system is also scalable, ensuring that operating system services will expand as the number of nodes or the application demand for I/O capacity rises.

PARAGON™ SUPERCOMPUTER ARCHITECTURE

System & Interconnect	
System organization	Nodes arranged in a 2D mesh
Communications paradigm	Message passing
Message latency (Hrdwre)	40 ns per hop across mesh
Node-to-node bandwidth	175 MB/sec full duplex
Total memory bandwidth	3.2-192 GB/sec
Bisection bandwidth	0.7-5.6 GB/sec
Processor	
Technology	Intel 75 MFLOPS, 64-bit i860™ XP RISC processor
Cache	16 KB instruction; 16 KB data
Physical memory	16-160† MB/node
Virtual address space per node	2 GB
Memory bus bandwidth	400 MB/sec
Cache bandwidth	1.2 GB/sec

Node Architecture

The Paragon supercomputer offers flexible, general-purpose (GP) nodes that can serve for computation, I/O or interactive use.

GP nodes are designed around Intel's i860™ XP RISC processor, the second generation of Intel's "supercomputer on a chip" i860 RISC processor series. The i860 XP operates at 50 MHz and integrates more than 2.5 million transistors. With an integrated floating point unit and a dual-instruction mode, it provides floating point performance of 75 MFLOPS double precision, 100 MFLOPS single precision.

For optimum node performance, the i860

processor's speed is matched by on-chip cache performance and high-bandwidth memory units. The chip includes 16 KB data and instruction caches, and the bandwidth between the processor and the cache memory peaks at 1.2 GB/sec.

General-purpose (GP) nodes have a single i860 XP application processor, as well as an additional i860 XP that is dedicated to improving latency and throughput of messaging operations. The performance of the message processor is not included in the system's performance ratings.

GP nodes provide an expansion port of adding an I/O interface such as the High Performance Parallel Interface (HiPPI) or Small Computer Systems Interface (SCSI). The nodes can be configured with 16, 32, 96† or 160† MB of high speed, error-correcting dynamic RAM, accessible at 400 MB/sec. The 64-bit memory unit provides interleaved, dual-bank memory accesses.

Interconnect Architecture

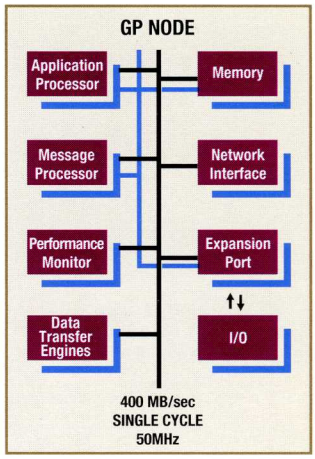
Along with node performance, inter-node communications is a determining factor for sustained multicomputer performance. The Paragon system's interconnect network provides high-bandwidth, low-latency communications and frees programmers from having to concern themselves with interconnect topology. All nodes appear to be connected to all other nodes, and communication performance is uniform. Bisection bandwidth, a measure of overall system capacity, scales from 0.7 GB for a Paragon Model XP/E 8N to 5.6 GB/sec for a Paragon XP/S Model 35.

The interconnect architecture is organized to provide optimum node-to-node communications while minimizing the involvement of the application processor(s) in messaging.

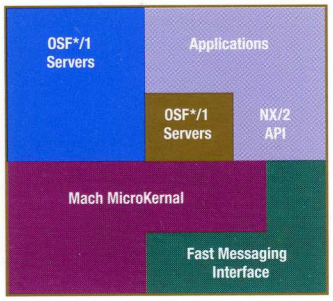
Message Processor. When an application decides to send a mesage, the node's i860 XP message processor handles message-protocol processing and frees the application processor to continue with numeric computing. Messaging software is executed from the message processor's internal cache, enabling overlapped communication and application processing to occur without incurring expensive context-switching delays. The message processor is also used to implement efficient global operations such as synchronization, broadcasting and global reduction calculations (e.g., global sum).

Message Routing. The actual transmission of messages is carried out by an independent routing system of custom-designed Mesh Router Controllers (MRCs), one for each node, arranged in a two-dimensional mesh. These fixed-function chips route messages between any two nodes in the system at hardware speeds of up to 175 MB/sec full duplex – more than enough bandwidth to directly service any peripheral device, high-speed point-to-point I/O channel, or local area network. Hardware latency, the time to set up the transfer of the first byte of a message, is so low – 40 nsec per MRC traversed – that physical location of nodes becomes unimportant for performance. Like the i860 XP processor itself, the MRCs are fabricated using Intel's triple-metal, submicron semiconductor process.

Node-Interconnect Interface. Completing the interconnect implementation is the Network Interface Controller (NIC), a custom VLSI chip that provides a full-bandwidth, pipelined interface between each node's MRC and node memory. Supported by two block transfer engines on the node, each NIC permits simultaneous inbound and outbound communication at hardware rates of 175 MB/sec and provides end-to-end error checking for each message transfer.



General-Purpose Node. *Each GP node dedicates one i860 XP processor to user applications and one to message processing. The GP node's expansion port allows the addition of an I/O or networking interface.*



Distributed OSF*/1 Operating System. *The microkernel-based OSF/1 operating system minimizes software latency and maximizes scalability. It also presents a single system image to the user, regardless of the number of nodes in the machine.*

Hardware-Aided Performance Monitoring Support

GP nodes include a data-capture function that can capture run-time data for performance analysis and applications tuning. The data-capture function also provides a 1-microsecond global clock, which aids in debugging and performance analysis by allowing accurate time correlation of events occurring in different parts of the system.

Distributed OSF*/1 Operating System

The Paragon supercomputer's operating system brings OSF/1, the Open Software Foundation's industry-standard version of the UNIX* operating system, to the performance-driven environment of scalable, distributed-memory computing. The operating system provides the portability benefits of industry-standard UNIX, plus a number of enhancements that increase application performance and ease of use in a scalable environment.

Standards-Compliant. The use of proven OSF/1 technology ensures an industrial-strength operating system, improves applications portability, and fosters integration into heterogeneous computing environments. The operating system is designed to comply with the POSIX, AT&T System V.3, Berkeley 4.3bsd and the X/Open's XPG3 standards. In addition, NX/2 libraries provide compatibility with the Intel iPSC family.

Transparent Distribution and Single System Image. The Paragon supercomputer's OSF/1 implementation eliminates entirely the need for a host, front-end computer, or dedicated control processor that creates bottlenecks and limits the scalability of many parallel supercomputers. Instead, users can log onto the supercomputer directly, and the machine appears as a single system with a single process ID space and file system, despite the fact that the operating system is running in parallel over hundreds or thousands of nodes. Every file, every process, and every network service is available to every authorized application.

To achieve this convenience and scalability, the operating system is fully and transparently distributed across the system's nodes. A Mach 3.0-based microkernel resides on each node and implements core operating system functions. Higher level services such as the file system server run on service nodes, with access transparently provided from each microkernel. In this way, the operating system can scale to support an increasing number of nodes or users or rising I/O requirements.

THE PARAGON™ SUPERCOMPUTER'S OSF*/1 OPERATING SYSTEM

- Proven OSF/1 technology
- Designed to comply with with POSIX1003.1, System V.3, 4.3bsd UNIX* operating system standards
- Mach 3.0 microkernel/server architecture with multitasking, multi-threaded support
- Transparently distributed, single system image
- Multi-user mesh management tools, including accounting and resource management
- Interactive and batch processing
- Virtual memory
- Shared virtual memory††
- Parallel file access for high I/O bandwidth
- iPSC/860 and Concurrent File System compatibility

Parallel Programming, Minimal Software Latency. The operating system presents a variety of opportunities for exploiting the machine's parallel architecture. Enhanced NX-style message passing is supported with the messaging processor, and fast messaging routines reduce software latency by enabling applications to access the interconnect network with minimum software overhead.

The Paragon system also offers enhanced Mach multi-threaded processes, as well as parallel fork and exec primitives to use in creating parallel UNIX processes.

Resource Management. For flexibility in managing the Paragon supercomputer's resources, the system offers simultaneous batch and interactive scheduling, using separate partitions for each. The system administrator controls the definition of partitions and can alter them to match the site's job mix. With the batch partition, applications are gang scheduled, ensuring that all processes associated with an application are scheduled in parallel.

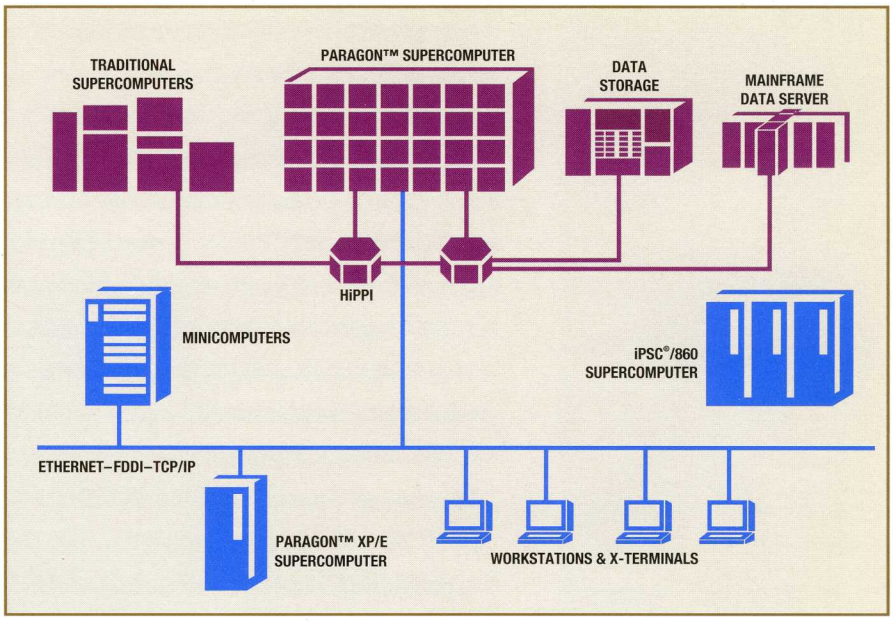
The Multi-User Accounting and Control System (MACS) provides facilities for scheduling jobs, tracking system resource usage and reporting on resource utilization and user activity. The Network Queuing System (NQS) works with MACS to provide support for batch operations.

Memory Management. In addition to the high performance of the message-passing architecture, the system provides programming flexibility through virtual memory on each node. Virtual memory simplifies software development and porting by enabling even extremely large programs to run on a single node before being distributed across multiple nodes.

Shared virtual memory†† permits multiple nodes to share a region of memory directly and makes it easier to port programs from shared memory systems.

File Systems. The OSF/1 operating system implements the standard UNIX file system, so existing applications can access files without the need for code modifications. In addition, disk striping provides a parallel file access method that improves I/O performance and allows users to create and access files as large as the amount of physical disk space available in the system.

The system is also source code-compatible with the iPSC/860 Concurrent File System.



Supercomputer connectivity. *With industry-standard connectivity options, X Window System support, and UniTree client interfaces the Paragon XP/S supercomputer provides easy interoperability in the networked environments.*

Mass Storage, I/O, Networking.

Scalable storage facilities and industry-standard connectivity options make the Paragon system ready for integration into production computing sites and research centers.

MASS STORAGE AND CONNECTIVITY

- Up to 272 GB internal disk storage, hardware RAID controllers
- Scalable bandwidth to 5.6 GB/sec
- UniTree* client support
- Scalable HiPPI and SCSI connections
- Ethernet, FDDI and HiPPI networking
- NFS and TCP/IP protocols
- HiPPI IPI-3††
- HiPPI to VME††

Mass Storage. The Paragon supercomputer provides reliable, cost-effective mass storage via arrays of commodity 3.5-in disks protected by hardware RAID (Redundant Arrays of Inexpensive Disks) technology. The system supports more than a quarter terabyte of scalable internal disk storage. In addition, a 4 mm DAT cassette tape drive allows convenient data exchange.

A client interface supports backups and restores to external network-based file servers supporting the widely-used UniTree* file

PARAGON™ SUPERCOMPUTER DEVELOPMENT ENVIRONMENT

Languages

Fortran-77
C
C++
High Performance Fortran translator††
Validated Ada* (from Verdix)

Tools

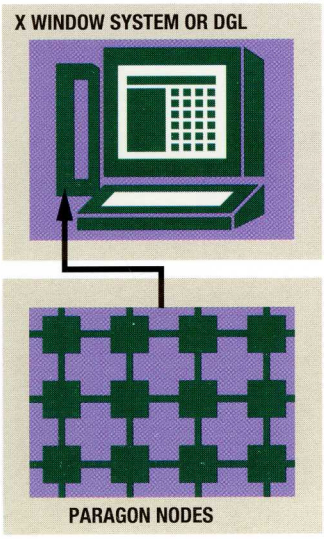
FORGE-90* parallelization tools (from APR)
ParAide interactive development environment – parallel debugger and performance analysis tools with OSF/Motif graphical user interface and System Performance Visualizer

Sequential Libraries

BLAS, FFT, NAG
Signal processing library (SEGlib)

Parallel Solvers & Libraries

ProSolver™ family
DES direct dense matrix solver
SES skyline sparse matrix solver
IES iterative sparse matrix solver
2D and 3D FFTs
*Available through a third party supplier



Network-based Visualization. *Using X and DGL client libraries on the supercomputer, applications can utilize the computational power of the Intel i860 XP processors to directly generate graphics for industry-standard display servers.*

management software. Also part of the system are standard UNIX file maintenance utilities such as backup, restore, and tar.

I/O Interfaces. GP nodes can be configured with high-performance, industry-standard interfaces, to create I/O nodes. Multiple I/O nodes can be added to meet the needs of I/O- or network-intensive applications, larger systems, or increasing numbers of users.

Available interfaces include:

- A multi-function I/O daughter card to support Ethernet and 8-bit fast SCSI and 16-bit fast SCSI†† interfaces.
- A HiPPI (High Performance Parallel Interface), for peak transfer rates of 100 MB/sec (full-duplex) to stand-alone disk arrays, HiPPI switches, HiPPI graphics frame buffers, and other computer systems.
- A HiPPI to VME†† connection provides a useful building block for special I/O systems integration.

Network Connectivity. For flexibility in client/server compute environments, the system's I/O nodes support standards-conforming interfaces for HiPPI, Ethernet and FDDI connections. Networking software includes standard utilities and protocols such as FTP and TCP/IP, and sockets, telnet, and NFS*. Standard HiPPI-IPI3†† interfaces will provide access to high speed external peripherals and file management systems.

Developing High Performance Applications

High performance compilers for industry-standard languages, plus a growing set of parallel development tools assist in creating new applications and porting existing codes to the MPP environment. Most tools can be used directly on the supercomputer or as cross-development tools on Silicon Graphics and Sun Microsystems workstations.

Languages and Compilers. The Paragon supercomputer's standards-compliant compilers for Fortran-77, C, C++†† and Validated Ada generate highly optimized code, foster applications portability and support a choice of programming paradigms. The system's ANSI-standard Fortran-77 and C languages are inter-callable, freeing programmers to choose the source language that best suits the tasks at hand. The Fortran-77 compiler supports standard Mil Spec enhancements such as VAR and DO/ENDDO.

High Performance Fortran is an emerging language that will allow developers to use the familiar Fortran programming style to create and modify applications to run on a variety of scalable systems. The language supports both MIMD- and SIMD-style programming, and includes a variety of features that aid in code parallelization and optimization. For High Performance Fortran, a source-to-source translator, uses standard HPF data decomposition directives to convert Fortran 77 and Fortran 90 programs into SPMD parallel programs that can be compiled by the Paragon Fortran-77 compiler.

C++ is an object oriented version of the C language that supports classes, subclasses, methods, objects, messages and inheritance (automatically sharing data types among classes and sub classes).

Debugging and Performance Analysis. The ParAide integrated development environment provides a graphical environment for debugging and performance analysis of large-scale parallel applications. It combines the functionality of the Interactive Parallel Debugger (IPD), the Unix* *prof* and *gprof* utilities, and an enhanced version of ParaGraph with an OSF/Motif graphical interface. ParAide is designed around a scalable Tools Application Monitor (TAM) that provides efficient data collection for both debugging and analysis.

ParAide includes special facilities for

debugging message passing programs and for minimizing the complexity of displays, tables and windows, thus boosting programmer productivity with applications spanning hundreds of nodes. Data collection can be turned on and off selectively at run-time, and information such as event traces and CPU usage can be displayed in numeric or visual formats.

A System Performance Visualization (SPV) tool provides a means for users, developers and system administrators to gain an overview of system usage or examine a particular node's or partition's activities in detail.

Performance-Tuned Application Tools

Intel's ProSolver family of parallel equation solvers are performance-optimized libraries that simplify application development by addressing the user's need to solve large systems of equations. The libraries handle both matrix assembly and equation solution, and can be applied to dense and sparse systems of equations.

Several widely used single-node math libraries are available, including the BLAS and NAG libraries, the signal processing library (SEGlib), and routines for Fast Fourier Transform (FFT) algorithms.

Visualizing the Grand Challenges

The Paragon system has multiple options for network-transparent, client/server visualization. The X Window System Version 11 enables client applications to run on the supercomputer and display 2D output on X-compliant devices. In addition, the system supports Silicon Graphics' de facto standard Distributed Graphics Library* (DGL) and OpenGL.

Both these capabilities are supported over Ethernet, FDDI, and HiPPI networks. HiPPI frame buffers can also be connected to the Paragon system to provide high-bandwidth connectivity for interactive visualization and rendering.

Convenient, Cost-Effective Operation

The Paragon supercomputer needs no more than the standard cooling and environmental protection of a typical computer-room. The system is air cooled and uses conventional AC power supplies, which eliminates the need for special cooling equipment, simplifies system installation, and reduces life-cycle costs.

Clean packaging for the system does away with long fragile cabling used in conventional vector supercomputers, and still found in many highly parallel machines. Each backplane section is connected to its horizontal and vertical neighbors by flexible, ultra-reliable printed circuits. No single link is more than seven inches long.

To maximize system availability, critical aspects of the system hardware, such as power supplies and disk subcontrollers, have redundant components. Nodes use error-correcting memory, and the interconnect mesh has its own error detection facilities. An internal diagnostic network monitors important system components and operation.

Software is subjected to rigorous regression and stress testing, and the commercially-proven OSF/1 operating system ensures a reliable software foundation.

Ensuring Success

Intel backs its Paragon supercomputers with sophisticated consultation and support services designed to ensure a smooth transition from traditional vector computers into the higher performance Paragon supercomputer environment.

Available services include application benchmarking, parallel programming classes, and long-term, applications consulting. System maintenance is supported by trained technicians in more than 100 offices worldwide. Users also benefit from membership in the large and growing Intel supercomputing community, with more than 400 installations around the world.

ProSolver™ EQUATION SOLVERS

ProSolver-DES

- Dense equation solver
- Real and complex double precision
- Pivoting

ProSolver-SES

- Direct solution of sparse systems of equations
- In core and out of core
- Real double precision
- Finite element interface

ProSolver-IES

- Interactive solutions of sparse systems of equations
- Single and real double precision
- Generalized minimal residual method (GMRES(k))
- Conjugate gradient

PARAGON™ XP/S AND XP/E SYSTEM SPECIFICATIONS

System Configuration

Model Number	XP/E-8N	XP/E-16N	XP/E-24N	XP/E-28N	SP/S-2	XP/S-5	XP/S-10	XP/S-15	XP/S-25	XP/S-35
Peak GFLOPS (64-bit)	.6	1.2	1.8	2	2	5	11	16	25	36
Number of Compute Nodes	8	16	24	28	32	66	140	208	336	480
Number of I/O Nodes	1	1	1	1	1	3	5	7	11	15
Number of Service Nodes	1	1	1	1	1	2	3	4	5	5
Distributed Memory (GB)	0.128	0.256	0.384	0.448	0.5	1.1	2.4	3.5	5.5	7.8
Maximum Memory (GB)*	1.28	2.56	3.84	4.48	5.1	10.8	22.9	33.9	54.6	77.6

I/O System (Standard)

Standard Internal Disk Capacity (GB)	4	4	4	4	4	8	16	24	40	56
Maximum Internal Disk Capacity (GB)	8	8	8	8	20	56	92	175	200	272
HiPPI Channels	Multiple HiPPI interfaces available as an option									
Standard Peripherals	One 4mm tape (2 GByte capacity) and one Ethernet Interface									
Total I/O Bandwidth Available (GB/sec)	1.4	1.4	1.4	1.4	1.4	3.8	4.9	5.6	5.6	5.6

Operating Environment

Power Consumption (KWatts)	1.2	2.2	3.2	3.8	4	8	16	23	37	51
Thermal dissipation (KBTU/hr)	4.2	7.7	11.2	13.3	14	28	54	78	124	174
AC Power & Cooling	50-60 Hz, Single Phase, 208 Volt, Air Cooled Plus 50-60 Hz single phase 110 V desktop Diagnostic Station				50-60 Hz, Three-Phase, 208 Volt, Air Cooled					

Physical Dimensions

Number of Cabinets	1 cabinet, 16 slots	1 cabinet, 32 slots Plus desktop Diagnostic Station	1	2	3	4	6	8
Width (ft)	1.8		1.8	3.6	5.3	7.1	10.7	14.2
Height (ft)	5.1		5.1					
Depth (ft)	2.1		3.5					
Weight (1000 lbs)	0.5		1	2	3	4	6	8
Required Floorspace (sq ft)	10		12	18	25	33	43	56

*Maximum memory available Q2 1994.

Intel Corporation
Supercomputer Systems Division
14924 N.W. Greenbrier Parkway
Beaverton, OR 97006
(503) 629-7600

Intel Corporation (UK) Ltd.
Supercomputer Systems Division
Pipers Way
Swindon, SN3 1RJ
United Kingdom
(+44)-793-696700

Intel Japan K.K.
Supercomputer Systems Division
5-6 Tokodai, Tsukuba City
Ibaraki-Ken 300-26
Japan
(+81)-29847-8511

Intel GmbH
Supercomputer Systems Division
Dornacher Strasse 1
8016 Feldkirchen bei Muenchen
Germany
(+49)-89-90-992-415

Intel Corporation
Supercomputer Systems Division
1 Rue Edison-BP303
78054 St. Quentin-en-Yvelines Cedex
France
(+33) 1-30-57-7000

† Available Q2 1994
†† Available Q4 1994