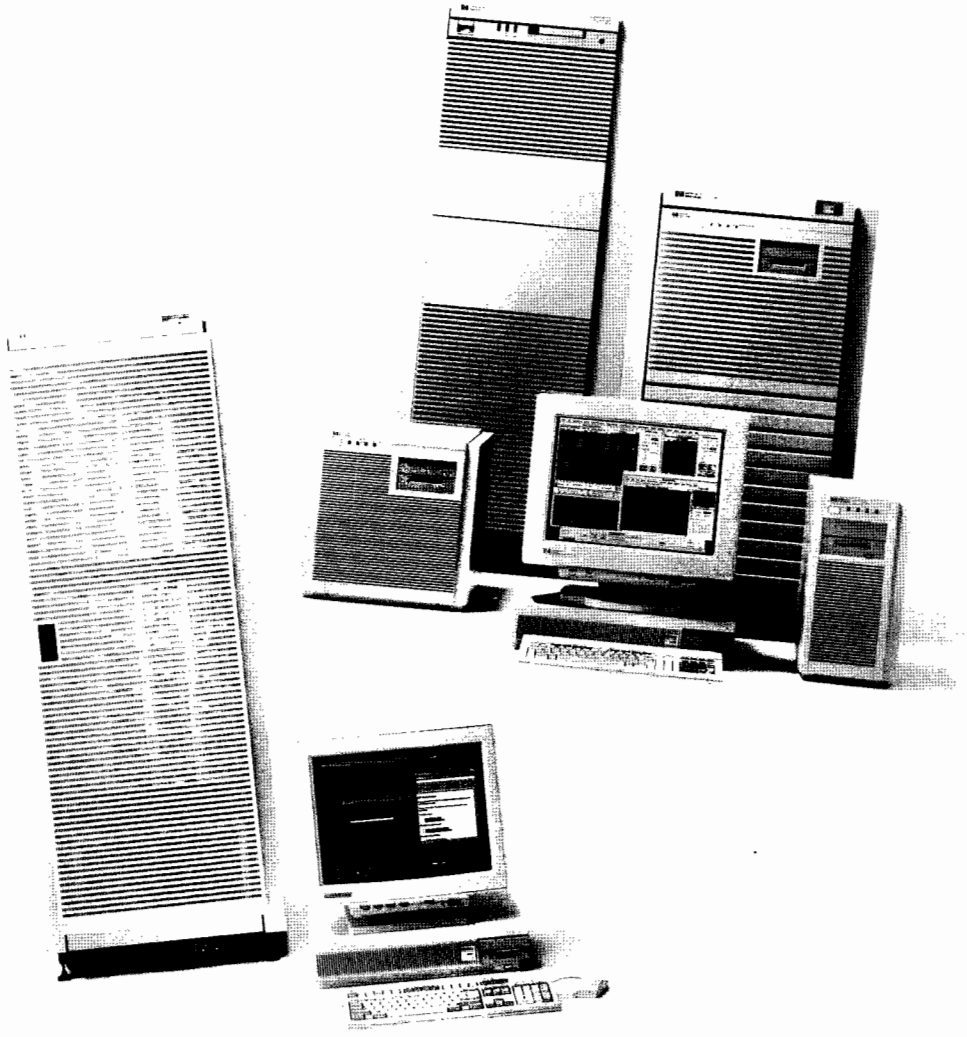
HEWLETT® PACKARD

Technical Server Sales Guide



HP and HP Channel Partner
Internal Use Only

Table of Contents

HP Computer Museum
www.hpmuseum.net

For research and education purposes only.

Technical Server Sales Guide

Purpose:
This guide provides the details to HP's High-End Technical Servers. It provides the background, unique selling opportunities, target markets, and competitive data you need to sell successfully.

Chapter 1 — The High Performance Compute Server Strategy
Chapter 1 articulates the overall performance message of the technical server compute paradigms which should be communicated to the customer.

Chapter 2 — Product Overview and Positioning
This chapter describes the strengths and weakness of the Series 800, the Series 735CL, and Convex's Exemplar Systems. By understanding the differences you will be able to lead with the most appropriate server to meet customer needs.

Chapter 3 — Series 800 Servers
Chapter 3 positions the Series 800 Servers into the Technical Server Market. There is an in-depth list of the solutions as well as useful information on the competition.

Chapter 4 — Clusters
This chapter describes Hewlett-Packard's Series 700 clusters. There is detailed information on the potential user, useful competitive information and information on the target markets.

Chapter 5 — The Convex Partnership
Chapter 5 provides an explanation of HP's selling agreement with Convex for the Exemplar Systems (SPP-1000XA and SPP-1000CD).

Appendix A – Technical Comparison of the Series 800 and Series 700 Clusters
A technical comparison to help you understand the differences.

Appendix B — Third Party Software for Series 700 Clusters

Appendix C — A 'pocket guide' to multiprocessing

1.0 The High Performance Compute Server Strategy

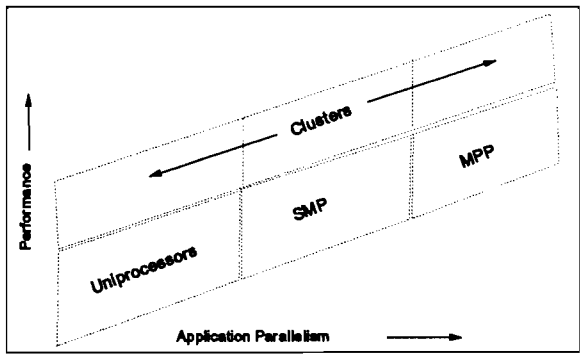
HP's customers span a broad range of markets and industries. To effectively meet the performance requirements of these customers, HP must exploit system level parallelism in several ways. Because each customer situation or customer application exhibits different opportunities, three multiprocessor computing techniques are currently used within HP's extended technical server family. These techniques are Symmetric Multiprocessors (SMP), Massively Parallel Processing (MPP) Systems and Compute Clusters. They are all based on HP's strong PA-RISC architecture and HP-UX application binary interface.

Fundamental to the success of these multiprocessor techniques is strong uniprocessor performance. Today's PA-RISC 7100 family, and future generations, will insure industry leading uniprocessor- and multiprocessor-based system performance throughout the decade.

The strength of uniprocessors is leveraged through processor-level parallelism within HP's SMP-based systems. SMP has helped HP grow to a number one position in the Unix server market. It will continue to play a key role in delivering industry leading Unix server performance across the technical and commercial markets.

Massively Parallel Processing leverages a higher degree of inherent parallelism within certain applications to dramatically boost the performance of these applications. HP's PRO partner, Convex Computer Corp., recently introduced the Exemplar family of scalable parallel processing systems configurable with up to 128 PA-RISC processors. These systems are ideally suited for those applications that exhibit a high degree of parallelism.

HP further builds on the strengths of its uniprocessor and SMP-based systems through system-level parallelism. Through the use of clustering techniques, these compute servers provide exceptional throughput capability rivaling the performance of supercomputers at a fraction of the price.



The combination of HP's strong uniprocessor strategy and its three multiprocessor computing techniques allows HP to deliver highest performing compute servers...optimal solutions to exploit parallelism within our customers' compute environments to provide leadership performance growth throughout the decade... architectures that provide a high degree of scalability and investment protection within a flexible and productive high-performance computing environment.

2.0 Technical Server Positioning

HP's extended technical server family is made up of the Series 800 Servers, the Series 700 Clusters, and Convex's Exemplar Scalable Parallel Processing (SPP) systems. The purpose of this section is to briefly articulate the relative strengths and weaknesses of these three product lines and to provide an overview and product comparison to aid the sales representative in understanding the product with which to lead, or recommend, in a given customer situation. For a more detailed discussion of each of these product families refer to the individual product sections within this sales guide.

This section is organized as follows:

a) A general overview of the product and its strengths, and target customers and markets.
b) A table summarizing the relative strengths, weaknesses, and possible customer characteristics and target markets.
c) A decision tree to illustrate the general positioning of the three product families.

The Workstation Cluster – Model 735CL

Clustering is a technology of networking multiple machines together to act as a single *computational* resource. (This is different from the commercial application purpose of tying machines together for high availability, like the VAXCluster). The Model 735CL computational cluster is a rack of 4 or 8 "headless" workstations networked together with Ethernet or FDDI.

The primary customer benefits are as follows:

1. Provides the computational throughput capacity of a supercomputer at a fraction of the price.
2. Is modular and scalable up to hundreds of workstations.
3. Provides an inexpensive alternative to those customers that want to develop in, or experiment with a parallel computational environment.

Clusters are typically sold to customers in need of a dedicated high-performance compute server to execute discrete, compute-intensive jobs.

Target market segments are as follows:

Market Segment	Type of Apps
1. Gov't Labs & Universities	Homegrown
2. Mechanical CAE	Third Party
3. Petroleum (Exploration and Analysis)	Third Party and Homegrown

Clusters do not share memory or disks and are, therefore, sometimes called a "shared nothing" system. Since each node within the cluster is essentially a self-contained workstation, clusters are usually used as "throughput" compute servers. That is, servers to which multiple, independent, compute-intensive jobs can be queued or "batched."

Most customers use servers in this manner—as throughput machines! However, many times sales decisions are made on the time it takes to complete one job—time-to-completion! Make sure that the customer is really judging system by how it really will be used in their environment.

Third party software is usually used to control batch queuing or load balancing. *See Appendix B for a comprehensive list of cluster related third party tools.*

Series 800 Servers

The Series 800 servers are uniprocessor and multiprocessing machines available from HP. The uniprocessor models are the E Class Models 25, 35, and 45 and F, G, H & I Class Models 50 and 60. The multiprocessing or SMP (Symmetric Multiprocessing) models are G, H and I Class Model 70 and the T500. The value of the Series 800 uniprocessor machines is that they provide:

- added expandability (memory, disk capacity, I/O),
- high availability features (mirrored disks, battery backup, automatic switchover), and
- growth (investment protecting board upgrades and box swaps).

In addition to the features available in the uniprocessor models, the SMP models also provide shared memory symmetrical multiprocessing, allowing multiple jobs to run on the same machine without degrading performance.

The Series 800 Servers have been traditionally marketed to commercial markets. While they offer a very strong platform for commercial applications, they are equally as strong for many scientific and engineering applications. HP is developing features for the Series 800 which are of importance to technical customers. In addition, HP is working with key ISVs to qualify their applications to the Series 800.

Series 800 servers are typically bought by mainstream "production" shops, seeking to augment or replace existing shared resource. They want painless scalability, cost-effective upgrades and minimal operating costs. In technical markets, the Series 800s are often used as NFS file servers, software compile servers and database engines for technical applications. They are also used as compute servers for:

1. Remote (non-desktop) computation
2. Throughput (running multiple jobs at the same time) – especially for cache-local or memory blocked applications
3. X-Terminal Servers

The Series 800 are strong throughput engines, but are not recommended for parallelized applications until HP-UX 10.0 and parallel Fortran are available. While the Series 800 can run a parallel job, a computational cluster or the SPP from Convex is likely to perform better on most parallel jobs. The SPP is specifically tuned to handle parallelization. The cluster, with the aid of parallelizing software (PVM, Linda), can run coarse-grained parallel applications very efficiently. The next generation of the Series 800 will also be better optimized for parallel jobs.

The Convex Exemplar

Convex has recently introduced their Exemplar Scalable Parallel Processing system. These systems are based on PA-RISC and an HP-UX application binary interface. These products are binary compatible with the Series 700 and Series 800 families although application re-qualification is recommended. This Exemplar family is made up of the SPP-1000XA (Extended Architecture) scalable from 8 to 128 CPUs and the SPP-1000CD (Compact Design) scalable from 2 to 16 CPUs. The key elements of Convex's SPP-1000 product family are as follows:

1. High speed, multi-tiered, shared memory subsystem that minimizes traffic jams between CPUs and allows the system to scale to a large number of processors.
2. High performance software environment which includes parallel compilers, and parallel software development tools. (Applications need to be developed and/or compiled to take advantage of MPP systems.)

The Convex Exemplar machines are excellent compute servers for parallel applications since they have kernel based threads and parallelizing compilers and are tuned for fast interprocessor communications. These systems can either be viewed as fast "time-to-completion" machines for single, large parallel jobs or high-performance "throughput" machines for lots of simultaneous parallel jobs.

For applications to be "parallelized," they either have to be compiled with Convex's parallelizing compilers, or modified using a message passing protocol such as PVM to run across several processors simultaneously. Relatively few third party applications have been parallelized to date, although ISVs are increasingly turning to this technique to improve the performance of their applications. Many customers with compute intensive home-grown code are also "parallelizing" their applications.

Application segments and industry segments that lend themselves well to MPP are as follows:

Market Segment	Type of Apps
1. Gov't Labs & Universities	Homegrown
2. Mechanical CAE	Third Party
3. Petroleum (Exploration and Analysis)	Third Party & Homegrown
4. Computational Chemistry	Third Party & Homegrown

The Exemplar Family is also designed to fit into the high-end SMP computation segment that the Series 800 does not cover — parallel technical applications. The SPP1000 systems do not provide the same system call performance of a native HP-UX product (like the Series 800 and Series 700). Therefore, it is not as good as the Series 800 for NFS file server, database and, of course, for commercial applications.

Product Comparison Table
The product comparison table on the next page summarizes the key points made above and lists relative strengths, weaknesses, and possible customer characteristics and target markets. It is designed to help clarify which product will best suit your customer's needs.

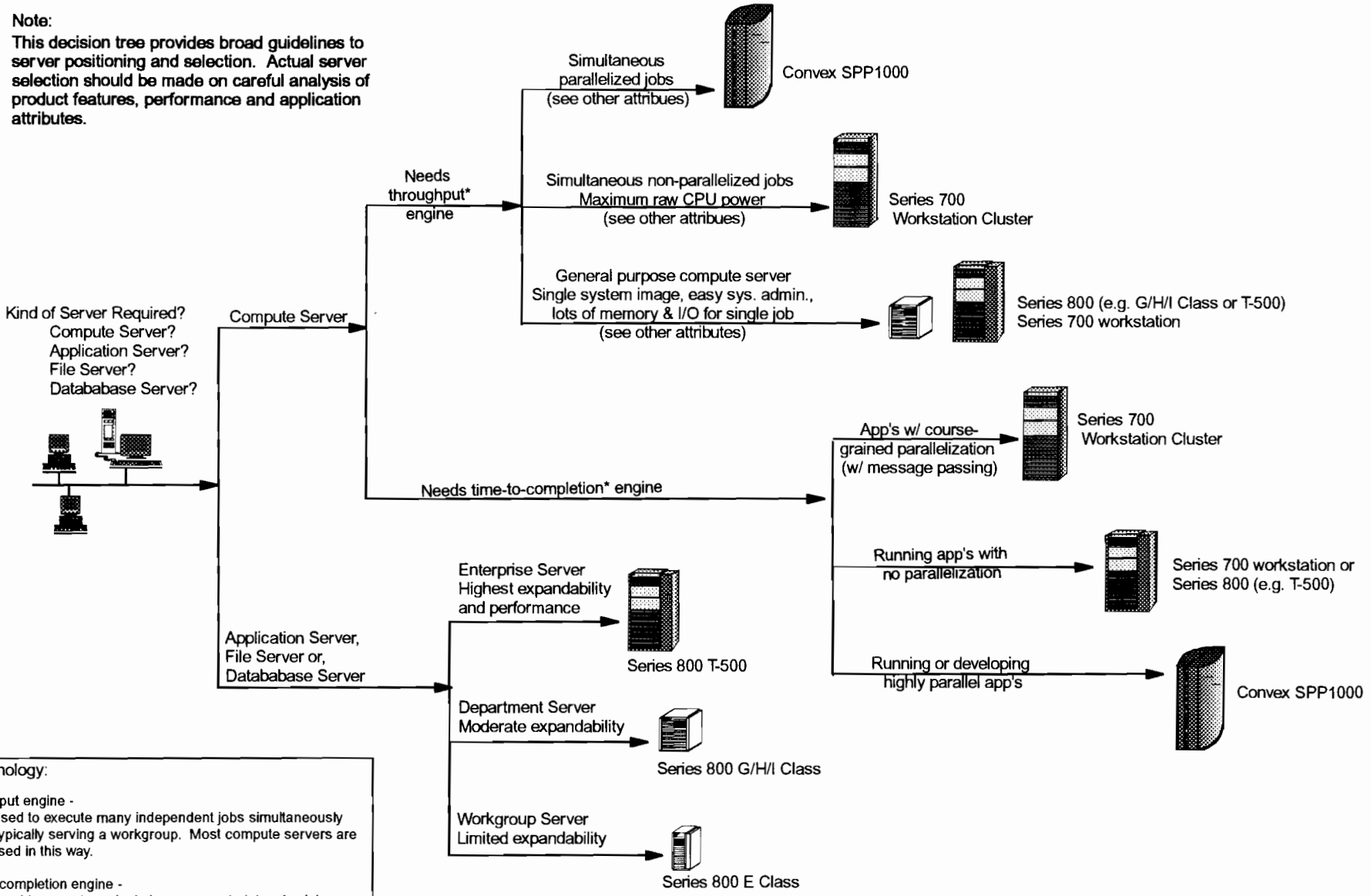
Technical Server Decision Tree
The technical server decision tree on page 12 is designed to graphically illustrate the key differentiating customer requirements that need to be considered when selecting a technical server. It also acts as a summary to the information stated above.

None of these tree paths have 'cut & dry' rules. There could be lots of cross-over depending on customer situations; the decision tree is only a guide.

	Positives	Issues	Customer Characteristics	Applications/ market
Series 700 Clusters	• Low initial & incremental cost/CPU • No contention for memory, each processor has its own memory bus, doesn't compete w/other CPUs for bus bandwidth or memory space • Coarse grained parallelism (via message passing) • Good throughput engine for compute intensive jobs that require high memory bandwidth • Inherent availability, cluster runs if one node dies • Cost-effective incremental scaling to 100s of CPUs	• No single system view for users, admin's, programmers • May require multiple copies of software licenses • Not recommended for fine-grained parallelism • Not recommended for interactive applications • Limited memory & I/O exansion per node compared to similar Series 800 Server	• Needs throughput engine (complete lots of compute-intensive jobs in a given time) • Want shared compute resource for compute hungry workgroup • Want as much raw CPU power as possible • Want the best price-performance • Want MPP, but are constrained by small budget • Have coarse-grained apps that are parallel (distributed memory parallel) and/or have expertise to parallelize own homegrown code • What to implement innovative high-performance solutions • Engineering group or lab • Leading edge customers	• National labs • Research institutions, universities • Mechanical computer-aided engineering (e.g. CFD) • Petroleum industry (e.g. reservoir simulation) • Computational chemistry • Software development (Compile/builds of code) • R&D groups within above industry segments • Homegrown applications • Possibly homogeneous job stream
Series 800 Servers	• Shared peripherals (disk drives, tape drives, etc.) • Global memory can be shared between processes • Larger cache may make large jobs run faster. • Single multithreaded OS (one copy of OS for multiple CPUs) • Single point of administration • Higher performing communication between processes (shared memory vs. internal cluster network) • System services (acc'ting, time-sharing, load balancing) are system supplied vs. user supplied • Tuned for database and file serving • Large, shared I/O capacity	• Different expansion bus from S700 family to handle large I/O requirements (HP-PB vs. EISA) • Check to ensure ISV supports technical apps on S800 until availability of HP-UX 10.0 • As with all SMP systems, shared memory bus may limit scaling with simultaneous memory intensive jobs • Synchronization & inter-CPU communication may slow performance for large parallel apps	• Mainstream "production" shops • Throughput engine for heterogenous mix of jobs • Wants easy system admin	• NFS file server • Software compile servers • Database engines • Application server/X station server • General purpose compute server - EDA - MCAE - Scientific - CASE - AEC/GIS • Best utilization with heterogeneous job stream • MIS departments
Convex SPP1000 Family	• Highest performance for parallel applications • Based on PA-RISC and HP-UX ABI • Scalable to 128 processors • High performance and high capacity memory subsystem • High bandwidth I/O subsystem • Multithreaded microkernel • Parallelizing compilers • Parallel application development performance tools • System software (subcomplex mgr) allows partitioning of user processes	• Higher cost per CPU (therefore lowest price-performance) • Uses microkernel and may require HP application re-qualification • Different expansion bus than S700 family (S-bus vs. EISA) • A Convex product • Some HP-UX layered software products may not be available at First Customer ship • Single CPU job performance not faster than comparable Series 700 or Series 800 • Large files & system file size, threads and checkpoin — - restart in a later release	• Wants high performance server for parallel applications • Is developing parallel apps • Requires high memory capacity and/or performance • Requires high I/O capacity and/or performance • Not price sensitive • Wants advanced MPP architecture	• National labs • research institutions, universities • Mechanical computer-aided engineering (e.g. CFD) • Petroleum industry (e.g. reservoir simulation) • Computational chemistry • R&D groups in above industry segments • Homegrown applications

Technical Server Decision Tree

Note:
This decision tree provides broad guidelines to server positioning and selection. Actual server selection should be made on careful analysis of product features, performance and application attributes.



* Terminology:

Throughput engine -
Used to execute many independent jobs simultaneously Typically serving a workgroup. Most compute servers are used in this way.

Time-to-completion engine -
Used to execute a single large compute intensive job the fastest.

3.0 The HP 9000 Series 800 Servers

One-Minute Sell

This document is designed to help you position and sell the Series 800 servers into engineering, scientific, and other technical markets. First, recognize that the worldwide technical UNIX server market in 1994 is $3 to $4 billion (Source: IDC), and this represents a sizable opportunity for you and HP. Today, many of HP's customers require the ability to purchase both servers and workstations from the same vendor. The technical server program will arm you with the information to position the broad product line against the competition. Most of HP's competitors already provide a product line which extends from the desktop to the "data center" for technical computing. Plus, you already know the customers who require these solutions; they're the same ones who are buying workstations from you! The technical server program gives you a strong offering to help you beat the competition.

Many customers and HP sales reps view the Series 800 Servers as the leading platform for commercial UNIX applications (OLTP, database, MRP). Yet the Series 800 is also an excellent platform for technical computing! Features developed for commercial applications are also extremely important in the tech space (FW SCSI, fast I/O throughput, high avail.). And remember, it contains the same core technology as the workstation product line.

3.1 Overview of Series 800 Line

The Series 800 Servers offer three different design centers to meet customer requirements:

- The workgroup level;
- The department level; and,
- The division level.

The following is a breakdown of the different classes of Series 800 product line:

The Workgroup Server – E-Class

The E-Class is the newest member of the Series 800 product line, and has the following key features:

- Uses PA-7100LC (same as in 712 workstation).
- Runs same HP-UX and has same applications availability as higher end products
- Offers the full high availability features as in other Series 800s
- Available at 44 (E25), 60 (E35) and 80 (E45) MHz clock cycles.
- Comes with 2 NIO slots standard (4 slots is an option)
- Offers strong performance in low cost package:

	E25	E35	E45
SPECint92	45	65.6	80
SPECfp92	66.6	98.5	122
Estimated TPS	80	125	155
SPECnfs	450	650	800

The Department Server – G/H/I Class

The G/H/I are the mainstay of the Series 800 product line. The Models 50, 60 and 70 use PA-7100 at 96MHz, the same chip as in the 735 and 755 workstations. The package offers expandability and growth without box swaps. The Model 70 provides two-way SMP in a moderately priced server.

Configurations:

- G = 4 NIO Slots
- H= 8 NIO Slots
- I = 12 NIO Slots
- 50 = 1 CPU w/256x256 KB cache
- 60= 1 CPU w/1MBx1MB cache
- 70 = 2 CPU w/1MBx1MB cache/CPU

	G/H/I 50	G/H/I 60	G/H/I 70
SPECint92	100	108.8	N/A
SPECfp92	158.5	195.3	N/A
SPECrate_int92	2,382	2,554	4,930
SPECrat_ fp92	3,790	4,626	8,310
Estimated TPS	190	280	411
SPECnfs	1,014	1,313	N/A

The Division Server – The T500

The T500 is the high-end of the Series 800 product line, and has the following features:

- It offers 1 to 12 way SMP
- Uses PA-7100 at 90MHz
- Large cache/CPU (1MBX1MB/CPU)

Expandability:

- 12 to 114 NIO Slots
- 256 MB to 2 GB memory
- Disk storage to 1.9 TB

	T500
SPECint92	98.3
SPECfp92	170.2
SPECrate_int92	2,310 (1 way) to 23,717 (12 way)
SPECrate_fp92	4,019 (1 way) to 38,710 (12 way)
Estimated TPS	300 to 2,000+
SPECnfs	Estimated up to 4,000 with HP-UX 10.0

For more detailed configuration information, see the Series 800 Configuration Guide, P/N 5091-9206E. For up-to-date performance benchmarks, see the UNIX Systems -Performance Quick Reference Card, P/N 5091-7097E (also on PowerTools and the Hotlines).

3.2 Target Market Segments for Series 800

The Series 800 will be targeted towards three segments or applications in the technical market:

- NFS file servers,
- Compute, and
- Database.

HP is focusing on markets segments where the Series 800 product line can be a strong player and where HP already has a leadership position with the Series 700 workstations. IDC recently estimated the size of the UNIX technical server market in 1994 to be between $3 and $4 billion.

Market Segment	% of UNIX technical server market
NFS File Server	50%
Compute Server	30%
Primary Market Segments:	
EDA	
Technical CASE	
MCAE	
Secondary Market Segments:	
Scientific	
AEC	
GIS	
Database	20%

NFS File Server Market Segment

The Network File System (NFS) is a file sharing application which allows users to access remote file systems transparently. Usually, the customer will have a centralized file server running NFS. Clients (workstations or X stations) mount the NFS server to read and write to files. NFS is part of Sun's Open Network Computing (ONC) technology. NIS (Network Information System) is also part of the ONC technology and is used with NFS to provide a single file for login names, host files, etc.). NFS and NIS come bundled with HP-UX. NFS File Servers can be configured for workgroups, departments, and even entire divisions.

The market size for NFS File Servers has been estimated to be approximately half of the market for UNIX based technical servers. NFS File Servers are used in almost every market segment, from automotive to petroleum.

NFS is an I/O intensive application. Therefore, it is important to select servers with not only the processing power to handle NFS File requests, but also the I/O capacity to quickly read and write to files on disk. Typically, S700 workstations quickly become I/O bound when being used as NFS File Servers. The Series 800 servers can solve that problem. The G/H/I Class servers provide from 4 to 12 I/O slots (but remember, each Fast-Wide SCSI card and FDDI card

take 2 slots!). Additionally, the T500 provides from 14 to 112 slots and a capacity of up to 256 MB/sec. The E Class servers have limited expansion, and therefore are best used for fairly small workgroups.

Today the NFS performance on HP platforms is measured on uniprocessor servers only. This covers a large segment of the NFS marketplace, and one PA-RISC processor provides more NFS performance than four SuperSPARC processors! While HP has the best uniprocessor performance in the industry, it is adding MP scaling to NFS (in HP-UX 10.0), which will allow NFS to take advantage of more than one processor.

Competition and Selling Strategies for NFS

1. Sun

Sun developed NFS as part of its Open Network Computing (ONC) software. Sun has become more aggressive recently in marketing its server line as NFS File Servers. In addition, Sun recently introduced an NFS Cluster, which is designed to provide a high level of NFS performance at a much lower price than the SPARCserver 1000 or SPARCcenter 2000. Here is the breakdown of product offerings:

SS Model 512: 200 NFS_OPS: This model offers less than half the NFS performance of an E25!

SS Model 514: 300 NFS_OPS: Even with four processors, Sun has not reached the NFS performance of an E25!

SS1000: 2,106 NFS-OPS (8 CPU): Sun only published performance using all 8 CPUs, since it takes all 8 SPARC chips to get good NFS performance.

SC2000: 2,575 NFS_OPS (8 CPU): Again, Sun only publishes NFS performance numbers on the high end of the systems. You have to spend at least $545,900 ($US List) on CPU, memory and disk to achieve this level of performance.

SPARCcluster 1: 480 to 3069 NFS_OPS: Sun advertises this as the low cost solution. It is neither. To get the maximum "potential" performance, you must spend over $500,000 US List. Plus, there is no load balancing software in the system to balance the NFS operations between processors. If all NFS requests go to one cluster node, all you get is the performance on one SS10 Model 40 machine (about 100 NFS_OPS). The entry level configuration provides the same level of NFS performance as the E-25, but at more than ten times the cost! How's that for low cost!

Clearly there are many opportunities for HP to sell NFS file servers successfully against Sun. And when we have MP Scaling for NFS (due in HP-UX 10.0), HP servers will outperform Sun at every level of NFS performance.

2. Auspex

Auspex is a leader in specialized NFS File Servers. Auspex Systems Inc. is a young company (founded in 1987), has approximately 300 employees and approximately $75 million in annual sales.

Auspex calls its NFS File Servers *"functional multiprocessors."* Rather than have multiple processors sharing equally in the NFS functions, the processors in the Auspex box are specialized. Some are Ethernet (or FDDI) processors and handle only network traffic. Some are I/O processors and handle disk traffic. One chip (a SPARC processor) handles NFS protocol. The remaining processors are Motorola 6800 chips. To add an additional LAN connection, you have to also add an additional Ethernet processor. Auspex therefore adds performance by adding a LAN connection (and processor). Series 800 File Servers have the traditional uniprocessor and symmetric multiprocessor architecture. Adding a LAN connection will not improve performance on an NFS File Server from HP. Adding a PA-RISC processor to the SMP machines (Model 70 and T500) will improve performance once NFS MP Scaling is added to the HP-UX. The Motorola chips in the Auspex are connected to a VME bus, which runs at a faster speed than the standard VME bus, making it very proprietary (and expensive). You can only purchase additional processors, disks, and network connections from Auspex.

Auspex can achieve some fairly high performance with its latest round of machines. Still they are very expensive and not very expandable.

Auspex NFS performance benchmarks:

NS/6000-2 (2 Ethernets) 546 NFS_OPS
NS/6000-4 (4 Ethernets) 1,108 NFS_OPS
NS/6000-8 (8 Ethernets) 2,037 NFS_OPS

Auspex got a jump in the market with the introduction of their functional multiprocessing systems. Yet the Auspex systems have not changed much in the past five years, and their architecture is dated. HP systems have kept up with the times. In fact some would say that they are ahead of the times. Plus, HP offers the ability to perform more than NFS on a file server. This is a great time to go head to head against Auspex.

Weaknesses of the Auspex Systems are:

- Very proprietary (not open at all)
- As better machines are available, cannot re-deploy Auspex to perform other functions
- Auspex NFS File Servers cannot perform any other functions than NFS File Serving – they can't do batch processing, compiling, email services, print services, boot services and other ancillary services placed on a general purpose server.

For more information, see the Auspex File Server Evaluation on the Networking and GSY Hotlines.

Compute Server Market Segment

Customers often buy Series 700 workstations as compute servers. While they offer leading compute performance, workstations can be limited in scalability and expansion (memory, I/O), especially when compared to other servers in the market. The Series 800 servers extend the workstation product line to compete against products such as Sun's SPARCcenter 2000 and SGI's Challenge Series.

The SMP servers , the models 70 and T500, are excellent solutions for running multiple jobs at the same time (commonly called "throughput"). A customer may want, for example, to run multiple simulations to test a circuit design. The Series 800 SMP products will allow the user to run multiple jobs at the same time. Your customer may wonder when he or she should purchase an SMP over a cluster of workstations. Clearly there are costs and benefits to each design. The SMP architecture is limited in the number of jobs it can run at the same time by the size of the memory bus. When a customer is running a job mix that is either:

1) largely cache local, or
2) has a mix of some large and many small to medium jobs, an SMP solution is generally the better solution.

If the customer is running mostly medium to large memory intensive jobs, then a *workstation cluster* might be the better answer. Because each cluster CPU has its own memory, it will be less likely to saturate the memory bus. Yet some customers may not want the complex system management required with a cluster. For them, the Series 800 SMP is the best solution.

The Series 800 is also an ideal X station server. For flexibility in configuration, the Series 800 supports multi-vendor clients. To optimize performance and manage multiple concurrent designs, the system supports large memory and disk capacity. It also provides UPS and backup features that protect critical design data. These features lower the cost per seat and provide easier administration.

Some customers are trying to off-load their applications from a supercomputer and have technical requirements out of the realm of the mainstream of the technical server marketplace. These customers should be directed to Convex to learn about their PA-RISC based SPP products.

The Electronic Design Automation (EDA) Market
EDA focuses on tools for:

- ASIC Design (Application Specific Integrated Circurits)/System Design Entry
- PCB (Printed Circuit Board) Design & Layout
- IC (Integrated Circuit) Design & Layout
- Circuit Simulation/Analysis/Routing

SPECint92 is the important benchmark to measure EDA applications performance. SPECrate_int92 is the benchmark for SMP systems. It measures the throughput capacity of SMP systems (multiple independent jobs running on multiple processors). EDA applications are integer intensive and do not require much floating point performance. EDA applications take advantage of larger and faster caches. The HP 9000 product line has a performance advantage over other workstation and server vendors with its large, fast caches. The Series 800 also offers enhanced uniprocessor performance with fast caches up to 1MB x 1MB per CPU. (The Series 700 goes as far as 256KB x 256KB). Also, the Series 800 offers these large caches in MP solutions, with both the 70 Models and the T500.

HP is a leader in the EDA market segment. Currently the second largest vendor of hardware in this market, HP is looking to increase its share through aggressive sales and marketing programs. The Series 800 is a key platform in this strategy. Customers in this market are broken down into the following segments:

1. Semiconductor Manufacturers/Designers (e.g. Intel)
2. Electronic Product Manufacturers/Designers (e.g. Sony)

HP has approximately a 25% market share. HP is aggressively trying to increase our market share in the EDA segment. And the Series 800 is a key ingredient for our success here.

Application Vendors in EDA

The following application vendors support or are in the process of supporting the Series 800:

NEW on Series 800

Mentor Graphics (MGC)

Mentor Graphics applications are now available on HP 9000 Series 800 Servers. Mentor Graphics solutions include the Falcon Framework design environment, and tools for ASIC, IC and PCB design, including MCMs. Utilizing HP's high performance Series 800 servers, Mentor Graphics offers engineering organizations the following:

1. significant cost-per-seat savings,
2. higher productivity, and
3. greater system performance through a client-server architecture.

With the addition of the Series 800, Mentor Graphic's EDA offering provides several significant benefits. Currently, engineering organizations are searching for ways to reduce time to market, lower cost, improve design quality and manage increasing design complexity. The Series 800 offers solutions to these challenges by maximizing computing resources in a client-server environment.

Increases in engineering productivity and reductions in time to market are benefits associated with the increased compute resources offered on the Series 800. These features enable engineers to offload compute intensive jobs, such as time consuming simulations, onto servers while freeing their workstations for other tasks. Cost per seat savings can be achieved when a Series 800 is used in an X station server configuration. UPS and high availability protect design data.

Mentor Graphics products available today include:

- BOLD
- Design Architect
- Design Viewpoint Editor
- ICgraph
- BOARD Station graphical application (excluding AutoTherm, Physical Cable and Board Station 500)
- QuickSim II
- SimView

For more information contact Bob Gannon at (503) 598-8246 or Bob Morrison at (503) 598-8295, or contact the GSY hotline, subject "MGC800.'

NEW on Series 800

Cadence Design Systems, Inc.

Cadence currently is qualifying some of their most compute-intensive, server-based tools on the Series 800. These include Gate Ensemble™, and Cell 3 Ensemble™ for IC place and route, Verilog-XL™ and Verilog Turbo™ for logic simulation, and Dracula™ for IC verification. All of these tools can take full advantage of both the performance and memory expansion capabilities of the Series 800.

Cadence design data can also be placed on a Series 800 being used as an NFS file server. Cadence customers can take advantage of the Series 800 high availability and back-up functionality to protect their critical design data. This helps improve time-to-market for Cadence users.

NFS file support is available immediately, Cadence application support is projected for later this year. Contact Dennis Brown at Cadence (408-428-4064) for updated information.

NEW on Series 800

Cooper and Chyan Technologies

The Cooper and Chyan Technologies (CCT) SPECCTRA™ Autorouter will be available on the Series 800 by 2QCY94. Many leading EDA suppliers have incorporated CCT's software into their technology.

One of the most difficult tasks in printed circuit board design is connecting components on the board. As printed circuit boards become denser, and have more design constraints, this becomes a more difficult problem. SPECCTRA™ uses a new shape-based technology which enables customers to avoid manual intervention.

The Series 800 provides customers the benefit of offloading multiple SPECCTRA™ routing jobs to a powerful server. By providing more platform choices, the customer can buy the most cost-effective configuration to match their client-server design environment.

For more information contact Sandy Sund at CCT 408-366-6966, or in Europe contact Ben Franklin at CAD Connection +49-89-260-6166.

Other EDA applications on the Series 800 (Source: HP-UX Solutions Catalog):

Zuken Quad Designs
GenRad Technology Modeling Associates
Computervision Theda
Pacific NUMERIX

EDA Competition at a Glance

	SUN	DEC	IBM
Positives	–#1 share –Low price leader/ high discounts	–Strong price/performance	–Gaining market share –Strong server position in EDA
Negatives	–Dated technology –Weak Sparc performance	–Losing market share (only 4%) –OS strategy confusing –OSF/1 not available on high-end servers	–No SMP solution
Actions	–Use Series 800 to block sales of SS1000, SC2000 –Series 800 paired with X stations provide low-cost solution	–Good time to attack DEC accounts	–Beat IBM's high-performance servers with Series 800 SMP solution

The Mechanical Design Automation (ME) Market Segment

In ME, servers are forecasted to be a growing segment of hardware sales (reaching $500 million by 1996) and represent a tremendous opportunity for technical servers. By application, the largest segment of the ME market is CAD, followed by CAE and then CAM. CAE, which includes finite element analysis, thermal analysis and computational fluid dynamics (CFD), is compute-intensive and requires the additional expansion and performance of the Series 800. CAE is a growing segment in ME and the Series 800 will help you capture that growth.

A new trend in the market is Product Data Management (PDM) systems, which track designs, parts, inventories, etc. PDM systems require hardware which can handle analysis of ME systems and also complex document management. Most of the major ME vendors are developing PDM systems. The Series 800 will be an excellent platform for these products since they are database oriented.

Application Vendors for ME

NEW

Computervision

Computervision announces two new product data management offerings for HP 9000 Series 800 Servers:

- Design Manager and
- Engineering Data Management (EDM)

Design Manager provides small workgroups with easy to use data management functionality that previously was only affordable for large sites. This is a new product from Computervision.

EDM plays a key role in facilitating simultaneous or concurrent engineering by providing a secure, enterprise repository for parts and design information.

Customers that are streamlining processes and utilizing new technologies like Computervision's EDM can drastically reduce lead times and lower cost while improving quality. EDM forms the backbone of concurrent engineering network where all the product teams work together to produce the product. The product provides:

- Interfaces for drafters, managers and administrators to access, track, sort and retrieve design data for revision, rework or reference.
- Links to production facilities with suppliers that enable companies to send 3D models to increase the accuracy of delivered parts.
- Controls, manages and distributes electronic design information from applications like CAE/CAD/CAM, word processing and spreadsheet programs.
- Enables design and manufacturing teams to share ideas and information.

HP 9000 Series 800 Servers give customers the option of choosing a mainframe class, central repository strategy or a multiple client-server site strategy for product data management. The 800 servers provide capabilities that support enterprise-wide global implementations of EDM engineering data management solutions. These capabilities include a wide array of networking, high availability features such as UPS and switchover, and strong OLTP performance that maximize access to important information and safeguard the company's strategic product assets.

EDM and Design Manager work with Computervision's CADDS5 Mechanical Design Package and have interfaces to manage designs generated from other popular ME-CAD applications. Contact: John Spindler, Global Partner Manager (617) 221-5003 Parametric Technology Corp. (PTC) to support Pro/PDM on thte Series 800. For more information contact Bruce Clay, Global Partner Manager at (617) 221-5240.

Other ME products available on the Series 800 (Source HP-UX Solutions Catalog):

- Sherpa DMS, PIM Series
- ISICAD
- Visionary Design Systems (Solid Design. Workmanager, ME10, ME30, DMS ME10D)
- Gerber (Sabre-5000 only)
- Swanson (ANSYS)
- MSC (Nastran)

MCAE Competition at a Glance

	SUN	IBM	SGI
Positives	–Large installed base –Price leader	–Strong in MCAE –Heavy investments in applications –Power2 architecture highest performer in industry	–Viewed as #1 up and coming in ME market –Challenge Series and powerful servers
Negatives	–Perceived focus away from ME market –Low performance at high-end	–No SMP solutions	–R4400 lags PA-RISC –No high-availability solutions –Lacks price/performance
Actions	–Attack Sun accounts with Series 800	–Position Series 800 SMP against IBM servers	–Position T500 against Challenge Servers

Scientific Market Segment

The scientific market is a catch-all category that contains a number of market segments including Oil & Gas, Analysis (using SAS, for example) and others.

Applications Availability for Scientific Market

NEW

CXSOFT

CXSOFT announces performance-boosting math software and clustering tools for HP's 9000 Series 800 Servers. CXSOFT, a newly formed business unit of Convex Computer Corporation, announced four software solutions for the HP 9000 Series 800 servers.

1. Convex MLIB: a math library which significantly boosts performance for numerical routines. The remaining three software products allow Series 800 Servers to run in a clustered configuration.
2. Convex LSF is an advanced and user-friendly cluster load balancing technology.
3. Convex PVM is a parallel programming tool for clusters of systems.
4. Convex NQS+ is an efficient batch queuing system for distributing jobs across multiple systems.

CXSOFT is an independent business unit of Convex Computer Corporation. Leveraging Convex's industry leadership and significant investment in its compilers, visualization and mathematical library software, CXSOFT markets and distributes very high-performance system software solutions for scientists and engineers using PA-RISC workstations and servers.

CXSOFT software products benefit customers who need tools to serve a large metacomputing environment where large number of users compete for system resources. CXSOFT's distributing computing tools provide new opportunities for HP 9000 Series 800 Servers offering the ability to cluster existing Series 800 systems. These tools also enable Series 800's, 700's and Convex Meta-Series to IOP and act as a "Virtual Machine" while taking advantage of the capabilities each system provides.

Product availability: ConvexMLIB summer '94, LSF immediately, ConvexNQS+ summer '94, and Convex PVM Spring '94.

For more information: Contact CXSOFT (214) 497-3300 for business office; (800) CXSOFT1 for orders.

Other Scientific Applications

- SAS – All SAS Software
- Petroleum
- InTera – ECLIPSE
- Computer Power Group – AP-Today, Operating Control Systems
- RMS Technologies – Digital Terrain Modeling (DTM), METSTAR Egret Technologies – EQuIS/GWM, EQuIS/LAB, EQuIS/RDB, EQuIS/SWM Lynx Geosystems Inc. – Lynx 3D Component Modeling Systems Oceanroutes Canada LTD – Oceanroutes Fisher Controls – SCADA Gas, SCADA Water

Competition in the Scientific Market at a Glance

	SUN	IBM	SGI
Positives	–Low-end price leader	–Aggressive in Oil and Gas –Tight relationship with national labs and universities –Power2 architecture	–Fastest growing vendor –Challenge servers –Make hay of yet to be released TFP
Negatives	–Poor performance in uniprocessors and multiprocessors	–Power2 architecture not available in desktop or SMP –No SMP offering	–Best chip available today is R4400; poor floating point –Lack price/performance advantage
Actions	–Beat Sun on low-end with workstations and on high-end with servers	–Position SMPs against high-end servers	–Position T500s against Challenge Series

A More Detailed Analysis of Competitive Product Lines for All Compute Segments

SUN

Sun has not kept up with the other RISC vendors in improving their chip technology. It is significant to note that the uniprocessor performance of HP's lowest end model (E-Class Model 35) is at least as good as Sun's best (SS Model 10-51). And the 8 way SPARCserver 1000 and SPARCcenter 2000 quickly run out of gas. You can accomplish the same level integer performance with four 90MHz PA-RISC chips as what it takes 8 SuperSPARC chips in the SS1000 or SC 2000! And the difference in floating point performance is even greater! While the eight CPU SS1000 is at the end of the line, you can still go up to 8 or even 12 CPU with the T500! Sun has had trouble getting more than 8 way support in Solaris. They recently announced that they are shipping servers with support up to 20 processors on the SPARCcenter 2000. It is unlikely that there will be many customers for the SS2000 with 20 CPUs. It will be difficult for most technical applications to take advantage of 20 processors (remember, they all share the same bus!). If customers are looking for scaling beyond what is available in the T500, you should bring in Convex to talk about their PA-RISC based, HP-UX compatible SPP.

IBM

Matching Solutions	
Sun SS Model 10-51	715 or E Class 25 or 35
Sun SS Model 10-512	735/755 or G/H/I 50 or 60
SPARCserver 1000	G/H/I 70 or T500
SPARCcenter 2000	T500

IBM is a strong player in the technical UNIX server market with the Power 2 architecture. That architecture, which contains 2 floating point and 2 integer units in a multi-chip design has the highest performance numbers on the market today. IBM has no SMP machines. Instead, it pitches the SP-1 for MP applications. The SP-1 is a glorified cluster. It is a rack containing multiple Model 370 workstations without their skins. There is no centralized bus, all communications between CPUs must converse the network (and TCP/IP) as opposed to pointers in memory (in shared memory). Please see section on clusters for a more detailed description of clusters and the SP-1.

IBM will be introducing an SMP machine this summer (1994), but it will use the PowerPC 601 architecture, not the Power2. The 601 is much lower performing than PA-RISC at 90 or 96 MHz.

When competing against the Power2 servers from IBM remember the following:

1. If your customer is interested in a shared resource and/or running multiple jobs at the same time, the Model 70 or T500 can run multiple jobs without degrading performance (there is a limit to how many jobs can run at the same time however, depending on the application). Since the Power2 is a uniprocessor, it is going to be slowed down by running multiple jobs. The SPECrate benchmarks demonstrate the difference.
2. If your customer only cares about running one large job faster (time to solution), then running the job in parallel is the best bet. If the application is coarse grained parallel (like many CFD and graphics applications), a cluster might be the best bet. Otherwise, you might want to recommend that the customer consider a low-end SPP from Convex.

SGI

The Challenge Series SMP servers offer 2-12 way processing (Challenge L) and 2-36 way processing (Challenge XL), a large system bus (1.25 GB/s) and large memory capacity (6GB on Ch L; 16 GB on Ch XL). The Challenge Series also supports kernel based threads, parallelizing compilers, large file sizes and file systems sizes (check this too). Still, the Challenge Series is not without its faults.

- The MIPS 4400 chip is a much lower performing chip than the 96 and 90MHz PA-RISC 7100 chip.
- SGI's prices for memory and disk are extremely high (even higher than HP's!).
- TFP, the new chip that is supposed to bring SGI chip technology up to the level of PA-RISC, hasn't shipped yet, despite the fact it was announced over a year ago. SGI does a good job of selling futures, but not delivering product.
- The MIPS architecture as tiny on-chip caches and large and slow off-chip caches. With hierarchical caches (on and off-chip) there are at least two clock cycle penalties to go to off-chip cache. Additionally, the off-chip cache runs at a slower clock cycle than the on-chip cache and the CPU, creating a greater delay when going to off-chip cache (off-chip cache for a 150 MHz R4400 chip runs at 75 MHz). In the MIPS architecture, this is especially painful because MIPS uses a super-pipelined architecture. It takes several stages (and clock cycles) to get to cache. Cache miss penalties are exceptionally time-consuming. PA-RISC uses large off-chip caches that have one clock cycle delay and run at the same speed as the CPU. Designs in the PA-RISC architecture handle cache misses more efficiently. Therefore we have big fat *fast* caches, which can greatly improve applications performance.

Matching HP Servers Against SGI:
Challenge M – 735, 755 or G/H/I 50/60
Challenge L – G/H/I 70 or T500
Challenge XL – T500

DEC

Digital's Alpha chips boost strong performance, but don't always deliver that performance for real applications. None of DEC's SMP machines support OSF/1 yet. Alpha, like the MIPS architecture for SGI, uses a hierarchical cache structure (8KB on chip, 512 to 4 MB on board) and is super-pipelined. Alpha therefore suffers some of the same weaknesses as SGI. That is why all of DEC's high clock speeds do not translate into leading application performance.

Digital's multiple OS strategy has slowed down applications porting efforts and system sales. While most vendors support more than one operating system, they generally have one which is strategic, while the remaining operating systems may be tactical (NeXTstep for financial markets, for example). DEC supports three operating systems as strategic: OpenVMS (for commercial), OSF/1 (for technical) and NT (for low end). Not only does it require considerable investment to support 3 different operating systems, but it also is confusing to ISV's. In addition, DEC has not been able to make all operating systems available on all platforms (OSF/1 is notably missing on the SMP platforms).

DEC's server product line in the following is composed of:

- DEC 2000 AXP Models 300 and 500 at the low end
- DEC 3000 AXP Models 600S and 800S
  (Models 400S and 500S are not currently marketed)
- DEC 4000 AXP Models 610, 620, 710 and 720 (Runs OpenVMS only)
  DEC 7000 AXP Models 610, 660 (Runs OpenVMS only)
- DEC 10000 AXP Models 610-660 (Runs OpenVMS only)

3.2 Database Applications

Database management is the smallest segment of the technical UNIX server market, but one where HP offers many strengths. HP has been a leader in the UNIX Database market, both in performance and market share. HP supports all the leading database vendors, and for many, is their leading UNIX platform.

The T500 is HP's most powerful database engine. It offers performance scaling with SMP, large data storage and I/O capacity and strong networking solutions. Strong database solutions are carried throughout the line, from the E-Class to the T500. Please see the Sales Guide (pub # 5091-7225E) for more help in selling in the database market.

3.3 What's Coming Down the Pipeline

The Series 800 Server product line is extremely competitive in the market today. HP is continuing to invest in technologies important to the technical market. A sample of the work in progress follows:

1. NFS MP Scaling: allows NFS user more than one processor. Will be available in HP-UX 10.0. With MP Scaling, the T500 is capable of producing more than 4000 NFS_OPS – the highest in the industry!
2. NFS Accelerator: improves the performance of NFS on all Series 800 servers (Sun uses accelerators to achieve its NFS performance numbers today!).
3. Parallelizing Compilers: the first of HP's parallelizing compilers will allow users to automatically parallelize applications written in Fortran. It will be sold and supported from HP.

In addition to the products listed above, HP will be delivering servers which better meet the needs of the technical markets. Features like large files and file systems sizes, faster networking products, better storage management technologies, more powerful chip technologies will be developed and integrated into the product line.

3.4 Channel Partner Solutions Summary

Over 100 Channel Partners already support both the HP 9000 Series 800 and the Series 700. These channel partners are listed in the HP Solutions Catalog and most have qualified without HP's assistance, which indicates the binary compatibility of the two platforms. In technical markets, over 40 key solutions such as the SAS Institute system and ESRI are already available.

There are also new channel partners with exciting announcements. The following list is a summary of material covered in this document:

ME

- Computervision – EDM Product Data Management Software – a new port to the HP platform.
- Computervision – Design Manager – a low-end Product Data Management solution.
- PTC – Pro/PDM to be supported on the Series 800.

EDA

- Mentor Graphics now available on the Series 800.
- Cadence Design Systems now qualifying solutions.
- Cooper and Chyan Technologies to support the Series 800.

Tools

- CX-Soft Tools to support the Series 800
  - –New Clustering capabilities for Series 800s
  - –New Math library, MLIB, to boost numerical performance

The following list is a sampling which shows the wide range of technical applications available.

- Mentor Graphics – EDA
- Zuken – EDA
- Quad Design – EDA
- GenRad – EDA
- Computervision Theda – EDA
- Technology Modeling Associates – EDA
- Pacific Numerix Corp. – EDA
- HP Mechanical Design Division – ME
- Computervision EDM – ME/Product Data Management
- Computervision Design Manager – ME/Product Data Management
- SAS – Statistical Analysis
- ESRI – Global Information Systems
- Cadre – CASE
- Centerline – CASE
- Sterling NQS – Tools
- IDE – CASE
- MSC – Mechanical
- Swanson – Mechanical
- Frame – Technical Document
- Sherpa – Product Data Management
- Genasys – GIS
- ISICAD – AEC
- HP WorkManager – Product Data Management
- BBN Scientific
- Clearcase, Atria Software – CASE
- PTC Pro/PROJECT – Project Management, ME
- SAS – All SAS Software
- InTera – ECLIPSE
- Computer Power Group – AP-Today, Operating Control Systems
  RMS Technologies – Digital Terrain Modeling (DTM), METSTAR
- Egret Technologies – EQuIS/GWM, EQuIS/RDB, EQuIS/SWM Lynx
  Geosystems Inc. – Lynx 3D Component Modeling Systems
  Oceanroutes Canada LTD – Oceanroutes Fisher Controls SCADA

3.5 Where to Get More Help

This is a quick summary of the technical server program, but there are additional resources that can be of help to you. First of all, there is the Sales Center which can help with positioning and more selling strategies. The WSG Sales Center in Fort Collins will be answering questions on the technical server program and assisting in any special pricing, as will Sales Centers in your geography. The Customer Response Center can help with configuration questions. In addition, the following are available on both the Hotlines and PowerTools from LDC:

- Customer Presentation on Series 800 in the Technical Markets
- NFS Configuration Guide
- NFS Customer Presentation
- Auspex Server Evaluation
- Series 800 Price Guide (P/N 5091-9206E)
- Series 800 Server Product Brief (P/N 5091-9132E)

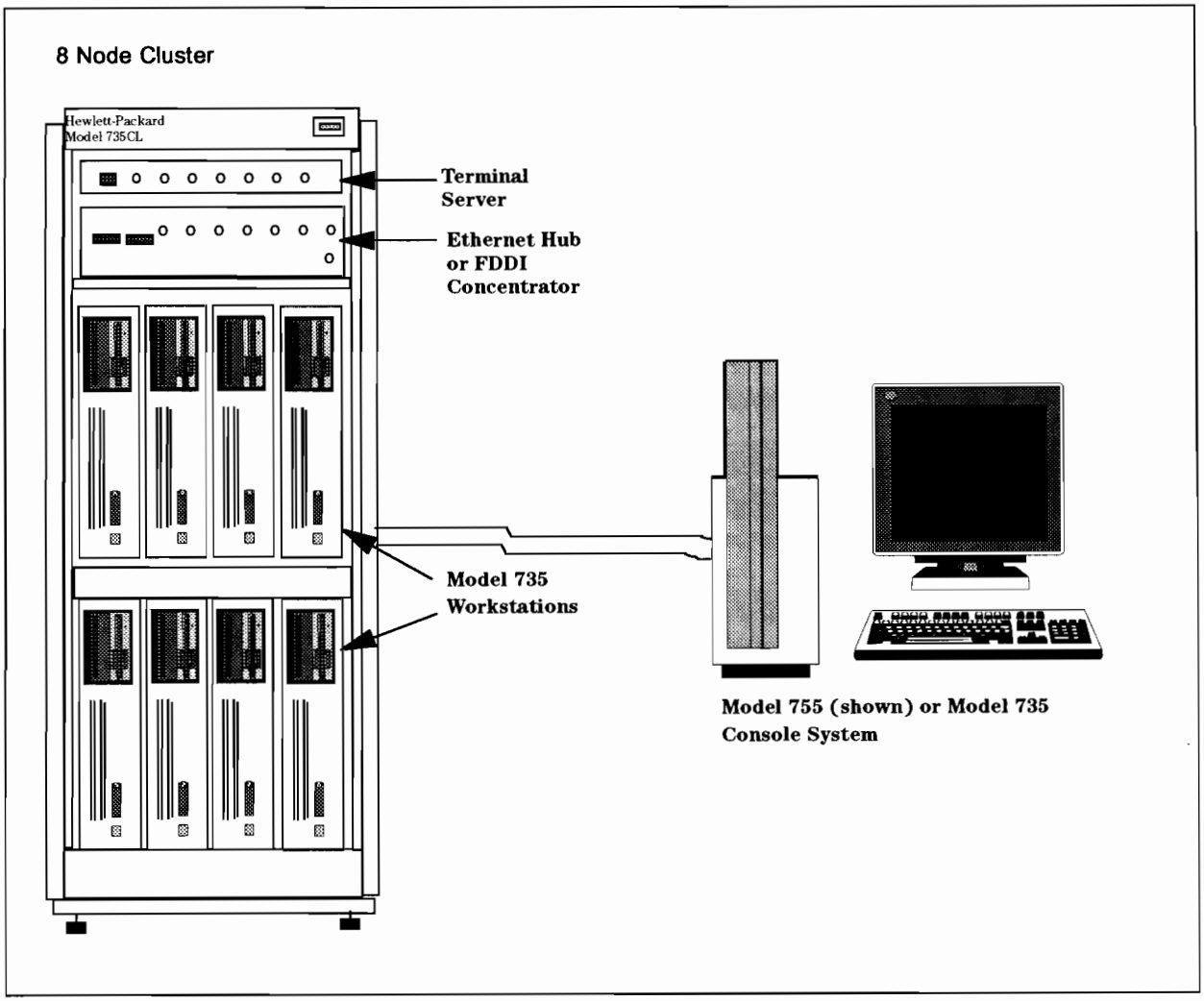
4.0 Series 700 Workstation Clusters

One-Minute Sell

Product Description

Workstation clusters are replacing traditional mainframes and supercomputers to perform batch or parallel processing. Today's Series 700 clusters are based on the Model 735 workstation and come in four factory configurations; a 4-node cluster and an 8-node cluster each available based on Ethernet or FDDI.

4.1 The Series 700 Compute Cluster – Model 735CL



The configurations of the cluster are as follows:

Model 735 Product Structure

	Model 735CL Ethernet Workstation Cluster–4 or 8 Node	Model 735CL FDDI Workstation Cluster–4 or 8 Node	Terminal Console Model 735 (Ethernet or FDDI)	Terminal Console Model 755 (Ethernet or FDDI)
Base Workstation Configuration:	• (4/8) Model 735 99MHz Headless Workstations • 80MB RAM (per workstation) • 1GB Fast-Wide SCSI-II Disk • ThinNet 802.3 Ethernet LAN • (1) EISA Slot (per work-station)	• (4/8) Model 735 99MHz Headless Workstations • 80MB RAM (per workstation) • 1GB Fast-Wide SCSI-II Disk • FDDI/9000 (per workstation) • (1) EISA Slot (per work-station)	• Model 735 • CRX Color Graphics • 32MB RAM • 1GB Fast-Wide SCSI-II Disk • ThinNet 802.3 Ethernet LAN, or FDDI/9000 LAN • (1) EISA Slot	• Model 755 • CRX Color Graphics • 64MB RAM • 2GB Fast-Wide SCSI-II Disk • ThinNet 802.3 Ethernet LAN, or FDDI/9000 LAN • (4) EISA Slots
Base Cabinet Configuration:	• (1) 1.6m, 19" Rack (includes factory wiring) • (1) Ethernet Concentrator • (1) Terminal Server • (1) 5m Ethernet Cable	• (1) 1.6m, 19" Rack (includes factory wiring) • (1) FDDI Concentrator • (1) Terminal Server • (1) 5m FDDI Cable		
Options:	• 735 Work-stations • Memory	• 735 Work-stations • Memory	• Memory • Disks • Mini-towers • EISA	• Memory • Disks • Mini-towers • EISA

A workstation connected to the cluster serves as a console to manage system administration to the cluster. While any existing Series 700 system can be used as a terminal console system, two preconfigured terminal console systems are offered. They are pre-configured Models 735 and 755 workstations to provide the performance and expandability appropriate to serve as a file server to the cluster.

A Series 800 server may also be connected to the cluster as a high-availability file server, providing back-up facilities, software, and database management to the cluster. Software residing on the server workstation manages the task distribution. The cluster may be linked to a number of different clients, including workstations and X-terminals.

The size of a cluster may vary according to the application and the processing mode (batch or parallel). The maximum number of workstations used in a batch processing cluster is unlimited, with the current installations topping 100 workstations. The optimal number of workstations in a parallel processing cluster ranges from 8 to 16. After 16 workstations, the message passing overhead between machines overcomes any additional useful work being done. The 4-workstation cluster can be incrementally expanded by adding pre-configured "cluster expansion nodes" shown on top of the next page. Once the cluster rack is full (8 workstations), more cluster racks can be daisy chained together to expand the capacity of the cluster.

4.2 Batch Processing in Clusters

When a computational cluster is used in a batch processing mode, jobs are sent to the cluster from workstations, X-terminals, and other available clients. Tasks are automatically sent to workstations within the cluster via a central queue or a distributed queue, depending on the batch queuing application used. Typical jobs sent to a queue are large compute-intensive calculations, background jobs, and other non-interactive tasks which typically require more than ten minutes processing time. Batch processing is useful for applications such as CAE, CAD, software development and scientific analysis. For example, a software developer uses batch queuing software to run compiles on a cluster while the developer's own workstation is freed up to do more interactive development. The batch software, such as Task Broker or NQS, performs functions such as task distribution and balancing the workload to maximize performance. (see Appendix B for third-party batch processing software)

4.3 Parallel Processing in Clusters

When a cluster operates in parallel processing mode, a large job sent to the cluster is divided among multiple CPUs. Parallel processing is an advanced form of clustering and requires modifications to the application software to break it up into relatively independent tasks. The customer uses software tools such as Linda or PVM to help parse the program into smaller pieces of code or information. In a parallel processing cluster, interaction between processors can be slower than in a traditional multiprocessor computer. Applications must have a high ratio of compute-intensive calculations to I/O communications to perform well in a cluster computing environment. This type of application is called a coarse-grained application. Molecular modeling, raytracing used in visualization, and seismic data analysis are examples of coarse-grained applications.

If an application requires that the calculations communicate with each other before completing processing (each calculation process-

ing on a separate CPU), the performance of the application on the cluster will be slowed down by the time required to pass information through the network. Transaction processing applications are examples of applications which would require more communications time and less processing time.

4.4 HP Distributed Computing Environment

DCE (Distributed Computing Environment) RPC (Remote Procedure Call). The RPC is just one part of the DCE core technology set. The HP DCE initial product offering is currently available.

Distributed client/server applications which incorporate RPC technology will work in a cluster environment without modification to the application. The DCE RPC should not be confused with batch or parallel processing software. In its simplest form, an RPC breaks the application into procedures which run one after the other. The difference between an RPC-based application and a parallelized application is that the parallelized application runs the procedures at the same time (in parallel) on the processors available. By combining an RPC with threads, a technique for spawning parallel work, sophisticated users may use RPCs to create parallel applications. Nevertheless, most users will prefer existing solutions like Linda or PVM.

The HP DCE product offering includes services in addition to the RPC that could provide value to the cluster environment. Timing services, security services and naming services are part of DCE and can aid in the effective use of a cluster.

4.5 The Model 735CL

The Model 735CL is HP's first fully corporate price listed workstation compute cluster system. Previous versions of HP's cluster systems were only available through the WSG Custom Products Group. The success of our Custom Product cluster systems and increased cluster activity by our competition has created growing interest in these viable alternatives for high-end compute servers. The Model 735CL cluster system combines the industry's leading price/performance Model 735 workstations into a single computational resource. These powerful computational servers opens the door to go head-on against SUN, DEC, SGI and IBM in every selling situation where price, CPU performance, scalability and investment protection are critical for an enterprise.

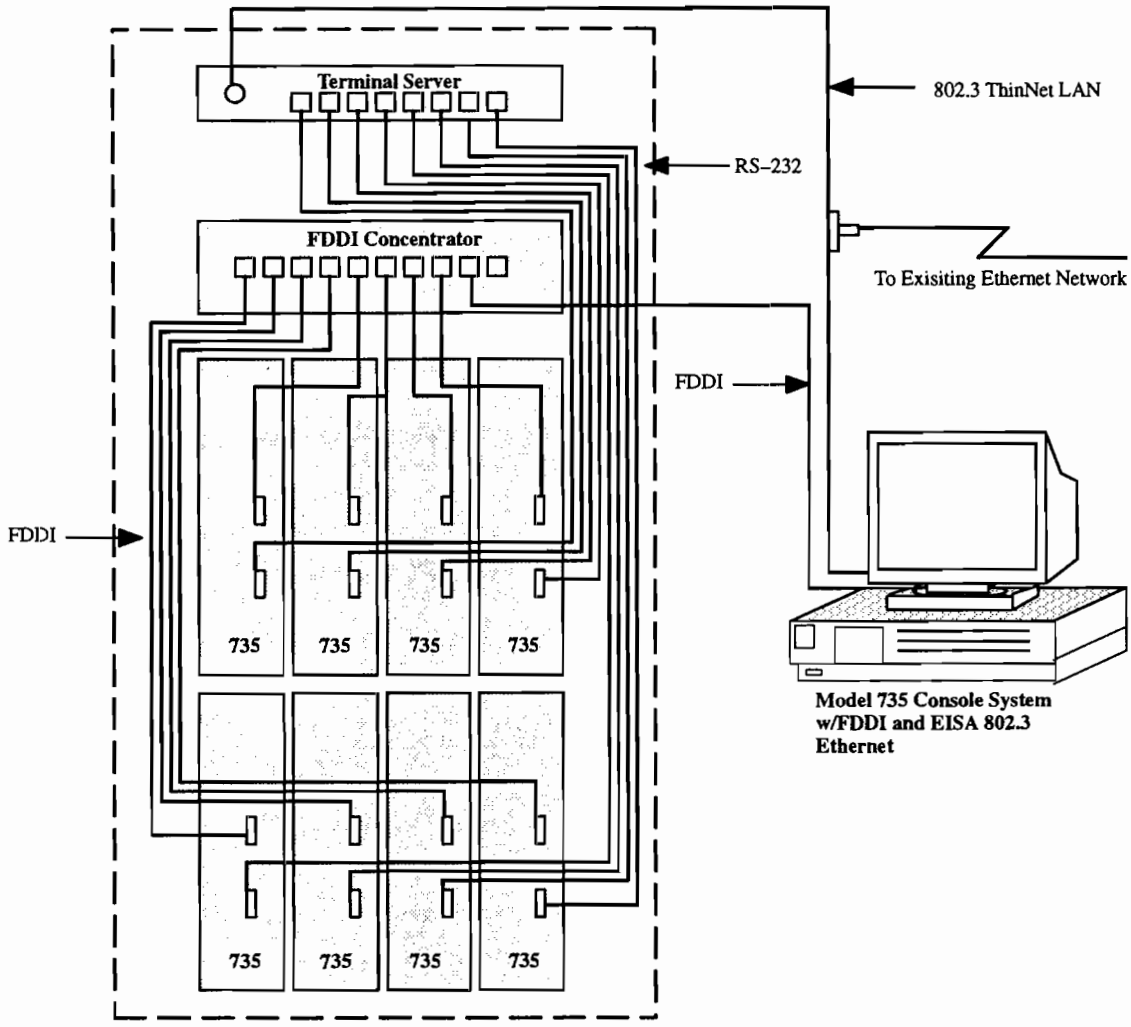
The Model 735CL is based on the 99MHz implementation of PA-RISC architecture. The 735CL systems are full binary compatible with the entire Series 700 workstation family. This product compatibility provides your customer with a complete investment protection story which includes CPU performance increases through simple system upgrades and the flexibility to deploy clustered workstations onto the desktop based on current needs.

4.5.1 Benefits Summary

Supports thousands of third party applications users buying supercomputers must write their own applications

- More affordable to purchase and operate than supercomputers
- Delivers high throughput acceleration on multiple batch jobs
- Runs sophisticated job management software such as load-balancing and batch queuing
- Supports a variety of parallelization software development tools
- Modular in design, with the ability to build a supercomputer one block at a time
- Exceptional CPU performance scalability at a fraction of the cost compared to traditional high performance systems

Typical Model 735CL FDDI Installation



Model 735CL with Attached Terminal Console System Being Used as a Network Bridge

The Model 735CL offers exceptional computational throughput because each compute processor has its own 1GB Fast-Wide disk sub-system and is configured with their own 256KB of cache and 80MB of RAM. All cluster nodes can be upgraded with additional memory and disk storage. The Model 735CL cluster systems are available with either factory installed 10MB Ethernet or 100MB FDDI interconnect.

4.6.1 Processor Performance

Model 735CL Compute Cluster Comparison Matrix

	4-node Ethernet Cluster	8-node Ethernet Cluster	4-node FDDI Cluster	8-node FDDI Cluster
# of Nodes	4	8	4	8
Processor	PA-7100	PA-7100	PA-7100	PA-7100
Clock Rate	99MHz	99MHz	99MHz	99MHz
SPECfp92	167.9	167.9	167.9	167.9
SPECint92	109.1	109.1	109.1	109.1
MFLOPS	45.4	45.4	45.4	45.4
Memory per node	80 – 400MB	80 – 400MB	80 – 400MB	80 – 400MB
Internal Disk Storage per node	2GB (max)	2GB (max)	2GB (max)	2GB (max)
Network Interconnect	802.3 ThinNet LAN (standard)	802.3 ThinNet LAN (standard)	FDDI/9000 (standard)	FDDI/9000 (standard)
Concentrator	HP ThinLAN Hub Plus 9-port, BNC	HP ThinLAN Hub Plus 9-port, BNC	Interphase FDDI/ 10-port	Interphase FDDI/ 10-port
Interconnect Transfer Rate (peak)	10MB/sec	10MB/sec	10MB/sec	10MB/sec
Terminal Server	Chase Research 10 LAN Rack, 8-port	Chase Research 10 LAN Rack, 8-port	Chase Research 10 LAN Rack, 8-port	Chase Research 10 LAN Rack, 8-port
19" 1.6m cabinet	yes	yes	yes	yes

4.6.2 Packaging

The Model 735CL is packaged as a complete system with pre-configured hardware, software and its own cabinet. The Model 735CL is available in two base configurations, a four node system, configured and priced for a low-cost entry into cluster computing and our fully configured eight node system. Each system includes a choice of either factory installed Ethernet or low latency FDDI interconnect.

Each factory base-configured cluster node includes the following:

- 80MB of RAM
- 1GB Fast-Wide SCSI II Disk
- Instant Ignited 9.x HP-UX software
- Choice of FDDI or Ethernet Networking LAN

4.6.3 System Memory

Memory increments of 144, 208 and 400 are offered from the factory. Standard Model 735 add-on memory increments of 16, 48, 112, 176 and 369 are available. It is highly recommended that all cluster nodes be configured with equal amounts of memory. From the factory, all nodes ship with 80MB's of base memory. When factory add-on memory is ordered each cluster node will be configured with the same amount of RAM. System balancing of RAM is important when configuring systems for optimal system performance. Series 735 memory has ECC single-bit error correction and multiple bit error correction logic is used.

4.6.4 System Internal Disk Storage

A second 1GB factory installed disk drive is not an available option for any of the four cluster systems. Lack of a second 1GB disk option for clusters is a factory product structuring constraint. If your customer wants to order additional internal storage for the cluster system, please quote and order the Model 735 (Field/Customer Installed) 1GB Fast-Wide disk drive.

Just as we recommend equal amounts of memory in each cluster node, we also recommend equal internal disk storage capacity. This insures each job has equal amounts of swap space available.

Remember, all factory ordered cluster nodes are configured with 1GB Fast-Wide disk. Fast-Wide and Single-Ended disks are incompatible in the Model 735. If you are ordering add-on field/customer installed storage make sure it's Fast-Wide.

4.6.5 Cabinet

The cluster cabinet is 1.6m, 19" wide and contains two factory installed shelves used for mounting the workstations. Rails are used for mounting the terminal server and the concentrator. Note: HP does not support any other devices internal to the cabinet. Both cabinet temperature and power requirements are very critical for overall system reliability. This cabinet has not been shock and vibration tested with the workstations mounted. Transport of the cabinet should be done without the workstations mounted in the cabinet.

The Model 735CL cabinet includes the following factory installed components:

- (2) shelves
- (4) 4" fans
- (2) AC distribution bus bars
- Terminal server
- Either an FDDI concentrator or Ethernet concentrator
- Either (15) 2m FDDI cables or (15) 2m Ethernet cables
- Either (1) 5m FDDI or (1) 5m Ethernet cable [terminal server to terminal console]
- Front and rear louvered doors

4.7 Receiving a Cluster

A cluster system is a more complex system than a workstation or a tightly integrated supercomputer system. However, we have bundled these systems and factory installed as many of the components as possible for a successful install.

Note: The only components that do not come factory installed in the cluster cabinet is the workstations themselves. As our cabinet strategy evolves in the near future we intend to factory install the workstation systems in the cabinet for customer shipment.

Because this constraint exists on today's factory ordered cluster systems, each quoted cluster system price includes 4 hours of HP Customer Engineer installation. Installation of the workstations by HP support representatives is not expected to be more than 1 hour. The additional installation time will include the installation of a terminal console system and basic network addressing administration. The goal of our installation program is to install and verify the integrity of the cluster system.

4.8 Site Planning and Prep

The Model 735CL system is tested and conforms to C1 open room environment standards. However, this system does have specific environmental needs such as wall space, power and networking requirements. In order to ensure each cluster sale is a success and proper customer expectations have been conveyed, each cluster system quote includes 1 hour of CE/SE site planning and prep customer consultation. There is a Model 735CL Cluster Site Planning Guide available through SRDO to Sales Centers and a list of primary site planning engineers residing in branch sales offices (P/N AL083-90602).

4.9 Terminal Console System

A cluster system requires a workstation with some form of graphics in order to administrate the networked cluster nodes. These administrative resources are called terminal console systems. To facilitate the ease of ordering and configuring, HP offers preconfigured Model 735 or 755 terminal console systems.

These systems are typical Model 735 or 755 workstations configured with standard factory memory, disk storage and CRX graphics, along with a 19" color monitor. These bundles also include factory installed networking. In addition, there is some factory software networking configuration done. The network cable needed to interface from the clusters terminal server to the terminal console system is bundled with the cluster system.

A customer can choose to use their own workstation as a terminal console system. System usage and configuration concerns should be understood prior to using or selling a terminal console system.

4.10 Selecting a Terminal Console System

Many times a terminal console system is used as a general purpose front-end file server and/or as a network router for the cluster system. Terminal console systems that offer performance and EISA expandability generally are good choices. This is why we have chosen to factory bundle the Model 735 and the 755 as terminal console systems for clusters. Both the Model 735 and 755 workstations deliver high performance and support a wide range of external disk subsystems. Other workstation models can be used as terminal consoles; however, system selection should be made based on the following criteria:

- Does the customer intend to use the terminal console system as a front-end network router for the cluster system?
- Does the customer have large batch files that must reside local in a large disk sub-system?
- Would the customer like to be able to route and support large external disk storage?

The Model 755 is an excellent choice as an administrative system and a network router. This system has 4 EISA slots for network and disk storage expandability. The Model 735 has 1 EISA slot and is limited when trying to configure it as a router along with external disk storage.

Key item to remember: if you configure a terminal console system with large disk storage subsystems and Fast-Wide disks you will ensure optimum file server performance

4.11 Adding Nodes to the Cluster

The Model 735CL is available in two base configurations, a four node system, configured and priced for a low-cost entry into cluster computing and our fully configured eight node system.

Five, six and seven node clusters can also be ordered either as a factory or field installed options. This is done by selecting one or more preconfigured cluster workstation option. Each individual cluster workstation is configured with 80MB's of RAM and 1GB of disk and comes with no keyboard or graphics. Additional system memory in 144, 208, and 400MB increments can be factory ordered. Make sure each add-on cluster workstation is configured with equal amounts of memory and Fast-Wide disk.

4.12 Daisy Chaining or Cascading Cluster Systems

It is not uncommon for a customer to request a cluster system larger than an eight node system. Support for this request should be handled by quoting multiple cluster systems and connecting them together. The Model 735CL cluster supports multiple systems daisy chained or cascaded together. A customer wanting to factory order a ten node cluster system will be unable to do so because our entry level cluster system is configured with a minimum of four nodes. In the case of a ten node system you would have to quote either an eight or twelve node system.

Our current cluster cabinet does not support two cabinets physically bolted together. The cluster systems can, however, be electrically daisy chained or cascaded together using the spare 5m cable from the second cluster system. This cable is connected between the concentrators of each cluster system.

4.13 Selecting Network Interconnect

Currently our factory bundled cluster systems are available with either FDDI or 802.3 Ethernet.

We do have plans to expand our network product offerings for future product. Price and performance are the two primary factors when selecting a network interconnect. In a workstation cluster, each individual node (CPU processor) is connected via this network interconnect. Therefore, network performance (latency) becomes very important in order to maintain a fast throughput connection between each processor.

4.14 Competition

This section describes the competitive activity in the market for computational clusters and provides suggestions for selling HP clusters against Sun, IBM, SGI, and DEC.

4.14.1 Players in the Field

IBM

IBM is HP's most aggressive competitor in the computational cluster market. For over a year, IBM customers have been networking RS/6000s together with Ethernet and FDDI, wrapping them up with third party software and calling it a cluster (we call it a carpet cluster). IBM supported this version of a cluster with its Cluster Service Offering (CSO) program.

IBM announced a cluster product in January 1993. The Scalable POWERparallel Series 9076 SP1 (SP1 for short) is basically little more than a rack mounted cluster of RS/6000 POWER workstations with a high speed interconnect. The current system is scalable from 8 to 64 processors, with a possibility of scaling up to hundreds of processors in the future. Features include:

- The high speed interconnect, called the "Vulcan Switch," will provide a 40MB/second transfer rate from processor to processor, and a 3 microsecond hardware latency (software latency, unpublished at this point, will be larger). The vulcan switch is very expensive (almost $10K per node).
- IBM offers AIX Parallel Environment Software, but this is nothing new and it is not from IBM. The AIX Parallel Environment Software are simply PVM, Express, and Network Linda – the same software as is available on an HP cluster.
- IBM also offers a queuing package called Condor. Condor was developed at the University of Wisconsin on HP platforms!

When IBM first announced its cluster offering, it supported a proprietary interconnect called the Serial Optical Channel Converter (SOCC). IBM dropped SOCC with the announcement of Fibre Channel.

The main site of IBM's cluster development is at IBM in Kingston, New York, called the Highly Parallel Supercomputing Systems Lab (HPSSL). To demonstrate cluster computing, IBM has built a cluster facility in Austin, TX. There is an additional research site in Rome, Italy. IBM also strongly promotes cluster computing at trade shows and through the press.

Several of IBM's installed sites include federally funded supercomputer labs. Over the past year, HP has been very successful breaking into the National labs which were traditionally IBM strongholds, including Sandia, Livermore, and Los Alamos.

IBM also aggressively targets research labs in universities.

Selling against IBM Clusters and the SP1

HP provides a more complete solution for clusters than IBM (for both carpet clusters and the SP1).

HP clusters use standard workstations in the racked cluster solution. IBM "re-engineers" their workstations to fit in the rack. This is more costly and delays release of the latest technology in the SP1. The processor used in the SP1, for example, was announced four months prior in a standard RS/6000 workstation. The latest HP technology is available immediately in an HP cluster.

HP workstations have a significantly more powerful processor and bigger and faster cache memory than IBM workstations.

System	HP 735CL	IBM 590 Cluster	IBM SP1	IBM 370 Cluster
Base Config.	(8) 735 PA-RISC 99MHz 80MB RAM per node 1 GB Fast-Wide Disk per node Rack System Ethernet	(4) 590 RS6000 POWER2 RISC 66 MHz 64 MB RAM per node 1 GB Disk per node	(8) 370 POWER2 RISC 62.5MHz 64MB RAM per node 1 GB Disk per node	(4) 370 RS6000 POWER RISC 62.5MHz 64MB RAM per node 1GB Disk per node
SPECint92	109.1	117	70.3	70.3
SPECfp92	167.9	242.4	121.1	121.1
MFLOPS	45.4	130	22.9	25.9

Applications that are more memory intensive—such as Monte Carlo simulations, Finite Element, and Fluid Dynamic applications—perform much better on HP workstations.

SGI

SGI is attacking the high-performance computing market with its line of SMP machines. SGI recently announced a new line of MP servers: the Power Challenge Series and the Onyx. The Power Challenge is a shared memory multiprocessing (SMP) super workstation aimed at high end compute-intensive users. The Onyx is the same box as the Power Challenge with graphic capabilities added.

In general, clusters outperform SMP machines in scalability and price/performance. Clusters may also outperform the SMP for applications that contain many independent routines. SMP machines may outperform the cluster for applications which contain many interrelated tasks.

The following is an analysis of SGI's Onyx SMP versus HP's cluster offering.

The Power Challenge/Onyx from SGI is a competitive offering vis-a-vis the HP/Convex cluster. The main advantage of SMP architecture over the HP/Convex cluster is the internal bus, which allows faster communication between processors than the current LANs (used for inter-node communication) available. The Power Challenge/Onyx has a 1GB/sec synchronous internal bus. This is very fast compared to the 100MB/sec FDDI interconnect between processors in a cluster. HP is developing new interconnects which will provide much higher bandwidths with much lower latencies.

The second advantage of the Power Challenge/Onyx is SGI's parallelizing compiler, which was developed for the Power Series SMP. A parallelizing compiler automatically parallelizes codes when running on the Onyx. This is not as efficient as code which has been explicitly parallelized using a parallel programming library such as PVM or Network Linda, but does not require the user to perform the arduous task of embedding PVM calls into the source code. Currently, there is no compiler to automatically parallelize code for a cluster.

There are many reasons why HP's Model 735CL cluster is a better choice.

The PA-RISC 7100 is a superior chip to the R4000 and the R4400 used in the Power Challenge/Onyx. To quote Dataquest, SGI is "compensating for uniprocessor performance by [using] a symmetric multiprocessor (SMP) based design." The less powerful SPU in the Power Challenge/Onyx makes it less attractive than the PA-RISC cluster for both use as a compute and a floating point engine. The PA-RISC cluster provides better floating point with fewer processors (12 Model 735s provide more MFlops than a 24 processor R4400

based system). The floating point in a 735 (167.9 SPECfp92) is almost twice that of the R4400 (121.1 SPECfp92). Even the TFP chip (for True Floating Point) due out in 1994 does not surpass the PA-RISC 7100 technology.

HP maintains a price/performance lead over the Power Challenge/ Onyx. A two processor SMP costs 25% more than a two node cluster of 735s. Scaling up Mflops for Mflops, a 12 node FDDI cluster provides the same Mflops performance as a 24 node Onyx, but at a much lower cost.

The Power Challenge/Onyx does not support RAID technology. The HP cluster does. Compute-intensive users are often moving large amounts of data to and from the cluster. The ability to store data quickly and reliably is an important factor in the choice of an architecture.

	System HP 735CL	**SGI Challenge L**	**SGI Challenge XL**
Base Configuration	(8) 735 PA-RISC 99MHz 80MB RAM per node 1GB Fast-Wide Disk per node Rack System Ethernet	8 Way SMP RS4400 150MHz 64MB RAM per node 2GB Disk per node Proprietary System Bus	12 Way SMP RS4400 150MHz 64MB RAM per node 2GB Disk per node Proprietary System Bus
SPECint92	109.1	70.3	70.3
SPECfp92	167.9	121.1	121.1
MFLOPS	45.4	22.9	25.9

4.14.2 On the Sidelines

Of the remaining competitors, none has come up with a program as strong as IBM's.

Sun

Sun Microsystems has not developed a cluster program and is placing emphasis on its multiprocessor machines. Still, most third-party software solutions have been ported to the Sparcstations due to the prevalence of Sun workstations in the market.

	System HP 735CL	SUN SS10 Model 402
Base Configuration	(8) 735 PA-RISC 99MHz 80MB RAM per node 1GB Fast-Wide Disk per node Rack System Ethernet	(4) SSS10 Model 40[2] 12 SuperSparc 40MHz 64MB RAM per node 1GB Disk per node (4) Carpet Clusters Ethernet
SPECint92	109.1	50.2
SPECfp92	167.9	60.2
MFLOPS	45.4	22.4

The SPARCcluster1 is designed to be an NFS solution and will perform poorly as a computational cluster (by Sun's own admission!).

DEC

With the announcement of the Alpha chip, DEC has announced a strategy to cluster workstations using technology from MasPar Computer Company and Intel's supercomputer division. They also announced their intention to sell low-end Crays. It does not appear that DEC has been successful in any of these ventures.

The DEC Series 800 Alpha Farm cluster currently has a narrow SPECint and SPECfp performance lead over the Model 735CL. However, their entry price is significantly higher and the Alpha architecture continues to lack application availability.

DEC also clusters the Series 600 workstations. SPECint and SPECfp are comparable to the Model 735CL. However, as with the DEC Series 800, the Model 735CL outperforms both DEC products in double precision LinPacks.

Cray will be using Alpha chips in its new MPP machine. But this system will not be compatible with Alpha workstations produced from DEC. There is no growth path from a cluster of DEC Alpha workstations to the Cray MPP.

System	HP 735CL	DEC 4000 610 AXP Farm	DEC 4000 810 AXP Farm
Base Configuration	(8) 735 PA-RISC 99MHz 80MB RAM per node 1GB Fast-Wide Disk per node Rack System Ethernet	8 DEC 600 AXP 21064 175MHz 64MB RAM per node 1GB Disk per node Rack System Ethernet	8 DEC 800 AXP 21064 175MHz 64MB RAM per node 1GB Disk per node Rack System Ethernet
SPECint92	109.1	114.1	130.2
SPECfp92	167.9	121.1	184
MFLOPS	45.4	22.9	38.9

HP's Differentiation

Despite the high level of competition from IBM, HP's cluster program maintains a competitive advantage in the market due to the following:

- HP leads in desktop package workstation performance.
- HP's position as the "hot box" has resulted in a larger selection of cluster-enabling software being ported or being developed for the Series 700.
- HP leads in price/performance. This leadership position has been maintained with the introduction of the full line of PA-7100 technology. HP workstations deliver more power for less cost. Since cost savings is the primary motivation to switch from a mainframe or supercomputer to a cluster, HP's price/performance position is certainly an advantage.
- The combination of Convex software and math/vector libraries with HP's PA-RISC architecture provides a distinct advantage over all competitors.

4.15 Markets and Opportunities for Clusters

Why Buy a Cluster?

The main reason to buy a computational cluster instead of a mainframe or supercomputer is to save money. Purchasing a cluster costs 10-100 times less than a mainframe or supercomputer. Additionally, the annual maintenance costs decrease by an equal magnitude. Performance in desktop workstations has increased at a much faster rate than in mainframes and supercomputers. Therefore, users can achieve similar levels of computing power at a fraction of the cost with the purchase of computational clusters.

Workstation clusters are also by their nature modular and flexible. In the past, a department would actually buy more supercomputing power than they needed to "grow into it" over time. Now with

workstation clusters, users can simply add more workstations to the cluster as their needs grow or swap out older systems with more powerful systems over time. And as history has shown, new workstations will not only be more powerful but also cheaper (supercomputers, by contrast, have actually grown more expensive over time).

Many people view clusters as a low-risk entry into MPP (Massively Parallel Processing) machines. People are selecting clusters over other high-end computing paradigms because they are made of general purpose workstations. Purchasing a specialized computer, such as a CM-5 from Thinking Machines, requires the user to write his own operating system, as well as application. This is a huge investment on the part of the end user, well beyond the millions invested just to purchase the hardware. And if the vendor stops supporting that specialized hardware, the end user is stuck. Workstation clusters free parallel processing advocates from the risks associated with today's MPP solutions.

Further strengthening the workstation cluster's position in the parallel computing market is Convex's ongoing effort to work with key ISV's to "parallelize" their code to run on HP workstation clusters or their new SPP-1000 family (which is compatible with the 735CL).

Who Buys Clusters?

The following bullets profile a potential computational cluster user:

- They are advanced, sophisticated early adopters
- They are often downsizing from a Cray or IBM 3090VF
- They usually develop their own software in-house
- They run many multiple batch jobs
- They run compute-intensive tasks, such as:
  - –Physics
  - –Finite element analysis
  - –Oil/gas reservoir analysis and seismic simulation
  - –Design verification
  - –Computational chemistry
  - –Financial analysis
  - –Fluid dynamics
  - –Software development

Most existing computational clusters perform batch processing with plans to migrate toward parallel processing. Parallel processing on a cluster could require higher speed lower latency networking such as FDDI. An application must lend itself towards parallel processing; for example, an application which performs multiple independent operations may be parallelized. An application that requires multiple interactions between processes may not. Additionally, applications which have a high ratio of compute-intensive calculations to I/O communications perform well in a cluster computing environment.

Example of a Cluster Customer

The customer is an international research laboratory performing four major experiments in high-energy physics. The applications they want to run on the cluster are simulations of atomic particles accelerating and colliding. The data sets needed to run these simulations can require terabytes of storage per experiment. And one set of simulations requires one CPU day on the server and produces up to 200 MB of data. The customer plans to submit jobs to the cluster via NQS (Network Batch Queuing). Eventually, the customer would like to off-load experiments from a Cray X/MP to the cluster, running simulations in parallel mode.

Advantages of Clusters over Mainframes and Supercomputers

Clusters have the following advantages over mainframes/ supercomputers:

- They are more affordable to purchase and to operate. The price of a cluster may be one tenth the price of a comparably powered mainframe or one one-hundredth the price of a supercomputer. The maintenance fees to operate a cluster are also a drastic reduction in cost. The annual maintenance costs for a Cray, for example, may be more than the purchase price for an eight node cluster.
- Workstation clusters are modular. Customers can "build" a virtual supercomputer one block at a time. Customers do not have to purchase additional processing power until they need it. And as workstation technology improves, they can buy more powerful machines at lower prices.
- Workstation clusters operate well with the desktop environment of workstations.
- Because they are built with off-the-shelf workstations, clusters support thousands of third party applications. Users buying supercomputers not only have to write their own applications, but often have to write their own operating systems!
- Because many high-performance applications are scalar (and not vectorized) a single PA-RISC based workstation can outperform a Cray. A cluster of workstations is even more effective.

Disadvantages of Clusters

Computational clusters have some disadvantages when compared to mainframes and supercomputers.

- Clusters require a lot of resources (money, people, time) on the part of the end user to configure the system, parallelize the software, etc.
- The user must have a high level of technical expertise available
- Workstation clusters have a slower interconnect between processors and less sophisticated memory systems than mainframes or supercomputers
- Workstation clusters are not a solution for all applications

Advantages of HP Workstation in Clusters

HP workstations offer the best price/performance ratios in the industry. Even with the advent of new technologies in the super-computing market, the most critical factor for high performance is the processor. The fastest processor in a cluster will outperform clusters of slower workstations (IBM's), SMP workstations with slower chips (SGI), and even supercomputers.

The availability of Convex software on the Series 700 will provide a superior computing environment to other vendors' offerings. HP's cluster program also provides preconfigured systems, support, and expertise in cluster computing.

Selling Resources

The Cluster Lab A cluster lab has been opened in Fort Collins, Colorado to demonstrate clusters. A second lab is also set up at Convex in Richardson, TX. The full suite of cluster software will be available on site so that customers may experiment to determine the best configuration for their needs. There will be on-site consulting and support services in both the Fort Collins and the Richardson, Texas labs. To set up a customer visit in the Fort Collins Cluster lab, contact Dan Nordhues at TN 229-2585.

The PSO The PSO (Professional Services Division) can provide consulting services to help customers design, implement and tune a cluster. The design phase determines the optimal mix of hardware and software in the cluster. For example, the number of applications running on the cluster will affect the hardware configuration. Implementation includes a project plan to guide the customer through installation of the system. Performance tuning, the third phase, is key to maximizing the throughput on the cluster. Performance tuning includes analyzing the network, operating system, and application to achieve maximum throughput. The PSO's services are billed out at an hourly rate, plus travel expenses.

Convex is both a VAR and an ISV for clusters. They can be a valuable resource for selling clusters. Many reps want to know when they should call Convex:

- When the customer already has a Convex C-Series on site. This is a perfect opportunity to add the processing power of a cluster to the Convex machine to make a MetaSeries. If the customer wants to replace the C-Series, he or she may still want to purchase Convex software (MLIB, ConvexPVM, ConvexNQS+).
- When the customer requires a faster network than HP's standard offering. Convex has a proprietary network solution with higher speeds and lower latencies than FDDI.
- When the customer requires consulting and systems integration (for example, if they need help downsizing from a Cray). Convex has the expertise and the resources to provide the additional services this customer needs.

Remember, with Convex's status as a VAR, you will receive credit and compensation for workstations sold through Convex. And with clusters come desktop sales.

4.16 Future Directions

HP intends to play a major role in defining and exploiting clustering technology. Our strategy is to work with both internal and external partners to strengthen our cluster software, networking and hardware product offerings. As our cluster product offerings expand, we plan to seek out new market segments for this technology.

HP currently has programs in development that will improve cluster system performance, ease of use and administration. In the area of hardware performance, our cluster program will review all CSO computer products and embrace those that are good candidates for clustering. It is our intent to remain with the price/performance curve of our desktop and standalone server products.

HP intends to deliver a 125MHz cluster product by the end of calendar Q2 as soon as hardware qualification is complete. This program will include 99MHz upgrades and new 125MHz cluster systems. If you have a customer requesting 125MHz NTE (not to exceed) price, and they meet the following NTE criteria, please contact the Sales Center.

Please note that it is probable that the 125MHz system will be made available as a desktop product first, followed by a cluster offering. In this scenario, the 125MHz workstations can be configured with cluster software and networked as "carpet clusters" (systems without a rack); however, we strongly recommend that these systems NOT be racked in an enclosed cabinet. As with any system, maintaining proper thermal characteristics is important for optimum reliability.

5.0 Overview of Convex Partnership and SPP-1000 Family

HP has had a strategic alliance with Convex Computer Company since March, 1992. In October, 1992 Convex and HP jointly announced the Convex Meta Series—a combination of Convex's C Series supercomputer and HP Workstation Clusters. Additionally, since that time, each company has sold its own version of a Series 700 workstation cluster.

5.1 The Convex SPP-1000 Product Line

On March 15, 1994, Convex introduced the SPP-1000 product line, which is the first generation of the Exemplar family of scalable parallel processing systems. The two principal models in the initial product line are the SPP-1000CD (2-16 processors) and SPP-1000XA (4-128 processors). Of significance to HP and its customers are the utilization of PA-RISC processors and the HP-UX operating system application binary interface in these Convex products. This will be the first supercomputer class (massively parallel processing – MPP) system to utilize an industry-leading commodity microprocessor and to be compatible with the application binary interface (ABI) of a widespread UNIX operating system. Convex decided their systems warranted a new method of classification other than MPP, due to their ease of use and scalability. Hence the term, scalable parallel processing, or SPP.

5.2 HP and Convex Sales Force Cooperation

Since this non-HP product has so much HP content and apparent overlap with our own products, it is important that the divisions accurately position it in the technical computing space (see section 2 for product positioning discussion). Due to the high level of compatibility between the Convex SPP-1000 Product family and the Series 700 and Series 800 systems, there may be opportunities to work with Convex Sales Reps to meet specific customer needs within your accounts.

The following is a description of the new mechanism for cooperation in the sale of SPP-1000. The goal is to encourage cooperation between the two sales forces. In this joint selling model, the HP sales force will be reference selling the SPP-1000 within the following targeted segments:

- Automotive
- University
- Aerospace
- Oil & Gas
- Government Research

Exceptions to this focused list will be few, as we are attempting to leverage the segments of greatest overlap between the two companies, and focus on industry and application segments where applications have been ported, or qualified on the SPP-1000. Feedback to the Sales Centers regarding opportunities outside these targeted segments is encouraged, as it will help in assessing the need for expansion of the joint marketing/selling relationship in the future. Remember that it is our intent to complement, not replace, our product line with the Convex solution where customer requirements dictate such; we should always strive to sell our own workstations, clusters, servers, X-terminals, etc., that bring direct HP revenue first.

5.3 Criteria for SPP-1000 Cooperation

The *minimum criteria* for a Convex SPP-1000 reference sale are presented below:

- Technical compute server application only, i.e., not a file, application, database, or multi-purpose server requirement
- Unsuccessful positioning of Series 800 servers or Model 735CL to meet the compute requirements due to performance (parallel processing support, most likely), scalability (e.g., number of CPUs, memory bandwidth, memory capacity, I/O bandwidth, etc.), ISV application availability (if not easily fixed with some market development dollars), or perhaps the system administration environment as in the case of clusters
- Planned availability of necessary ISV applications on the SPP-1000 in FY 94
- Perceived opportunity for cooperation between the HP and Convex sales centers – we will not be "forcing" our way into deals that Convex has pending in our accounts, this must either be new business, or a mutually agreed upon partnership between the SRs from each company

The process for qualification and management of these reference sale deals will be the exclusive domain of the geographic sales centers. Please avoid calling local Convex offices or headquarters independently looking for literature or partner SRs, as this will bypass the compensation management process from the sales center. While we do not want to limit interaction within existing HP/Convex field relationships, we must be able to track these deals.

5.4 Lead Qualification Process

We have developed a complete process for lead qualification and referral which will result in partnership of HP and Convex SRs for approved deals. If you determine that a Convex SPP-1000 solution may be the only means to meet a customer's requirements, call or send electronic mail to your sales center requesting a *Convex SPP-1000 Lead Qualification Form.* The Sales Center will also be able to provide you with the following additional documents:

- *Convex SPP-1000 Lead Qualification Criteria,*
- *Convex SPP-1000 Lead Qualification Process*
- Convex FY 94 ISV application plan
- Convex produced SPP-1000 data sheets and literature; any HP produced training materials on SPP-1000
- A disclaimer letter regarding the Convex products

The details for the SPP-1000 quota credit and commission plan, the mechanisms for billing field selling costs to PL5X, and specific contacts in the different Sales Centers have been released separately through normal sales management channels. Please keep in mind that while the Convex sales force is aware that HP will be referring qualified leads to them, they do not know the specifics of our compensation plan or criteria for these situations. This should at all times remain HP Confidential.

5.5 Convex Computational Clusters vs. HP Model 735CL

Convex continues to sell workstation clusters based on the HP Model 735. The relationship from the cluster standpoint is structured as a VAR relationship which insures that the HP sales reps receive a commission on HP hardware sold in their region.

Convex did offer a proprietary shared memory interconnect (SMI) for their cluster products and Meta Series but has since backed away from this in favor of using standard HP interconnects like FDDI and FibreChannel. Convex ships up to eight workstations pre-assembled in their own rack. Convex also sells, and has the ability to preload, cluster specific software for the Model 735 workstation clusters. This software is: ConvexPVM, Convex NQS+, Convex MLIB, and LSF. (See Appendix B for details.) HP, in comparison, ships system separately and assembles the cluster in the customer site. Additionally, the Convex software mentioned above is available for purchase on a "software only" basis from CXSoft, Convex's software distribution business segment. Other than the differences mentioned here, the HP's and Convex's workstation cluster products are essentially the same.

5.6 Disclaimer

It should be specifically noted that HP is currently not reselling Convex products. The SPP-1000 remains a product of Convex Computer Corporation. When presenting the Convex products in conjunction with the rest of the HP product family, the Sales Rep should be careful not to position the Convex products as HP products, or implicitly or explicitly guarantee the supply, quality, or performance in any way. Even verbal mention of any explicit or implicit guarantee exposes HP to possible expensive litigation. Therefore, you will see the following disclaimer in brochures, slides, and possibly on HP letterheads.

"The Convex SPP-1000 Series is designed, manufactured, distributed, and supported by Convex Computer Corp. HP and Convex are independent parties; HP is not responsible for the supply or performance of the SPP-1000 Series. Customers shall rely on their own evaluation when making purchasing decisions of any products from other companies."

5.7 Convex SPP-1000CD and SPP-1000XA Overview

Section 2 of this sales guide reviews may of the features of the SPP 1000 product family. For your reference the following table provides additional details on the Convex SPP-1000CD and the SPP-1000XA.

The Sales Centers will be able to provide details on how to obtain Convex product literature once the qualification process has been started. This process is started by completing the SPP-1000 lead qualification form.

SPP-1000CD	SPP-1000XA
• 2-16 Processors (up to 2 nodes)	• 4-128 Processors (up to 16 nodes)
• 256MB-4GB memory	• 256MB-32GB memory
• Low Cost Package (1 node)	• Exemplar Package (up to 2 nodes)
– up to 8 processors	– up to 16 processors
– up to 10 internal disks	– up to 20 internal disks with "hot" replacement
– external I/O cabinet in similar form factor	– external I/O cabinet in similar form factor
– upgradable to Exemplar Package	
• 4-8 way memory interleaving per node	• 8 way memory interleaving per node
• 4-16 S-bus slots	• 4-128 S-Bus slots
• CPU/Memory bandwidth:	• CPU/Memory bandwidth:
– 1.2GB/sec X-bar switch per node	–1.2GB/sec X-bar switch per node
– 1.2-2.4GB/sec SCI ring(s) between nodes	– 2.4GB/sec SCI ring(s) between nodes
• I/O Bandwidth: 250MB/sec per node	• I/O Bandwidth: 250MB/sec per node
	• Redundant Power Supplies

Appendix A

Series 700 Workstation Cluster versus Series 800 Servers

The purpose of this section is to aid sales reps and channel partners to understand the similarities and differences between HP 9000 Series 800 Servers and Series 700 Workstation Cluster. This is not designed to be a technical document.

Comparison Between the Series 700 Cluster and Series 800 Servers

The following is a comparison of the features and attributions of the Series 700 workstation cluster and the Series 800 Servers. These two product lines should be treated as one product family.

	Series 700 Workstation Cluster	Series 800 Servers	Comments
CPU	PCXT	same	
OS	HP-UX 9.0x	same (see below)	
Networking	Ethernet, FDDI	same	
Disk	FW SCSI, SE SCSI	same	
Max slots (I/O, LAN)	735 – 1 EISA, 755 – 4 EISA	E Class – 2 or 4 HP-PB G Class – 4 HP-PB H Class – 8 HP-PB I Class – 12 HP-PB T Class – 14-112 HP-PB	
High Perf Networking	FibreChannel (EISA based)	SNA gateway, FDDI, Ethernet, Token Ring	Integrated FibreChannel will be available on the next generation Series 800
Bus	EISA	HP-PB (Precision Bus)	
Maximum Memory	Model 735 – 400 MB Model 755 – 768 MB (console)	E-Class – 512MB F/G/I Class – 768MB T 500 – 2GB	Memory for Series 800 may be cheaper in high memory config's because you may use more low density memory boards.
Memory Bandwidth	264MB/sec	264 MB/sec (G/H/I 50, 60, 70) 480 MB/s (T500)	T500 currently uses 1/2 the 960 MB/s memory bandwidth
Max. Disk Capacity	735 – 126 GB, 755 – 297.5 GB	E/G Class – 144 GB H Class – 288 GB I Class – 330 GB T Class – 1900 GB	
SMP	no	2 way (Model 70) 12 way (Model T500)	
High Availability	multiple workstations – Cluster still runs if one workstation dies	Redundance through Switchover/UX Powerfail UPS Automatic Processor Deconfiguration (T500)	Mirrored Disks
Upgrade Paths	Board upgds to faster CPUs 735 box swap to 755	Board upgds with G/H/I Box swap to T500 CPU add-on inside T-500	

Issues to Consider

There are a few caveats in putting a Series 800 in the Series 700 environment. You must plan for these issues when bringing a Series 800 into a Series 700 environment.

Bus and I/O Issues

- Different bus structures: The Series 700 has an open bus with EISA slots. The Series 800 has a high-speed proprietary bus with HP-PB slots. The Series 800 may not support all peripherals used by the Series 700. For example, the EISA-based FibreChannel card available today on the Series 700 will not work in the NIO back-plane of the Series 800 machine. Different bus structures may be less of an issue with NFS File Servers.
- Fast-Wide SCSI cards and FDDI cards require two HP-PB slots each on today's Series 800 product line. Both Fast-Wide SCSI and FDDI are integrated into the Models 735 and 755, and do not use an EISA slot.

Operating System and Applications Issues

- Differences in HP-UX ABI. While the majority of the HP-UX ABI is binary compatible (>98%), there are still a few differences between each OS. This means that the customer may want to qualify his application on the Series 800. This does not require a porting effort! Most applications run on both with the same binary code. The SAS system, for example, has one binary which runs on both platforms. Mentor Graphics also qualified their most complex application (QuickSim) without modification to the code. HP-UX will be even more streamlined with HP-UX 10.0. And HP-UX 10.0 will be a smooth transition from HP-UX 9.0.
- Availability of third party applications. Many applications available on workstations are also available on Series 800 servers (check the HP Solutions Catalog). GSY strategy is to work with the ISV's (Independent Software Vendors) which are most strategic to the Workstation Group business and to continue to recruit software vendors and channel partners to support their applications on the Series 800. There are still quite a few, however, that have not yet been qualified. Please contact GSY product marketing (Adrian Albin @ T447-7808 or Diana Headrick @T436-5212) if you encounter an application needed on the 800 that is not yet available. Software pricing structures between server and desktop platforms may differ. While many products are priced the same between the Series 700 and Series 800, it is important to watch for those applications which differ.

Appendix B

Cluster Third-party Software

Note: Do not base software purchase decisions only on the information contained within this list. Contact the software vendor for features, price and availability information. This list is subject to change without notice.

This section includes the following parts:

B1.0 Software for Batch Processing Clusters;
B2.0 Parallel Pocessing Software for Parallel Processing Clusters, and;
B3.0 System Administration and Management Software.

B1.0 Software for Batch Processing Clusters

There are several software programs available on the Series 700 to manage the cluster network for batch processing mode:

B1.1 Task Broker
B1.2 NQS
B1.3 ConvexNQS+
B1.4 Load Balancer
B1.5 LSF
B1.6 DQS

B1.1 Task Broker

Terry Graf, Hewlett-Packard Co., (508) 436-5931
Task Broker is an HP product available on all Series 700 workstations and Domain/OS systems, as well as the Series 800 servers. Task Broker is also available from a third-party vendor to run on Sun SPARCstations. And it will soon be available on IBM, SGI, and DEC Alpha systems. Task Broker maximizes the performance of the workstation cluster by efficiently distributing tasks to the best server available.

Features

- A graphic user interface (GUI) has been added, greatly improving the product's ease-of-use. The GUI provides a visual interface to most of the Task Broker command set and configuration information, simplifying Task Broker administration. In addition, task status monitoring and control is provided for the end user.
- Centralized configuration management has been added to the new release. This feature allows the entire Task Broker installation to be initialized using a single group configuration, and to be administered from any single machine site.
- An integrated forms-based configuration editor provides for easier and more robust administration of Task Broker information. The configuration syntax is simpler and checking is done during the editing session.

- An on-line, context-sensitive help sub-system, utilizing HP's CACHE CREEK product, contributes to Task Broker's overall ease-of-use by providing usage information when it's needed.
- Application code does not have to be recompiled to be run by Task Broker.
- Task Broker automatically sends results back to a user-specified file on the user's computer. Mail messages can be sent to notify the user the job is complete.

Servers and clients are added to a network equipped with Task Broker simply by hooking the machine to a LAN, updating the central Task Broker configuration file to include the new client/server, and starting the Task Broker daemon.

Task Broker provides an accounting of services used on a given server.

The number of applications running on a computer at any one time can be limited. This prevents degradation of server performance.

Users can control the use of their computer as a server. For example, users can specify that their workstations be accessed only at night.

Task Broker Benefits

In situations where jobs need to be executed on remote computers, Task Broker offers more flexibility and power than existing network utilities, such as those for remote copying of files (ftp) or remote shell execution (remsh). Moreover, Task Broker enables you to define what services are to be provided on which computers, how and when each service will be provided, as well as who can access these services. Task Broker has the following user benefits:

- More efficient access to compute resources (i.e., load balancing)—Task Broker assigns jobs to the server node, whether a specialized server platform or unused workstation, whichever is most appropriate at the time. For example, a high-end compute server is obviously better for solving compute-intensive simulations such as finite element analysis. However, if this server is overloaded and a workstation that could do an adequate job is left idle, the overall productivity of the computer environment suffers.
- Improved productivity and products—by allowing multiple compute jobs to run in parallel and/or on faster systems, performance is dramatically improved. Users may use the time saved to run more jobs in a given time period in order to improve product design or shorten development time.
- Greater flexibility and ease of accessing multiple servers—the fact that all users get access to each other's idle cycles makes for a winning proposition for the entire work group.
- Access to heterogeneous computers—you can build a cooperative computing environment. For example, a task submitted from an HP 9000 Series 700 workstation may be serviced by a Series 300 or 400 workstation, or vice versa.

A new release of TaskBroker (Version 1.2) will be available in June, 1994.

The process by which a task is distributed to the cluster is as follows:

1. Once installed, Task Broker creates "daemons," which run in the background waiting for service requests.
2. Tasks (service requests) are submitted from the user's machine.
3. The "daemon" on the user's machine queries the "daemons" on other workstations in the cluster about their ability to perform the task. The criteria for selecting a server include: the type of application, the time of day, the number of services currently using the server, the CPU load, and the available disk space.
4. The daemon on the user's workstation selects the workstation that is both available and best suited to perform the task.
5. If all servers are busy, the task waits in a queue until a server is free.
6. Files needed to complete the task may be transferred to the server by Task Broker or accessed remotely via NFS.
7. The task is completed on the selected workstation, and the results are sent back to the originating workstation.

Task Broker resides on each client and server workstation in the cluster. The version of Task Broker must be appropriate for that model server. There is no centralized server which controls the brokering and queuing of tasks. For more information on Task Broker, refer to the Technical Data Sheet, P/N 5952-2264, and the Application Note 428, P/N 5952-3041. Please see the pricing and configuration guide for the price of a Task Broker license.

B1.2 NQS and NQS/Exec

Thomasine Bailey, Sterling Software, (415) 964-9900

NQS (Network Queuing Software) is a UNIX-based product available from Sterling Software which facilitates batch processing on a computational cluster. NQS is a batch-queuing program which allows the user to submit a job to the cluster via a system of queues. The system administrator creates the queues according to the system demands and computing resources available. NQS/Exec adds automatic load leveling to the NQS product. With NQS and NQS/Exec, when the user submits a job to the queue, the software automatically selects the best processor available to complete the task. NQS will not move the data over to the workstation automatically unless the file system has been previously mounted via NFS. Additional features of NQS and NQS/Exec include:

- POSIX batch-queuing standard
- Interoperability with multiple platforms
- Ability to submit a job to a selected queue and track the progress of the task
- The ability to monitor server use

B1.3 ConvexNQS+

Jon Gelsey/CXSoft, (214) 497-4000
ConvexNQS+ is based on NASA's Network Queuing System, the de facto industry standard. It permits jobs to run on the most appropriate processor. The enhancements include demand queuing for load balancing, file importing, direct remote submission, removal of jobs, and extensive documentation.

B1.4 Load Balancer

Dan Freedman, Freedman Sharp and Associates, 403-251-2729
Load Balancer is a batch-queuing product from Freedman Sharpe and Associates which operates in heterogeneous networked environments. Load Balancer is available on both the Series 700 and 800, as well as on Sun, DEC, IBM, and SGI workstations.

Load Balancer offers a central configuration file for simple system administration and maintenance. In addition, Load Balancer allows the user to customize the configuration down to the swap space needed by an application on a specific workstation. Other key features in Load Balancer include:

- Automatically queues and distributes jobs
- 256 job priority levels
- Queues, pauses, restarts, or defers batch jobs
- Understands fixed and floating licenses, access control node availability schedules
- Tracks server usage and performance.

B1.5 LSF

Ms. Bing Wu, Platform Computing Corporation, 416-978-0458; e-mail: info@platform.com
Platform Computing Co. is a UNIX systems software company dedicated to open distributed computing by providing integrated solutions to systems downsizing. Platform's customers span a wide variety of industries, such as software development, electronics, aerospace, automobile, petroleum, pharmaceutical, military and civilian research, consumer products, and education.

Platform's LSF (Load Sharing Facility) is a distributed computing system that supports fully transparent load sharing across heterogeneous UNIX systems. LSF, when deciding where to run a job, automatically takes into consideration the architecture, the operating system, and the amount of resources required by the job, such as memory, disk space, and software licenses. The system is highly fault tolerant—services are available as long as one host is up, and no pending job is lost even if all hosts are down. A comprehensive job accounting facility is available to allow management to "keep pulse"

of the system. LSF supports all types of applications—parallel and serial, submitted either interactively or in batch mode. LSF has an open system architecture with a network API supporting a wide variety of distributed applications and software packages, including:

Isbatch: A distributed batch-queuing system with powerful yet simple single-host-like batch command interface, configurable job queues, system-wide load balancing, strong resource sharing control, partial NQS compatibility, and a GUI interface.

Ismake: Parallel make fully compatible with GNU make.

Istcsh: Load sharing UNIX command shell fully compatible with tsch.

IsPVM: Load sharing PVM for parallel computing fully compatible with PVM.

Islogin: Load sharing login to start a user session on the best host.

Istools: Load sharing tool kit for custom load sharing applications built as shell scripts.

B1.6 DQS

Supercomputer Computations Research Institute
Florida State University
UNIX System Manager
400 Dirac Science Library B-186
Tallahassee, Florida 32306-4052
(904) 644-0190

The Florida State University's Supercomputer Research Institute (SCRI) is recognized as one of the finest environments for computationally oriented research in the world. Authorized by Congress in 1984, SCRI was the first university-based industrial/government partnership in basic scientific research using high-performance computers. The Institute serves a multidisciplinary group of scientists and faculty whose research ranges from simulating the inner workings of the proton to modeling lightning in thunderstorms. SCRI maintains three high-performance computer systems: a Thinking Machines CM-2 Connection Machine, a Cray Y-MP4/432, and a large cluster of super workstations.

SCRI is a leading center for research in the development of heterogeneous parallel processing. Presently, researchers at SCRI are developing the next generation of software that will integrate these different computer systems to address more difficult and complex problems that lay beyond the scope of a single system.

Distributed Queuing System (DQS) supports queuing of single machine jobs, multinode parallel jobs, PVM parallel processing jobs, and interactive sessions across a heterogeneous networked cluster of UNIX and OSF/1 workstations. DQS will automatically load balance the submitted work across available machines and support a

variety of machine resource requests. DQS supports a general hierarchical structure of queues that can span architectures, or groups of machines within an architecture. The user interface to DQS is from either the command line or through a menu- and icon-driven Motif graphical user interface (GUI).

Features include:

- Targeted at "cluster environments"
- Simple to install
- Support for batch, interactive, and multinode jobs
- Hierarchical management facilities
- Load balancing
- PVM directly supported
- Virtual Queue support allowing submission to a "group of queues" rather than a "specific queue"

DQS is distributed freely with full source and documentation via anonymous ftp from ftp.scri.fsu.edu in the file pub/DQS/DQS.tar.Z, or by contacting the UNIX system manager listed above.

B2.0 Parallel Processing Software for Parallel Processing Clusters

There are several software programs available which enable parallel processing on a distributed network. The software programs available on the HP 9000 Series 700 workstations are:

B2.1 ConvexPVM
B2.2 ConvexMLIB
B2.3 Express
B2.4 Forge 90
B2.5 ISIS
B2.6 Network Linda
B2.7 PVM
B2.8 Tuplex
B2.9 NetShare SDK and NetMake
B2.10 Fortran M

B2.1 Convex PVM

John Gelsey, CXSoft, (214) 497-4000
ConvexPVM, Parallel Virtual machine, enables software developers to thread applications and take advantage of the power of parallel processing. Many scientific applications consist of processes that differ significantly from each other. Developers, using ConvexPVM, can run these processes on the most efficient platform depending on their parallel, vector and scalar characteristics. A biochemical research organization added ConvexPVM and enabled a large number of HP 9000 systems to run jobs in parallel, thereby reducing run-times significantly.

B2.2 ConvexMLIB

Jon Gelsey, CXSoft, (214) 497-4000
ConvexMLIB mathematical library provides a set of subroutines to accomplish most of the commonly occurring tasks in scientific programming. The subroutines are highly optimized for use on HP's PA-RISC architecture. ConvexMLIB delivers dramatically improved performance by using advanced algorithms explicitly chosen to match HP's PA-RISC architecture and by coding these algorithms in PA-RISC assembly language. This set of subprograms enables developers to maximize their productivity by providing a set of often used subroutines that dramatically decrease run times. An astrophysicist cut his computational run time from two days to a few hours by adding ConvexMLIB to a Series 700 workstation application.

B2.3 Express

ParaSoft Corporation, (818) 792-9941
Express is a programming environment marketed by Parasoft. Express includes a five-phase procedure in the conversion of a sequential application to parallel application. The phases are:

1. Project evaluation
2. Code development
3. Debugging and testing
4. Performance evaluation
5. Maintenance and upgrades

The code-development phase of Express contains a tool called ASPAR, which aids the user in converting sequential applications code to parallel code. Express is available for applications written in C or Fortran. Express runs in heterogeneous (multivendor) environments.

B2.4 FORGE 90

Bob Enk, Applied Parallel Research, Inc., (301) 718-3733
FORGE 90, from Applied Parallel Research, Inc., is an interactive analysis environment for the exploration and parallelization of industrial strength Fortran programs. FORGE 90 is a series of layered products composed of the following:

- Baseline FORGE 90 is a an interprocedural Fortran program browser/analyser, providing tracing and in-context search functions across subprograms and through calls and COMMON. FORGE 90 locates undefined and useless variables, scrutinizes COMMON block usages, displays the relationships between routines, blocks, and statements (including data dependencies), and identifies the CPU intensive parts of a program down to the DO loop level. This tool is instrumental in providing information necessary for parallelization and optimization.

- Parallelizer for Distributed Memory Systems (DMP) is an interactive tool used with Baseline FORGE 90 to parallelize Fortran applications. DMP works with the information gathered by baseline FORGE 90 to assist the user in parallelizing high level loops (to optimize efficiency). The result is a Fortran 77 SPMD (Single Program, Multiple Data) program with parallelized DO loops and partitioned data arrays spread across processors, and with communications and synchronization calls to APR's parallel runtime library PVM, Express, or Linda interface. DMP also contains a parallel runtime profiler, which identifies performance bottlenecks and communications overloads, and predicts performance on scalable systems based on single node performance.
- xHPF and xHPF77 are batch command tools that convert programs with High Performance Fortran (HPF) Subset Data Decomposition Derivatives and Fortran 90 Array syntax or Fortran 77 DO loops into SPMD parallelized DO loops with calls to APR's runtime library.

FORGE 90 is available on most UNIX workstation platforms.

B2.5 ISIS

Richard Moran, ISIS Distributed Systems, Inc., (212) 979-7729

ISIS is a product line of distributed computing tools developed by ISIS Distributed Systems, Inc. The main product is the ISIS Distributed Toolkit, which provides a software programming environment to develop distributed parallelized applications in C, C++, Fortran, Common LISP, and Ada. ISIS offers software fault tolerance to the distributed computing environment.

Additional products which can be layered on to the Toolkit are:

- ISIS Distributed Resource Manager: This product is used to create a fault-tolerant "supercomputer" out of a network of workstations. The ISIS Distributed Resource Manager actually enables load-balanced batch processing in a distributed computing environment.
- ISIS Distributed News: Designed for banking and brokerage applications, this product is a fault-tolerant message publication-subscription system that serves as a communications backplane for applications requiring subject-based broadcasting.
- ISIS Reliable Network File System: This product is a fault-tolerant network file system based on NFS protocol. Completely compatible with NFS, the ISIS RFNS product requires no changes to the applications software, or the NFS servers, and no additional hardware.
- ISIS Distributed Sensors: This product is a system administrative tool designed to monitor server performance and availability.

Isis is available on the major workstation platforms.

B2.6 Network Linda

Leigh Cagan, Scientific Computing Associates, (203) 777-7442
Network Linda is a parallel processing language sold by Scientific Computing Associates (SCA). Network Linda contains six commands that are inserted into applications programs (available for C, C++, and Fortran) to run independent tasks simultaneously. SCA markets Network Linda to customers with high-level computing needs who have developed their own applications programs. Examples of these customers in the commercial market include the oil, gas, and finance industries.

Network Linda operates on a "Master/Worker" paradigm. The "Master" breaks the application into tasks via Linda commands embedded into the source code. When the program runs, the "worker" grabs a task, completes it, returns it to the master, and grabs another task. Tasks are contained in a unit called a "Tuple," and several tuples together are called a "Tuple Space."

Tuples may also contain data or results. Network Linda coordinates the interaction between Tuples so that the parallel operation is invisible to the end user. The only visible difference is that the application runs several times faster.

Features in Network Linda include:

- Language/architecture independence
- Linda compilers
- Parallel runtime libraries

B2.7 PVM/HeNCE

Oak Ridge National Laboratory
PVM (Parallel Virtual Machine) is a software environment coordinating parallel processing in a heterogeneous network. Developed by the Oak Ridge National Laboratory in Tennessee, PVM is public domain software; it is available by sending electronic mail to "netlib@ornl.gov" with the message "send index from pvm."
PVM achieves parallel processing via a message sending and receiving paradigm. PVM is well suited for both applications with interrelated subtasks and for traditional parallel applications (those without a great deal of interaction). The Oak Ridge Lab developed HeNCE to serve as a graphical interface tool and methodology to use PVM. PVM/HeNCE contain the tools to configure a heterogeneous network, manage multiple object modules in each component of the application, execute the application in parallel mode, debug, and monitor performance.

PVM has been used mainly to convert sequential scientific applications to parallel applications.

B2.8 Tuplex

Scott Rafer, Torque Systems, (415) 321-1200

Tuplex is a Linda-based parallel programming environment developed specifically for commercial applications. Torque works with application developers to parallelize third party applications for business imaging, spreadsheet/financial analysis, and other highly parallelizable applications. Like Network Linda, Tuplex uses the "Master/slave" paradigm to achieve the distribution of tasks across a cluster.

B2.9 NetShare SDK and NetMake

Lauren Seidl, Aggregate Computing Incorporated, (800) 966-1666, (612) 546-5579

Aggregate Computing Incorporated is a leader in the design and development of distributed computing technologies. Its products—NetShare SDK and NetMake—are designed to bridge the gap between distributed computing tools and distributed applications by providing network transparency. Aggregate Computing was founded in 1990.

Aggregate Computing provides distributed computing software tools and software products that help developers build applications that transparently distribute work across networks of machines. One of these products, NetShare SDK, combines resource management with remote execution to intelligently distribute work to the most appropriate machines throughout the network. NetShare DSK includes APIs, a set of services, and a graphical configuration tool for the central administration of services. Applications use the APIs to contact the services to find and use network-wide resources appropriately.

NetShare performs five major functions:

- Determines which computers are in the environment
- Determines the availability of the computers in real time
- Determines which resource fits the application's need
- Filters information about the available resources to select the best one for the need
- Supports remote execution

The key concept for NetShare is resource management. NetMake breaks the development task of compiling software into pieces and performs the compile in parallel across available machines on the network. The result is a five- to seven-fold decrease in compile time and a dramatic improvement in the productivity of software developers. NetMake can be thought of as the UNIX "make" utility with a NetShare backend. NetShare accepts "make" commands and coordinates the parceling out of compile jobs to available processors. This allows for the completion of compiles in parallel streams, rather than in one serial process.

For more product and sales information, contact Aggregate Computing.

B2.10 Fortran M

Ian Foster, Argonne National Laboratory,
Mathematics and Computer Science Division;
e-mail: fortran-m@mcs.anl.gov
(708) 252-4619

Argonne National Laboratory is a general program laboratory of the U.S. Department of Energy, administered by the University of Chicago. The Mathematics and Computer Science (MCS) division conducts basic and applied research in applied mathematics and computer science. In particular, it has been working in the area of parallel computing since 1984, when it established the Advanced Computing Research Facility; this has been expanded into the High-performance Computing Research Facility, with a focus on collaborative projects in computational science. MCS has a strong track record of producing high-quality research software systems that are widely adopted in the scientific and engineering communities. These include LINPACK, p4, MINPACK, OTTER, and PCN.

MCS recently developed Fortran M (FM), a preprocessor and runtime system for a small set of extensions to Fortran designed specifically for parallel programming. FM has a number of unique features that make it especially suitable for users new to parallel computing, who wish to write programs that use the power of workstation clusters:

- The FM extensions to Fortran are few, and can be learned in 3-4 hours. They use familiar Fortran syntax and concepts.
- FM is a "safe" language: programs can be guaranteed deterministic, and the FM preprocessor can detect incorrect usage.
- The FM runtime system can be tuned for the characteristics of different parallel computer systems, providing optimal performance.

For product information and access to the software, please e-mail request to Argonne National Laboratory.

B3.0 System Administration and Management Software

The management and administration of a cluster are key issues in a cluster environment. HP offers system administration and management software to aid in critical network administration practices, such as back-up, recovery, and performance management in a cluster. The tools that provide these services are:

B3.1 Network Node Manager
B3.2 PerfView
B3.3 GlancePlus
B3.4 RemoteWatch
B3.5 Omniback and Omniback/Turbo.

B3.1 Network Node Manager and PerfView

Network Node Manager, available as part of PerfView, dynamically generates and maintains a map of the network. It can be used to collect, analyze, and graph data about network layout and usage.

B3.2. PerfView

PerfView is a system-management software tool that collects data on and analyzes network performance. Using a color graphical map, PerfView displays the CPU, disk, and LAN utilization in the cluster. PerfView contains sophisticated alarm-filtering technology to identify potential problems in the cluster.

B3.3. HP GlancePlus

GlancePlus is used for online monitoring of individual systems. Each system's activity within the cluster can be displayed within a window. It can characterize system performance and resolve problems. Glance Plus monitors CPU, memory, and I/O usage by process and by user. GlancePlus can track historical trends and it can help balance disk use and tune file systems.

B3.4 RemoteWatch

RemoteWatch runs nightly, noting changes and errors in system configuration. It can send reports to designated system managers via e-mail. RemoteWatch also documents and reports configuration changes. With the customer's permission, HP service can remotely access error log files to diagnose and repair problems faster.

B3.5 Omniback and Omniback/ Turbo OmniBack

Omniback and Omniback/Turbo OmniBack are an automated global network backup management solution that provides sophisticated scheduling and journaling facilities to make centralized backup and recovery on a cluster a simpler procedure. OmniBack/Turbo is a high-speed version of Omniback used to provide backup services for users with large databases (using local raw disk).

OmniBack/Turbo is available on both the HP 9000 Series 700 workstations and the Series 800 business servers, and is especially valuable in clusters with an 800 as a high-availability file server.

For further information on HP's system and network tools, consult the HP OpenView sales kit, LDC part number 5091-4378E, and the WSADMIN presentation from the WSG Slide Hotline.

Appendix C

Pocket Guide to Multiprocessor Architectures—SMP, Clusters, & MPP

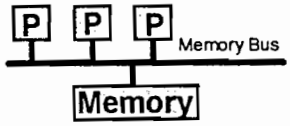
In order to scale system performance beyond the capability of a single processor, parallelism is employed by having multiple CPUs working in parallel. There are three major approaches to system-level parallelism:

- Symmetric Multiprocessors (SMP) used in some Series 800s,
- Clusters (Series 700 clusters), and
- MPP (Massively Parallel Processing) Convex's Exemplar Product Family.

The key differences are in the way resources are shared and how these systems are seen from an application side.

Symmetric Multiprocessors (SMP)

SMP is utilized today in the Series 800 G, H and I Class, Model 70 and T500 systems with scalability up to 12 processors.



In SMP, processes can be "load balanced" across multiple processors by the operating system, thus providing added performance. The primary limiting factor in SMP systems is that all processors share a common bus to main memory and thus contend for memory bandwidth.

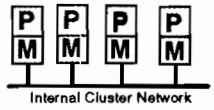
The memory bandwidth per processor is defined by the application requirements. Applications with relatively few data dependencies between processes will minimize contention for memory access. As the number of processors in an SMP system increases, the required memory bandwidth increases thereby reducing the performance improvements realized with each additional processor.

An often used analogy is trying to put out multiple fires from one fire hydrant. The fire hydrant is a shared resource able to deliver only so much water. This resource must be shared by, in this case, up to 12 fires.

The advantages of SMP system is that all memory is available to all processors. This makes load balancing and parallel work of multiple processors on the same data relatively fast. This scenario is often found in multiuser applications accessing a common database or for large parallelized technical applications that work on a data model, that must be shared.

Compute Clusters

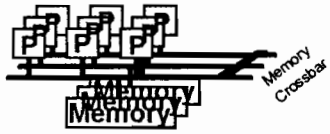
The Compute Cluster is a group of independent computers, each with its own memory, disks, and operating systems, connected through a high speed network.



Specific software is used to make the multiple computers act as a single compute server. Clusters are used in either batch mode or parallel mode. In batch mode, clusters are used to improve the throughput of multiple jobs. Queuing software distributes the incoming jobs on the cluster nodes and takes care of load-balancing, choosing the appropriate system and making the necessary data available on it. Applications do not have to be changed to utilize the cluster in this manner. Most clusters are used in this manner and are often referred to as a "throughput engine." In parallel mode, clusters are used to improve the time to completion of a single job. There are software tools available that facilitate the restructuring, or "parallelizing" of an application such that it can run on multiple nodes at once. The parallelization itself, however, has to be done by the software developer. As this is a significant investment not many applications have been parallelized for clusters. The network between the cluster systems is also much slower than the memory bus in an SMP system, so special attention has to be paid to data dependencies and interprocess communication that cause traffic on this internal network. Faster networks will improve the situation, but will still limit the communication speed compared to other multiple CPU architectures.

Massively Parallel Processors (MPP)

MPP systems are very much like SMP systems, except they allow for many more processors, usually more than 64, and they are attached to main memory using very high speed memory crossbar.



This reduces the memory bus bottleneck that is found with SMP systems, but drives system cost up significantly. MPP's are generally used to improve the time-to-completion of a single problem that can be highly parallelized. For most systems this parallelization has to be done by hand in order to achieve a high degree of scalability and performance. This is highly dependent on the application characteristics.

Convex's Exemplar is based on PA-RISC and an HP-UX ABI providing binary compatible to the existing HP-UX/PA-RISC platforms.

As described above, SMP systems, Clusters, and MPP systems, will increasingly be used in a complementary way to meet the technical server requirements of a diverse array of application types. It is Hewlett-Packard's strategy to ensure interoperability between the various multiple CPU architectures to best meet our customers' demand for increased performance. While Hewlett-Packard develops its SMP systems and Clusters, Hewlett-Packard teams with its PRO* partner Convex to develop an MPP based system based on PA-RISC technology that will provide binary compatibility with today's HP products.

HEWLETT® PACKARD

For more information, contact any of our worldwide sales offices or HP Channel Partners (In the U.S., call 1-800-637-7740; in Canada, call 1-800-387-3867).

United States:
Hewlett-Packard Company
2101 Gaither Road
Rockville, MD 20850
(301) 670 4300

Hewlett-Packard Company
5201 Tollview Drive
Rolling Meadows, IL 60008
(708) 255 9800

Hewlett-Packard Company
5805 Sepulveda Blvd., Suite 800
Van Nuys, CA 91411
(818) 786 5800

Hewlett-Packard Company
2015 South Park Place
Atlanta, GA 30339
(404) 955 1500

Canada:
Hewlett-Packard (Canada) Ltd.
5150 Spectrum Way
Mississauga, Ontario L4W 5G1
Canada
(416) 206 4725

Latin America:
Hewlett-Packard
Latin America Region Headquarters
Monte Pelvoux No. 111-2nd Floor
Lomas de Chapultepec
11000 Mexico, D.F.
(525) 202 0155

Japan:
Yokogawa-Hewlett-Packard Ltd.
Keio Fuchu 1-Chome Bldg.
1-9, Fuchu-Cho, Fuchu-shi
Tokyo 183
Japan
+81/423 30 7800

Asia Pacific:
Hewlett-Packard Asia Pacific Ltd.
17-21/F, Shell Tower
Times Square
1 Matheson Street
Causeway Bay
Hong Kong
+852/599 7777

Australia/New Zealand:
Hewlett-Packard Australia Ltd.
31-41 Joseph Street
Blackburn, Victoria 3130
Melbourne, Australia
+61 3/272 2895

European Headquarters & Multicountry Sales Region:
Hewlett-Packard S.A.
Route du Nant-d'Avril 150
CH-1217 Meyrin 2
Switzerland
+41/22/780 8111

Austria:
Tel:+43/222/2500-0

Belgium:
Tel:+32/2/778/3111

Commonwealth of Independent States (CIS):
Tel:+7/095/9235001

Czech Republic:
Tel:+42/2/4717321

Denmark:
Tel:+45/45/99/1000

Finland:
Tel:+358/0/88721

France:
Tel:+33/1/69918000

Germany:
Tel:+49/6172/16-0

Greece:
Tel:+30/1/6896411

Hungary:
Tel:+36/1/2527300

Iceland:
Tel:+354/1/671000

Ireland:
Tel:+353/1/2844633

Italy:
Tel:+39/2/92/121

Netherlands:
Tel:+31/20/5476911

Norway:
Tel:+47/67/15/97/00

Poland:
Tel:+48/22/375065

Portugal:
Tel:+351/1/3017330

Spain:
Tel:+34/1/6261600

Sweden:
Tel:+46/8/750/2000

Switzerland:
Tel:+41/57/312111

Turkey:
Tel:+90/1/224/5925

United Kingdom:
Tel:+44/344/360000

For Countries not listed:

European Headquarters & Multicountry Sales Region:
Tel:+41/22/780/8111

Middle East and Africa Operations:
Tel:+41/22/780/4111

Printed in U.S.A. 3/94
5962-8471E